CHEBYSHEV POLYNOMIALS

J.C. MASON
D.C. HANDSCOMB

CHAPMAN & HALL/CRC
A CRC Press Company
Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Mason, J.C.
Chebyshev polynomials / John C. Mason, David Handscomb.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-0355-9 (alk. paper)
1. Chebyshev polynomials. I. Handscomb, D. C. (David Christopher) II. Title..
QA404.5 .M37 2002
515′.55—dc21 2002073578

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

International Standard Book Number 0-8493-0355-9
Library of Congress Card Number 2002073578
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

In memory of recently departed friends
Geoff Hayes, Lev Brutman

Preface

Over thirty years have elapsed since the publication of Fox & Parker's 1968 text *Chebyshev Polynomials in Numerical Analysis*. This was preceded by Snyder's brief but interesting 1966 text *Chebyshev Methods in Numerical Approximation*. The only significant later publication on the subject is that by Rivlin (1974, revised and republished in 1990) — a fine exposition of the theoretical aspects of Chebyshev polynomials but mostly confined to these aspects. An up-to-date but broader treatment of Chebyshev polynomials is consequently long overdue, which we now aim to provide.

The idea that there are really four kinds of Chebyshev polynomials, not just two, has strongly affected the content of this volume. Indeed, the properties of the four kinds of polynomials lead to an extended range of results in many areas such as approximation, series expansions, interpolation, quadrature and integral equations, providing a spur to developing new methods. We do not claim the third- and fourth-kind polynomials as our own discovery, but we do claim to have followed close on the heels of Walter Gautschi in first adopting this nomenclature.

Ordinary and partial differential equations are now major fields of application for Chebyshev polynomials and, indeed, there are now far more books on 'spectral methods' — at least ten major works to our knowledge — than on Chebyshev polynomials *per se*. This makes it more difficult but less essential to discuss the full range of possible applications in this area, and here we have concentrated on some of the fundamental ideas.

We are pleased with the range of topics that we have managed to include. However, partly because each chapter concentrates on one subject area, we have inevitably left a great deal out — for instance: the updating of the Chebyshev–Padé table and Chebyshev rational approximation, Chebyshev approximation on small intervals, Faber polynomials on complex contours and Chebyshev ($\mathcal{L}_\infty$) polynomials on complex domains.

For the sake of those meeting this subject for the first time, we have included a number of problems at the end of each chapter. Some of these, in the earlier chapters in particular, are quite elementary; others are invitations to fill in the details of working that we have omitted simply for the sake of brevity; yet others are more advanced problems calling for substantial time and effort.

We have dedicated this book to the memory of two recently deceased colleagues and friends, who have influenced us in the writing of this book. Geoff Hayes wrote (with Charles Clenshaw) the major paper on fitting bivariate polynomials to data lying on a family of parallel lines. Their algorithm retains its place in numerical libraries some thirty-seven years later; it exploits the idea that Chebyshev polynomials form a well-conditioned basis independent of the spacing of data. Lev Brutman specialised in near-minimax approximations and related topics and played a significant role in the development of this field.

In conclusion, there are many to whom we owe thanks, of whom we can mention only a few. Among colleagues who helped us in various ways in the writing of this book (but should not be held responsible for it), we must name Graham Elliott, Ezio Venturino, William Smith, David Elliott, Tim Phillips and Nick Trefethen; for getting the book started and keeping it on course, Bill Morton and Elizabeth Johnston in England, Bob Stern, Jamie Sigal and others at CRC Press in the United States; for help with preparing the manuscript, Pam Moore and Andrew Crampton. We must finally thank our wives, Moya and Elizabeth, for the blind faith in which they have encouraged us to bring this work to completion, without evidence that it was ever going to get there.

This book was typeset at Oxford University Computing Laboratory, using Lamport's LaTeX 2_ε package.

John Mason
David Handscomb
April 2002

Contents

Appendices:

CHAPTER 1

Definitions

1.1 Preliminary remarks

> "Chebyshev polynomials are everywhere dense in numerical analysis."

This remark has been attributed to a number of distinguished mathematicians and numerical analysts. It may be due to Philip Davis, was certainly spoken by George Forsythe, and it is an appealing and apt remark. There is scarcely any area of numerical analysis where Chebyshev polynomials do not drop in like surprise visitors, and indeed there are now a number of subjects in which these polynomials take a significant position in modern developments — including orthogonal polynomials, polynomial approximation, numerical integration, and spectral methods for partial differential equations.

However, there is a different slant that one can give to the quotation above, namely that by studying Chebyshev polynomials one is taken on a journey which leads into all areas of numerical analysis. This has certainly been our personal experience, and it means that the Chebyshev polynomials, far from being an esoteric and narrow subject, provide the student with an opportunity for a broad and unifying introduction to many areas of numerical analysis and mathematics.

1.2 Trigonometric definitions and recurrences

There are several kinds of Chebyshev polynomials. In particular we shall introduce the first and second kind polynomials $T_n(x)$ and $U_n(x)$, as well as a pair of related (Jacobi) polynomials $V_n(x)$ and $W_n(x)$, which we call the 'Chebyshev polynomials of the third and fourth kinds'; in addition we cover the shifted polynomials $T_n^*(x)$, $U_n^*(x)$, $V_n^*(x)$ and $W_n^*(x)$. We shall, however, only make a passing reference to 'Chebyshev's polynomial of a discrete variable', referred to for example in Erdélyi et al. (1953, Section 10.23), since this last polynomial has somewhat different properties from the polynomials on which our main discussion is based.

Some books and many articles use the expression 'Chebyshev polynomial' to refer exclusively to the Chebyshev polynomial $T_n(x)$ of the first kind. Indeed this is by far the most important of the Chebyshev polynomials and, when no other qualification is given, the reader should assume that we too are referring to this polynomial.

Clearly some definition of Chebyshev polynomials is needed right away and, as we shall see as the book progresses, we are spoiled for a choice of definitions. However, what gives the various polynomials their power and relevance is their close relationship with the trigonometric functions 'cosine' and 'sine'. We are all aware of the power of these functions and of their appearance in the description of all kinds of natural phenomena, and this must surely be the key to the versatility of the Chebyshev polynomials. We therefore use as our primary definitions these trigonometric relationships.

1.2.1 The first-kind polynomial T_n

Definition 1.1 *The Chebyshev polynomial $T_n(x)$ of the first kind is a polynomial in x of degree n, defined by the relation*

$$T_n(x) = \cos n\theta \quad \text{when } x = \cos\theta. \tag{1.1}$$

If the range of the variable x is the interval $[-1, 1]$, then the range of the corresponding variable θ can be taken as $[0, \pi]$. These ranges are traversed in opposite directions, since $x = -1$ corresponds to $\theta = \pi$ and $x = 1$ corresponds to $\theta = 0$.

It is well known (as a consequence of de Moivre's Theorem) that $\cos n\theta$ is a polynomial of degree n in $\cos\theta$, and indeed we are familiar with the elementary formulae

$$\cos 0\theta = 1, \quad \cos 1\theta = \cos\theta, \quad \cos 2\theta = 2\cos^2\theta - 1,$$
$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \quad \cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1, \quad \ldots .$$

We may immediately deduce from (1.1), that the first few Chebyshev polynomials are

$$\begin{gathered} T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \\ T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \quad \ldots . \end{gathered} \tag{1.2}$$

Coefficients of all polynomials $T_n(x)$ up to degree $n = 21$ will be found in Tables C.2a, C.2b in Appendix C.

In practice it is neither convenient nor efficient to work out each $T_n(x)$ from first principles. Rather by combining the trigonometric identity

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

with Definition 1.1, we obtain the fundamental recurrence relation

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n = 2, 3, \ldots, \tag{1.3a}$$

which together with the initial conditions

$$T_0(x) = 1, \quad T_1(x) = x \tag{1.3b}$$

recursively generates all the polynomials $\{T_n(x)\}$ very efficiently.

It is easy to deduce from (1.3) that the leading coefficient (that of x^n) in $T_n(x)$ for $n > 1$ is double the leading coefficient in $T_{n-1}(x)$ and hence, by induction, is 2^{n-1}.

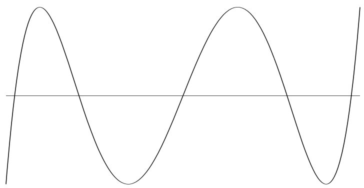


Figure 1.1: $T_5(x)$ on range $[-1, 1]$

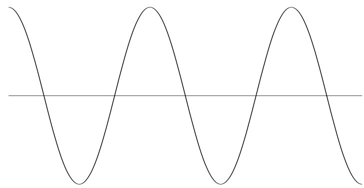


Figure 1.2: $\cos 5\theta$ on range $[0, \pi]$

What does the polynomial $T_n(x)$ look like, and how does a graph in the variable x compare with a graph of $\cos n\theta$ in the variable θ? In Figures 1.1 and 1.2 we show the respective graphs of $T_5(x)$ and $\cos 5\theta$. It will be noted that the shape of $T_5(x)$ on $[-1, 1]$ is very similar to that of $\cos 5\theta$ on $[0, \pi]$, and in particular both oscillate between six extrema of equal magnitudes (unity) and alternating signs. However, there are three key differences — firstly the polynomial $T_5(x)$ corresponds to $\cos 5\theta$ *reversed* (i.e., starting with a value of -1 and finishing with a value of $+1$); secondly the extrema of $T_5(x)$ at the end points $x = \pm 1$ do not correspond to zero gradients (as they do for $\cos 5\theta$) but rather to rapid changes in the polynomial as a function of x; and thirdly the zeros and extrema of $T_5(x)$ are clustered towards the end points ± 1, whereas the zeros and extrema of $\cos 5\theta$ are equally spaced.

The reader will recall that an *even* function $f(x)$ is one for which

$$f(x) = f(-x) \text{ for all } x$$

and an *odd* function $f(x)$ is one for which

$$f(x) = -f(-x) \text{ for all } x.$$

All even powers of x are even functions, and all odd powers of x are odd functions. Equations (1.2) suggest that $T_n(x)$ is an even or odd function, involving only even or odd powers of x, according as n is even or odd. This may be deduced rigorously from (1.3a) by induction, the cases $n = 0$ and $n = 1$ being supplied by the initial conditions (1.3b).

1.2.2 The second-kind polynomial U_n

Definition 1.2 *The Chebyshev polynomial $U_n(x)$ of the second kind is a polynomial of degree n in x defined by*

$$U_n(x) = \sin(n+1)\theta / \sin\theta \quad \text{when } x = \cos\theta. \tag{1.4}$$

The ranges of x and θ are the same as for $T_n(x)$.

Elementary formulae give

$$\sin 1\theta = \sin\theta, \quad \sin 2\theta = 2\sin\theta\,\cos\theta, \quad \sin 3\theta = \sin\theta\,(4\cos^2\theta - 1),$$
$$\sin 4\theta = \sin\theta\,(8\cos^3\theta - 4\cos\theta), \quad \ldots,$$

so that we see that the ratio of sine functions (1.4) is indeed a polynomial in $\cos\theta$, and we may immediately deduce that

$$U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1,$$
$$U_3(x) = 8x^3 - 4x, \quad \ldots. \tag{1.5}$$

Coefficients of all polynomials $U_n(x)$ up to degree $n = 21$ will be found in Tables C.3a, C.3b in Appendix C.

By combining the trigonometric identity

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

with Definition 1.2, we find that $U_n(x)$ satisfies the recurrence relation

$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x), \quad n = 2, 3, \ldots, \tag{1.6a}$$

which together with the initial conditions

$$U_0(x) = 1, \quad U_1(x) = 2x \tag{1.6b}$$

provides an efficient procedure for generating the polynomials.

A similar trigonometric identity

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

leads us to a relationship

$$U_n(x) - U_{n-2}(x) = 2T_n(x), \quad n = 2, 3, \ldots, \tag{1.7}$$

between the polynomials of the first and second kinds.

It is easy to deduce from (1.6) that the leading coefficient of x^n in $U_n(x)$ is 2^n.

Note that the recurrence (1.6a) for $\{U_n(x)\}$ is identical in form to the recurrence (1.3a) for $\{T_n(x)\}$. The different initial conditions [(1.6b) and (1.3b)] yield the different polynomial systems.

In Figure 1.3 we show the graph of $U_5(x)$. It oscillates between six extrema, as does $T_5(x)$ in Figure 1.1, but in the present case the extrema have magnitudes which are not equal, but increase monotonically from the centre towards the ends of the range.

From (1.5) it is clear that the second-kind polynomial $U_n(x)$, like the first, is an even or odd function, involving only even or odd powers of x, according as n is even or odd.

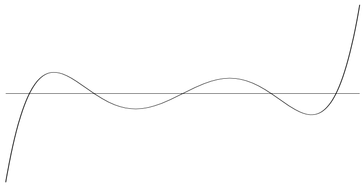


Figure 1.3: $U_5(x)$ on range $[-1, 1]$

1.2.3 The third- and fourth-kind polynomials V_n and W_n (the airfoil polynomials)

Two other families of polynomials V_n and W_n may be constructed, which are related to T_n and U_n, but which have trigonometric definitions involving the half angle $\theta/2$ (where $x = \cos\theta$ as before). These polynomials are sometimes[1] referred to as the 'airfoil polynomials', but Gautschi (1992) rather appropriately named them the 'third- and fourth-kind Chebyshev polynomials'. First we define these polynomials trigonometrically, by a pair of relations parallel to (1.1) and (1.4) above for T_n and U_n. Again the ranges of x and θ are the same as for $T_n(x)$.

Definition 1.3 *The Chebyshev polynomials $V_n(x)$ and $W_n(x)$ of the third and fourth kinds are polynomials of degree n in x defined respectively by*

$$V_n(x) = \cos(n + \tfrac{1}{2})\theta / \cos \tfrac{1}{2}\theta \tag{1.8}$$

and

$$W_n(x) = \sin(n + \tfrac{1}{2})\theta / \sin \tfrac{1}{2}\theta, \tag{1.9}$$

when $x = \cos\theta$.

To justify these definitions, we first observe that $\cos(n + \frac{1}{2})\theta$ is an odd polynomial of degree $2n+1$ in $\cos\frac{1}{2}\theta$. Therefore the right-hand side of (1.8) is an even polynomial of degree $2n$ in $\cos\frac{1}{2}\theta$, which is equivalent to a polynomial of degree n in $\cos^2\frac{1}{2}\theta = \frac{1}{2}(1+\cos\theta)$ and hence to a polynomial of degree n in $\cos\theta$. Thus $V_n(x)$ is indeed a polynomial of degree n in x. For example

$$V_1(x) = \frac{\cos(1+\frac{1}{2})\theta}{\cos\frac{1}{2}\theta} = \frac{4\cos^3\frac{1}{2}\theta - 3\cos\frac{1}{2}\theta}{\cos\frac{1}{2}\theta} = 4\cos^2\tfrac{1}{2}\theta - 3 = 2\cos\theta - 1 = 2x - 1.$$

We may readily show that

$$\begin{gathered} V_0(x) = 1, \quad V_1(x) = 2x - 1, \quad V_2(x) = 4x^2 - 2x - 1, \\ V_3(x) = 8x^3 - 4x^2 - 4x + 1, \quad \ldots. \end{gathered} \tag{1.10}$$

[1] See for example, Fromme & Golberg (1981).

Similarly $\sin(n+\frac{1}{2})\theta$ is an odd polynomial of degree $2n+1$ in $\sin\frac{1}{2}\theta$. Therefore the right-hand side of (1.9) is an even polynomial of degree $2n$ in $\sin\frac{1}{2}\theta$, which is equivalent to a polynomial of degree n in $\sin^2\frac{1}{2}\theta = \frac{1}{2}(1-\cos\theta)$ and hence again to a polynomial of degree n in $\cos\theta$. For example

$$W_1(x) = \frac{\sin(1+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta} = \frac{3\sin\frac{1}{2}\theta - 4\sin^3\frac{1}{2}\theta}{\sin\frac{1}{2}\theta} = 3-4\sin^2\frac{1}{2}\theta = 2\cos\theta+1 = 2x+1.$$

We may readily show that

$$\begin{gathered} W_0(x) = 1, \quad W_1(x) = 2x+1, \quad W_2(x) = 4x^2+2x-1, \\ W_3(x) = 8x^3+4x^2-4x-1, \quad \ldots . \end{gathered} \tag{1.11}$$

The polynomials $V_n(x)$ and $W_n(x)$ are, in fact, rescalings of two particular Jacobi[2] polynomials $P_n^{(\alpha,\beta)}(x)$ with $\alpha = -\frac{1}{2}$, $\beta = \frac{1}{2}$ and vice versa. Explicitly

$$\binom{2n}{n} V_n(x) = 2^{2n} P_n^{(-\frac{1}{2},\frac{1}{2})}(x), \qquad \binom{2n}{n} W_n(x) = 2^{2n} P_n^{(\frac{1}{2},-\frac{1}{2})}(x).$$

Coefficients of all polynomials $V_n(x)$ and $W_n(x)$ up to degree $n = 10$ will be found in Table C.1 in Appendix C.

These polynomials too may be efficiently generated by the use of a recurrence relation. Since

$$\cos(n+\tfrac{1}{2})\theta + \cos(n-2+\tfrac{1}{2})\theta = 2\cos\theta\ \cos(n-1+\tfrac{1}{2})\theta$$

and

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

it immediately follows that

$$V_n(x) = 2xV_{n-1}(x) - V_{n-2}(x), \quad n = 2, 3, \ldots, \tag{1.12a}$$

and

$$W_n(x) = 2xW_{n-1}(x) - W_{n-2}(x), \quad n = 2, 3, \ldots, \tag{1.12b}$$

with

$$V_0(x) = 1, \quad V_1(x) = 2x-1 \tag{1.12c}$$

and

$$W_0(x) = 1, \quad W_1(x) = 2x+1. \tag{1.12d}$$

Thus $V_n(x)$ and $W_n(x)$ share precisely the same recurrence relation as $T_n(x)$ and $U_n(x)$, and their generation differs only in the prescription of the initial condition for $n = 1$.

---

[2]See Chapter 22 of Abramowitz and Stegun's *Handbook of Mathematical Functions* (1964).

It is immediately clear from (1.12) that both $V_n(x)$ and $W_n(x)$ are polynomials of degree n in x, in which all powers of x are present, and in which the leading coefficients (of x^n) are equal to 2^n.

In Figure1.4 we show graphs of $V_5(x)$ and $W_5(x)$. They are exact inverted mirror images of one another, as will be proved in the next section (1.19).

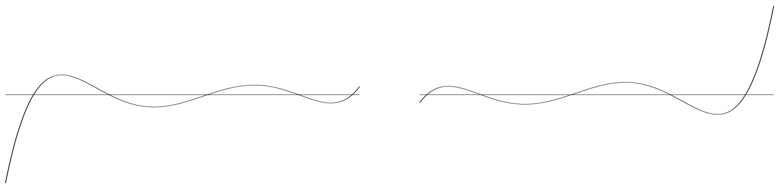


Figure 1.4: $V_5(x)$ and $W_5(x)$ on range $[-1, 1]$

1.2.4 Connections between the four kinds of polynomial

We already have a relationship (1.7) between the polynomials T_n and U_n. It remains to link V_n and W_n to T_n and U_n. This may be done by introducing two auxiliary variables

$$u = [\tfrac{1}{2}(1+x)]^{\frac{1}{2}} = \cos\tfrac{1}{2}\theta, \quad t = [\tfrac{1}{2}(1-x)]^{\frac{1}{2}} = \sin\tfrac{1}{2}\theta. \tag{1.13}$$

Using (1.8) and (1.9) it immediately follows, from the definitions (1.1) and (1.4) of T_n and U_n, that

$$T_n(x) = T_{2n}(u), \quad U_n(x) = \tfrac{1}{2}u^{-1}U_{2n+1}(u), \tag{1.14}$$

$$V_n(x) = u^{-1}T_{2n+1}(u), \quad W_n(x) = U_{2n}(u). \tag{1.15}$$

Thus $T_n(x)$, $U_n(x)$, $V_n(x)$, $W_n(x)$ together form the first- and second-kind polynomials in u, weighted by u^{-1} in the case of odd degrees. Also (1.15) shows that $V_n(x)$ and $W_n(x)$ are directly related, respectively, to the first- and second-kind Chebyshev polynomials, so that the terminology of 'Chebyshev polynomials of the third and fourth kind' is justifiable.

From the discussion above it can be seen that, if we wish to establish properties of V_n and W_n, then we have two main options: we can start from the trigonometric definitions (1.8), (1.9) or we can attempt to exploit properties of T_n and U_n by using the links (1.14)–(1.15).

Note that V_n and W_n are neither even nor odd (unlike T_n and U_n). We have seen that the leading coefficient of x^n is 2^n in both V_n and W_n, as it is in U_n. This suggests a close link with U_n. Indeed if we average the initial conditions (1.12c) and (1.12d) for V_1 and W_1 we obtain the initial condition

(1.6b) for U_1, from which we can show that the average of V_n and W_n satisfies the recurrence (1.6a) subject to (1.6b) and therefore that for all n

$$U_n(x) = \tfrac{1}{2}[V_n(x) + W_n(x)]. \tag{1.16}$$

The last result also follows directly from the trigonometric definitions (1.4), (1.8), (1.9) of U_n, V_n, W_n, since

$$\begin{aligned} \frac{\sin(n+1)\theta}{\sin\theta} &= \frac{\sin(n+\frac{1}{2})\theta\ \cos\frac{1}{2}\theta + \cos(n+\frac{1}{2})\theta\ \sin\frac{1}{2}\theta}{2\sin\frac{1}{2}\theta\ \cos\frac{1}{2}\theta} \\ &= \tfrac{1}{2}\left[\frac{\cos(n+\frac{1}{2})\theta}{\cos\frac{1}{2}\theta} + \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right]. \end{aligned}$$

Equation (1.16) is not the only link between the sets $\{V_n\}$, $\{W_n\}$ and $\{U_n\}$. Indeed, by using the trigonometric relations

$$\begin{aligned} 2\sin\tfrac{1}{2}\theta\ \cos(n+\tfrac{1}{2})\theta &= \sin(n+1)\theta - \sin n\theta, \\ 2\cos\tfrac{1}{2}\theta\ \sin(n+\tfrac{1}{2})\theta &= \sin(n+1)\theta + \sin n\theta \end{aligned}$$

and dividing through by $\sin\theta$, we can deduce that

$$V_n(x) = U_n(x) - U_{n-1}(x), \tag{1.17}$$

$$W_n(x) = U_n(x) + U_{n-1}(x). \tag{1.18}$$

Thus V_n and W_n may be very simply determined once $\{U_n\}$ are available. Note that (1.17), (1.18) are confirmed in the formulae (1.5), (1.10), (1.11) and are consistent with (1.16) above.

From the evenness/oddness of $U_n(x)$ for n even/odd, we may immediately deduce from (1.17), (1.18) that

$$\begin{aligned} W_n(x) &= V_n(-x) \qquad (n \text{ even}); \\ W_n(x) &= -V_n(-x) \quad (n \text{ odd}). \end{aligned} \tag{1.19}$$

This means that the third- and fourth-kind polynomials essentially transform into each other if the range $[-1, 1]$ of x is reversed, and it is therefore sufficient for us to study only one of these kinds of polynomial.

Two further relationships that may be derived from the definitions are

$$V_n(x) + V_{n-1}(x) = W_n(x) - W_{n-1}(x) = 2T_n(x). \tag{1.20}$$

If we were asked for a ‘pecking order’ of these four Chebyshev polynomials T_n, U_n, V_n and W_n, then we would say that T_n is clearly the most important and versatile. Moreover T_n generally leads to the simplest formulae, whereas results for the other polynomials may involve slight complications. However, all four polynomials have their role. For example, as we shall see, U_n is useful in numerical integration, while V_n and W_n can be useful in situations in which singularities occur at one end point ($+1$ or -1) but not at the other.

1.3 Shifted Chebyshev polynomials

1.3.1 The shifted polynomials T_n^*, U_n^*, V_n^*, W_n^*

Since the range $[0,1]$ is quite often more convenient to use than the range $[-1,1]$, we sometimes map the independent variable x in $[0,1]$ to the variable s in $[-1,1]$ by the transformation

$$s = 2x - 1 \text{ or } x = \tfrac{1}{2}(1+s),$$

and this leads to a shifted Chebyshev polynomial (of the first kind) $T_n^*(x)$ of degree n in x on $[0,1]$ given by

$$T_n^*(x) = T_n(s) = T_n(2x-1). \tag{1.21}$$

Thus we have the polynomials

$$\begin{gathered} T_0^*(x) = 1, \quad T_1^*(x) = 2x - 1, \quad T_2^*(x) = 8x^2 - 8x + 1, \\ T_3^*(x) = 32x^3 - 48x^2 + 18x - 1, \quad \ldots. \end{gathered} \tag{1.22}$$

From (1.21) and (1.3a), we may deduce the recurrence relation for T_n^* in the form

$$T_n^*(x) = 2(2x-1)T_{n-1}^*(x) - T_{n-2}^*(x) \tag{1.23a}$$

with initial conditions

$$T_0^*(x) = 1, \quad T_1^*(x) = 2x - 1. \tag{1.23b}$$

The polynomials $T_n^*(x)$ have a further special property, which derives from (1.1) and (1.21):

$$T_{2n}(x) = \cos 2n\theta = \cos n(2\theta) = T_n(\cos 2\theta) = T_n(2x^2 - 1) = T_n^*(x^2)$$

so that

$$T_{2n}(x) = T_n^*(x^2). \tag{1.24}$$

This property may readily be confirmed for the first few polynomials by comparing the formulae (1.2) and (1.22). Thus $T_n^*(x)$ is precisely $T_{2n}(\sqrt{x})$, a higher degree Chebyshev polynomial in the square root of the argument, and relation (1.24) gives an important link between $\{T_n\}$ and $\{T_n^*\}$ which complements the shift relationship (1.21). Because of this property, Table C.2a in Appendix C, which gives coefficients of the polynomials $T_n(x)$ up to degree $n = 20$ for even n, at the same time gives coefficients of the shifted polynomials $T_n^*(x)$ up to degree $n = 10$.

It is of course possible to define T_n^*, like T_n and U_n, directly by a trigonometric relation. Indeed, if we combine (1.1) and (1.24) we obtain

$$T_n^*(x) = \cos 2n\theta \text{ when } x = \cos^2\theta. \tag{1.25}$$

This relation might alternatively be rewritten, with θ replaced by $\phi/2$, in the form

$$T_n^*(x) = \cos n\phi \text{ when } x = \cos^2 \phi/2 = \tfrac{1}{2}(1 + \cos\phi). \tag{1.26}$$

Indeed the latter formula could be obtained directly from (1.21), by writing

$$T_n(s) = \cos n\phi \text{ when } s = \cos\phi.$$

Note that the shifted Chebyshev polynomial $T_n^*(x)$ is neither even nor odd, and indeed all powers of x from $1 = x^0$ to x^n appear in $T_n^*(x)$. The leading coefficient of x^n in $T_n^*(x)$ for $n > 0$ may be deduced from (1.23a), (1.23b) to be 2^{2n-1}.

Shifted polynomials U_n^*, V_n^*, W_n^* of the second, third and fourth kinds may be defined in precisely analogous ways:

$$U_n^*(x) = U_n(2x-1), \quad V_n^*(x) = V_n(2x-1), \quad W_n^*(x) = W_n(2x-1). \tag{1.27}$$

Links between U_n^*, V_n^*, W_n^* and the unstarred polynomials, analogous to (1.24) above, may readily be established. For example, using (1.4) and (1.27),

$$\begin{aligned}\sin\theta\, U_{2n-1}(x) &= \sin 2n\theta = \sin n(2\theta) = \sin 2\theta\, U_{n-1}(\cos 2\theta) \\ &= 2\sin\theta\cos\theta\, U_{n-1}(2x^2-1) = \sin\theta\,\{2xU_{n-1}^*(x^2)\}\end{aligned}$$

and hence

$$U_{2n-1}(x) = 2xU_{n-1}^*(x^2). \tag{1.28}$$

The corresponding relations for V_n^* and W_n^* are slightly different in that they complement (1.24) and (1.28) by involving T_{2n-1} and U_{2n}. Firstly, using (1.13), (1.15) and (1.27),

$$V_{n-1}^*(u^2) = V_{n-1}(2u^2-1) = V_{n-1}(x) = u^{-1}T_{2n-1}(u)$$

and hence (replacing u by x)

$$T_{2n-1}(x) = xV_{n-1}^*(x^2). \tag{1.29}$$

Similarly,

$$W_{n-1}^*(u^2) = W_{n-1}(2u^2-1) = W_{n-1}(x) = U_{2n}(u)$$

and hence (replacing u by x)

$$U_{2n}(x) = W_n^*(x^2). \tag{1.30}$$

Because of the relationships (1.28)–(1.30), Tables C.3b, C.2b, C.3a in Appendix C, which give coefficients of $T_n(x)$ and $U_n(x)$ up to degree $n = 20$, at the same time give the coefficients of the shifted polynomials $U_n^*(x)$, $V_n^*(x)$, $W_n^*(x)$, respectively, up to degree $n = 10$.

1.3.2 Chebyshev polynomials for the general range $[a, b]$

In the last section, the range $[-1, 1]$ was adjusted to the range $[0, 1]$ for convenience, and this corresponded to the use of the shifted Chebyshev polynomials T_n^*, U_n^*, V_n^*, W_n^* in place of T_n, U_n, V_n, W_n respectively. More generally we may define Chebyshev polynomials appropriate to any given finite range $[a, b]$ of x, by making this range correspond to the range $[-1, 1]$ of a new variable s under the linear transformation

$$s = \frac{2x - (a + b)}{b - a}. \tag{1.31}$$

The Chebyshev polynomials of the first kind appropriate to $[a, b]$ are thus $T_n(s)$, where s is given by (1.31), and similarly the second-, third- and fourth-kind polynomials appropriate to $[a, b]$ are $U_n(s)$, $V_n(s)$, and $W_n(s)$.

EXAMPLE 1.1: The first-kind Chebyshev polynomial of degree three appropriate to the range $[1, 4]$ of x is

$$T_3\left(\frac{2x-5}{3}\right) = 4\left(\frac{2x-5}{3}\right)^3 - 3\left(\frac{2x-5}{3}\right) = \frac{1}{27}(32x^3 - 240x^2 + 546x - 365).$$

Note that in the special case $[a, b] \equiv [0, 1]$, the transformation (1.31) becomes $s = 2x - 1$, and we obtain the shifted Chebyshev polynomials discussed in Section 1.3.1.

Incidentally, the 'Chebyshev Polynomials $S_n(x)$ and $C_n(x)$' tabulated by the National Bureau of Standards (NBS 1952) are no more than mappings of U_n and $2T_n$ to the range $[a, b] \equiv [-2, 2]$. Except for C_0, these polynomials all have unit leading coefficient, but this appears to be their only recommending feature for practical purposes, and they have never caught on.

1.4 Chebyshev polynomials of a complex variable

We have chosen to define the polynomials $T_n(x)$, $U_n(x)$, $V_n(x)$ and $W_n(x)$ with reference to the interval $[-1, 1]$. However, their expressions as sums of powers of x can of course be evaluated for any real x, even though the substitution $x = \cos\theta$ is not possible outside this interval.

For x in the range $[1, \infty)$, we can make the alternative substitution

$$x = \cosh\Theta, \tag{1.32}$$

with Θ in the range $[0, \infty)$, and it is easily verified that precisely the same polynomials (1.2), (1.5), (1.10) and (1.11) are generated by the relations

$$T_n(x) = \cosh n\Theta, \tag{1.33a}$$

$$U_n(x) = \frac{\sinh(n+1)\Theta}{\sinh\Theta}, \tag{1.33b}$$

$$V_n(x) = \frac{\cosh(n+\frac{1}{2})\Theta}{\cosh\frac{1}{2}\Theta}, \tag{1.33c}$$

$$W_n(x) = \frac{\sinh(n+\frac{1}{2})\Theta}{\sinh\frac{1}{2}\Theta}. \tag{1.33d}$$

For x in the range $(-\infty, -1]$ we can make use of the odd or even parity of the Chebyshev polynomials to deduce from (1.33) that, for instance,

$$T_n(x) = (-1)^n \cosh n\Theta$$

where

$$x = -\cosh\Theta.$$

It is easily shown from (1.33) that none of the four kinds of Chebyshev polynomials can have any zeros or turning points in the range $[1, \infty)$. The same applies to the range $(-\infty, -1]$. This will later become obvious, since we shall show in Section 2.2 that T_n, U_n, V_n and W_n each have n real zeros in the interval $[-1, 1]$, and a polynomial of degree n can have at most n zeros in all.

The Chebyshev polynomial $T_n(x)$ can be further extended into (or initially defined as) a polynomial $T_n(z)$ of a complex variable z. Indeed Snyder (1966) and Trefethen (2000) both start from a complex variable in developing their expositions.

1.4.1 Conformal mapping of a circle to and from an ellipse

For convenience, we consider not only the variable z but a related complex variable w such that

$$z = \tfrac{1}{2}(w + w^{-1}). \tag{1.34}$$

Then, if w moves on the circle $|w| = r$ (for $r > 1$) centred at the origin, we have

$$w = r\mathrm{e}^{\mathrm{i}\theta} = r\cos\theta + \mathrm{i}r\sin\theta, \tag{1.35}$$

$$w^{-1} = r^{-1}\mathrm{e}^{-\mathrm{i}\theta} = r^{-1}\cos\theta - \mathrm{i}r^{-1}\sin\theta, \tag{1.36}$$

and so, from (1.34),

$$z = a\cos\theta + \mathrm{i}b\sin\theta \tag{1.37}$$

where

$$a = \tfrac{1}{2}(r + r^{-1}), \quad b = \tfrac{1}{2}(r - r^{-1}). \tag{1.38}$$

Hence z moves on the standard ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{1.39}$$

centred at the origin, with major and minor semi-axes a, b given by (1.38). It is easy to verify from (1.38) that the eccentricity e of this ellipse is such that

$$ae = \sqrt{a^2 - b^2} = 1,$$

and hence the ellipse has foci at $z = \pm 1$.

In the case $r = 1$, where w moves on the unit circle, we have $b = 0$ and the ellipse collapses into the real interval $[-1, 1]$. However, z traverses the interval twice as w moves round the circle: from -1 to 1 as θ moves from $-\pi$ to 0, and from 1 to -1 as θ moves from 0 to π.

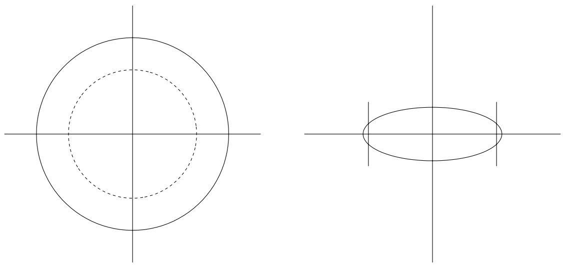


Figure 1.5: The circle $|w| = r = 1.5$ and its image in the z plane

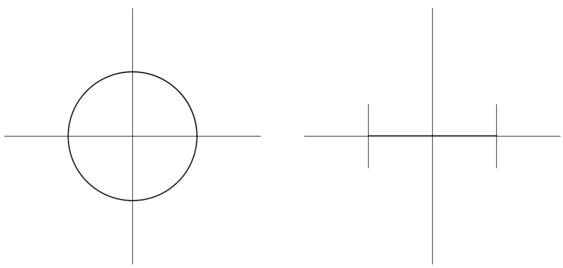


Figure 1.6: The circle $|w| = 1$ and its image in the z plane

The standard circle (1.35) and ellipse (1.39) are shown in Figure 1.5, and the special case $r = 1$ is shown in Figure 1.6. See Henrici (1974–1986) for further discussions of this mapping.

From (1.34) we readily deduce that w satisfies

$$w^2 - 2wz + 1 = 0, \tag{1.40}$$

a quadratic equation with two solutions

$$w = w_1, w_2 = z \pm \sqrt{z^2 - 1}. \tag{1.41}$$

This means that the mapping from w to z is 2 to 1, with branch points at $z = \pm 1$. It is convenient to define the complex square root $\sqrt{z^2 - 1}$ so that it lies in the same quadrant as z (except for z on the real interval $[-1, 1]$, along which the plane must be cut), and to choose the solution

$$w = w_1 = z + \sqrt{z^2 - 1}, \tag{1.42}$$

so that $|w| = |w_1| \geq 1$. Then w depends continuously on z along any path that does not intersect the interval $[-1, 1]$, and it is easy to verify that

$$w_2 = w_1^{-1} = z - \sqrt{z^2 - 1}, \tag{1.43}$$

with $|w_2| \leq 1$.

If w_1 moves on $|w_1| = r$, for $r > 1$, then w_2 moves on $|w_2| = \left|w_1^{-1}\right| = r^{-1} < 1$. Hence both of the concentric circles

$$C_r := \{w : |w| = r\}, \quad C_{1/r} := \left\{w : |w| = r^{-1}\right\}$$

transform into the same ellipse defined by (1.37) or (1.39), namely

$$E_r := \left\{z : \left|z + \sqrt{z^2 - 1}\right| = r\right\}. \tag{1.44}$$

1.4.2 Chebyshev polynomials in z

Defining z by (1.34), we note that if w lies on the unit circle $|w| = 1$ (i.e. C_1), then (1.37) gives

$$z = \cos\theta \tag{1.45}$$

and hence, from (1.42).

$$w = z + \sqrt{z^2 - 1} = \mathrm{e}^{\mathrm{i}\theta}. \tag{1.46}$$

Thus $T_n(z)$ is now a Chebyshev polynomial in a real variable and so by our standard definition (1.1), and (1.45), (1.46),

$$T_n(z) = \cos n\theta = \tfrac{1}{2}(\mathrm{e}^{\mathrm{i}n\theta} + \mathrm{e}^{-\mathrm{i}n\theta}) = \tfrac{1}{2}(w^n + w^{-n}).$$

This leads us immediately to our general definition, for all complex z, namely

$$T_n(z) = \tfrac{1}{2}(w^n + w^{-n}) \tag{1.47}$$

where

$$z = \tfrac{1}{2}(w + w^{-1}). \tag{1.48}$$

Alternatively we may write $T_n(z)$ directly in terms of z, using (1.42) and (1.43), as

$$T_n(z) = \tfrac{1}{2}\{(z + \sqrt{z^2 - 1})^n + (z - \sqrt{z^2 - 1})^n\}. \tag{1.49}$$

If z lies on the ellipse E_r, the locus of (1.48) when $|w| = r > 1$, then it follows from (1.47) that we have the inequality

$$\tfrac{1}{2}(r^n - r^{-n}) \le |T_n(z)| \le \tfrac{1}{2}(r^n + r^{-n}). \tag{1.50}$$

In Fig. 1.7 we show the level curves of the absolute value of $T_5(z)$, and it can easily be seen how these approach an elliptical shape as the value increases.

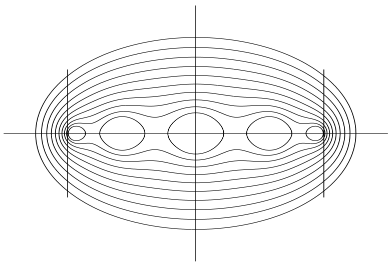


Figure 1.7: Contours of $|T_5(z)|$ in the complex plane

We may similarly extend polynomials of the second kind. If $|w| = 1$, so that $z = \cos\theta$, we have from (1.4),

$$U_{n-1}(z) = \frac{\sin n\theta}{\sin\theta}.$$

Hence, from (1.45) and (1.46), we deduce the general definition

$$U_{n-1}(z) = \frac{w^n - w^{-n}}{w - w^{-1}} \tag{1.51}$$

where again $z = \tfrac{1}{2}(w + w^{-1})$. Alternatively, writing directly in terms of z,

$$U_{n-1}(z) = \tfrac{1}{2}\frac{(z + \sqrt{z^2 - 1})^n - (z - \sqrt{z^2 - 1})^n}{\sqrt{z^2 - 1}}. \tag{1.52}$$

If z lies on the ellipse (1.44), then it follows directly from (1.51) that

$$\frac{r^n - r^{-n}}{r + r^{-1}} \le |U_{n-1}(z)| \le \frac{r^n + r^{-n}}{r - r^{-1}}; \tag{1.53}$$

however, whereas the bounds (1.50) on $|T_n(z)|$ are attained on the ellipse, the bounds (1.53) on $|U_{n-1}(z)|$ are slightly pessimistic. For a sharp upper bound, we may expand (1.51) into

$$U_{n-1}(z) = w^{n-1} + w^{n-3} + \cdots + w^{3-n} + w^{1-n} \tag{1.54}$$

giving us

$$\begin{aligned} |U_{n-1}(z)| &\le \left|w^{n-1}\right| + \left|w^{n-3}\right| + \cdots + \left|w^{3-n}\right| + \left|w^{1-n}\right| \\ &= r^{n-1} + r^{n-3} + \cdots + r^{3-n} + r^{1-n} \\ &= \frac{r^n - r^{-n}}{r - r^{-1}}, \end{aligned} \tag{1.55}$$

which lies between the two bounds given in (1.53). In Fig. 1.8 we show the level curves of the absolute value of $U_5(z)$.

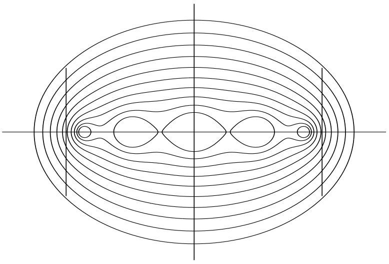


Figure 1.8: Contours of $|U_5(z)|$ in the complex plane

The third- and fourth-kind polynomials of degree n in z may readily be defined in similar fashion (compare (1.51)) by

$$V_n(z) = \frac{w^{n+\frac{1}{2}} + w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} + w^{-\frac{1}{2}}}, \tag{1.56}$$

$$W_n(z) = \frac{w^{n+\frac{1}{2}} - w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} - w^{-\frac{1}{2}}} \tag{1.57}$$

where $w^{\frac{1}{2}}$ is consistently defined from w. More precisely, to get round the ambiguities inherent in taking square roots, we may define them by

$$V_n(z) = \frac{w^{n+1} + w^{-n}}{w + 1}, \tag{1.58}$$

$$W_n(z) = \frac{w^{n+1} - w^{-n}}{w - 1} \tag{1.59}$$

It is easily shown, by dividing denominators into numerators, that these give polynomials of degree n in $z = \frac{1}{2}(w + w^{-1})$.

1.4.3 Shabat polynomials

Shabat & Voevodskii (1990) introduced the concept of 'generalised Chebyshev polynomials' (or *Shabat polynomials*), in the context of trees and number theory. The most recent survey paper in this area is that of Shabat & Zvonkin (1994). They are defined as polynomials $P(z)$ with complex coefficients having two critical values A and B such that

$$P'(z) = 0 \Longrightarrow P(z) = A \text{ or } P(z) = B.$$

The prime example of such a polynomial is $T_n(z)$, a first-kind Chebyshev polynomial, for which $A = -1$ and $B = +1$ are the critical values.

1.5 Problems for Chapter 1

1. The equation $x = \cos\theta$ defines infinitely many values of θ corresponding to a given value of x in the range $[-1, 1]$. Show that, whichever value is chosen, the values of $T_n(x)$, $U_n(x)$, $V_n(x)$ and $W_n(x)$ as defined by (1.1), (1.4), (1.8) and (1.9) remain the same.

2. Determine explicitly the Chebyshev polynomials of first and second kinds of degrees 0, 1, 2, 3, 4 appropriate to the range $[-4, 6]$ of x.

3. Prove that

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

   and that

   $$U_{m-1}(T_n(x))U_{n-1}(x) = U_{n-1}(T_m(x))U_{m-1}(x) = U_{mn-1}(x).$$

4. Verify that equations (1.33) yield the same polynomials for $x > 1$ as the trigonometric definitions of the Chebyshev polynomials give for $|x| \leq 1$.

5. Using the formula

   $$z = \tfrac{1}{2}(r + r^{-1})\cos\theta + \tfrac{1}{2}\mathrm{i}(r - r^{-1})\sin\theta, \quad (r > 1)$$

   which defines a point on an ellipse centred at 0 with foci $z = \pm 1$,

   (a) verify that

   $$\sqrt{z^2 - 1} = \tfrac{1}{2}(r - r^{-1})\cos\theta + \tfrac{1}{2}\mathrm{i}(r + r^{-1})\sin\theta$$

   and hence

   (b) verify that $\left|z + \sqrt{z^2 - 1}\right| = r$.

6. By expanding by the first row and using the standard three-term recurrence for $T_r(x)$, show that

$$T_n(x) = \begin{vmatrix} 2x & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2x & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2x & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2x & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & x \end{vmatrix} (n \times n \text{ determinant}) .$$

Write down similar expressions for $U_n(x)$, $V_n(x)$ and $W_n(x)$.

7. Given that the four kinds of Chebyshev polynomial each satisfy the same recurrence relation

$$X_n = 2xX_{n-1} - X_{n-2},$$

with $X_0 = 1$ in each case and $X_1 = x,\ 2x,\ 2x+1,\ 2x-1$ for the four respective families, use these relations *only* to establish that

(a) $V_i(x) + W_i(x) = 2U_i(x)$,
(b) $V_i(x) - W_i(x) = 2U_{i-1}(x)$,
(c) $U_i(x) - 2T_i(x) = U_{i-2}(x)$,
(d) $U_i(x) - T_i(x) = xU_{i-1}(x)$.

8. Derive the same four formulae of Problem 7, this time using only the trigonometric definitions of the Chebyshev polynomials.

9. From the last two results in Problem 7, show that

(a) $T_i(x) = xU_{i-1}(x) - U_{i-2}(x)$,
(b) $U_i(x) = 2xU_{i-1}(x) - U_{i-2}(x)$.

CHAPTER 2

Basic Properties and Formulae

2.1 Introduction

The aim of this chapter is to provide some elementary formulae for the manipulation of Chebyshev polynomials and to summarise the key properties which will be developed in the book. Areas of application will be introduced and discussed in the chapters devoted to them.

2.2 Chebyshev polynomial zeros and extrema

The Chebyshev polynomials of degree $n > 0$ of all four kinds have precisely n zeros and $n + 1$ local extrema in the interval $[-1, 1]$. In the case $n = 5$, this is evident in Figures 1.1, 1.3 and 1.4. Note that $n - 1$ of these extrema are interior to $[-1, 1]$, and are 'true' alternate maxima and minima (in the sense that the gradient vanishes), the other two extrema being at the end points ± 1 (where the gradient is non-zero).

From formula (1.1), the zeros for x in $[-1, 1]$ of $T_n(x)$ must correspond to the zeros for θ in $[0, \pi]$ of $\cos n\theta$, so that

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

Hence, the zeros of $T_n(x)$ are

$$x = x_k = \cos \frac{(k - \frac{1}{2})\pi}{n}, \quad (k = 1, 2, \ldots, n). \tag{2.1}$$

EXAMPLE 2.1: For $n = 3$, the zeros are

$$x = x_1 = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad x_2 = \cos \frac{3\pi}{6} = 0, \quad x_3 = \cos \frac{5\pi}{6} = -\frac{\sqrt{3}}{2}.$$

Note that these zeros are in decreasing order in x (corresponding to increasing θ), and it is sometimes preferable to list them in their natural order as

$$x = \cos \frac{(n - k + \frac{1}{2})\pi}{n}, \quad (k = 1, 2, \ldots, n). \tag{2.2}$$

Note, too, that $x = 0$ is a zero of $T_n(x)$ for all odd n, but not for even n, and that zeros are symmetrically placed in pairs on either side of $x = 0$.

The zeros of $U_n(x)$ (defined by (1.4)) are readily determined in a similar way from the zeros of $\sin(n+1)\theta$ as

$$x = y_k = \cos\frac{k\pi}{(n+1)}, \quad (k = 1, 2, \ldots, n) \tag{2.3}$$

or in their natural order

$$x = \cos\frac{(n-k+1)\pi}{n+1}, \quad (k = 1, 2, \ldots, n). \tag{2.4}$$

One is naturally tempted to extend the set of points (2.3) by including the further values $y_0 = 1$ and $y_{n+1} = -1$, giving the set

$$x = y_k = \cos\frac{k\pi}{(n+1)}, \quad (k = 0, 1, \ldots, n+1). \tag{2.5}$$

These are zeros not of $U_n(x)$, but of the polynomial

$$(1 - x^2)U_n(x). \tag{2.6}$$

However, we shall see that these points are popular as nodes in applications to integration.

The zeros of $V_n(x)$ and $W_n(x)$ (defined by (1.8), (1.9)) correspond to zeros of $\cos(n+\frac{1}{2})\theta$ and $\sin(n+\frac{1}{2})\theta$, respectively. Hence, the zeros of $V_n(x)$ occur at

$$x = \cos\frac{(k-\frac{1}{2})\pi}{n+\frac{1}{2}}, \quad (k = 1, 2, \ldots, n) \tag{2.7}$$

or in their natural order

$$x = \cos\frac{(n-k+\frac{1}{2})\pi}{n+\frac{1}{2}}, \quad (k = 1, 2, \ldots, n), \tag{2.8}$$

while the zeros of $W_n(x)$ occur at

$$x = \cos\frac{k\pi}{n+\frac{1}{2}}, \quad (k = 1, 2, \ldots, n) \tag{2.9}$$

or in their natural order

$$x = \cos\frac{(n-k+1)\pi}{n+\frac{1}{2}}, \quad (k = 1, 2, \ldots, n). \tag{2.10}$$

Note that there are natural extensions of these point sets, by including the value $k = n + 1$ and hence $x = -1$ in (2.7) and the value $k = 0$ and hence $x = 1$ in (2.9). Thus the polynomials

$$(1 + x)V_n(x) \text{ and } (1 - x)W_n(x)$$

have as zeros their natural sets (2.7) for $k = 1, \ldots, n+1$ and (2.9) for $k = 0, 1, \ldots, n$, respectively.

The internal extrema of $T_n(x)$ correspond to the extrema of $\cos n\theta$, namely the zeros of $\sin n\theta$, since

$$\frac{\mathrm{d}}{\mathrm{d}x}T_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\cos n\theta = \frac{\mathrm{d}}{\mathrm{d}\theta}\cos n\theta \bigg/ \frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{-n\sin n\theta}{-\sin\theta}.$$

Hence, including those at $x = \pm 1$, the extrema of $T_n(x)$ on $[-1, 1]$ are

$$x = \cos\frac{k\pi}{n}, \quad (k = 0, 1, \ldots, n) \tag{2.11}$$

or in their natural order

$$x = \cos\frac{(n-k)\pi}{n}, \quad (k = 0, 1, \ldots, n). \tag{2.12}$$

These are precisely the zeros of $(1-x^2)U_{n-1}(x)$, namely the points (2.5) above (with n replaced by $n-1$). Note that the extrema are all of equal magnitude (unity) and alternate in sign at the points (2.12) between -1 and $+1$, as shown in Figure 1.1.

The extrema of $U_n(x)$, $V_n(x)$, $W_n(x)$ are not in general as readily determined; indeed finding them involves the solution of transcendental equations. For example,

$$\frac{\mathrm{d}}{\mathrm{d}x}U_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin(n+1)\theta}{\sin\theta} = \frac{-(n+1)\sin\theta\cos(n+1)\theta + \cos\theta\sin(n+1)\theta}{\sin^3\theta}$$

and the extrema therefore correspond to values of θ satisfying the equation

$$\tan(n+1)\theta = (n+1)\tan\theta \neq 0.$$

All that we can say for certain is that the extreme values of $U_n(x)$ have magnitudes which increase monotonically as $|x|$ increases away from 0, until the largest magnitude of $n+1$ is achieved at $x = \pm 1$.

On the other hand, from the definitions (1.4), (1.8), (1.9), we can show that

$$\begin{aligned}
\sqrt{1-x^2}\,U_n(x) &= \sin(n+1)\theta, \\
\sqrt{1+x}\,V_n(x) &= \sqrt{2}\cos(n+\tfrac{1}{2})\theta, \\
\sqrt{1-x}\,W_n(x) &= \sqrt{2}\sin(n+\tfrac{1}{2})\theta;
\end{aligned}$$

Hence the extrema of the weighted polynomials $\sqrt{1-x^2}\,U_n(x)$, $\sqrt{1+x}\,V_n(x)$, $\sqrt{1-x}\,W_n(x)$ are explicitly determined and occur, respectively, at

$$x = \cos\frac{(2k+1)\pi}{2(n+1)}, \quad x = \cos\frac{2k\pi}{2n+1}, \quad x = \cos\frac{(2k+1)\pi}{2n+1} \quad (k = 0, 1, \ldots, n).$$

2.3 Relations between Chebyshev polynomials and powers of x

It is useful and convenient in various applications to be able to express Chebyshev polynomials explicitly in terms of powers of x, and vice versa. Such formulae are simplest and easiest to derive in the case of the first kind polynomials $T_n(x)$, and so we concentrate on these.

2.3.1 Powers of x in terms of $\{T_n(x)\}$

The power x^n can be expressed in terms of the Chebyshev polynomials of degrees up to n, but, since these are alternately even and odd, we see at once that we need only include polynomials of alternate degrees, namely $T_n(x)$, $T_{n-2}(x)$, $T_{n-4}(x)$, ... Writing $x = \cos\theta$, we therefore need to express $\cos^n\theta$ in terms of $\cos n\theta$, $\cos(n-2)\theta$, $\cos(n-4)\theta$, ..., and this is readily achieved by using the binomial theorem as follows:

$$\begin{aligned}(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta})^n &= \mathrm{e}^{\mathrm{i}n\theta}+\binom{n}{1}\mathrm{e}^{\mathrm{i}(n-2)\theta}+\cdots+\binom{n}{n-1}\mathrm{e}^{-\mathrm{i}(n-2)\theta}+\mathrm{e}^{-\mathrm{i}n\theta}\\ &= (\mathrm{e}^{\mathrm{i}n\theta}+\mathrm{e}^{-\mathrm{i}n\theta})+\binom{n}{1}(\mathrm{e}^{\mathrm{i}(n-2)\theta}+\mathrm{e}^{-\mathrm{i}(n-2)\theta})+\\ &\quad+\binom{n}{2}(\mathrm{e}^{\mathrm{i}(n-4)\theta}+\mathrm{e}^{-\mathrm{i}(n-4)\theta})+\cdots. \end{aligned} \tag{2.13}$$

Here we have paired in brackets the first and last terms, the second and second-to-last terms, and so on. The number of such brackets will be

$$\lfloor n/2\rfloor+1$$

where $\lfloor m\rfloor$ denotes the integer part of m. When n is even, the last bracket in (2.13) will contain only the one (middle) term $\mathrm{e}^{0\theta}$ [= 1].

Now using the fact that

$$(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta})^n = (2\cos\theta)^n = 2^n\cos^n\theta$$

we deduce from (2.13) that

$$2^{n-1}\cos^n\theta = {\sum_{k=0}^{\lfloor n/2\rfloor}}{}' \binom{n}{k}\cos(n-2k)\theta,$$

where the dash ($\sum'$) denotes that the kth term in the sum is to be halved if n is even and $k = n/2$. Hence, from the definition (1.1) of $T_n(x)$,

$$x^n = 2^{1-n}{\sum_{k=0}^{\lfloor n/2\rfloor}}{}'\binom{n}{k}T_{n-2k}(x), \tag{2.14}$$

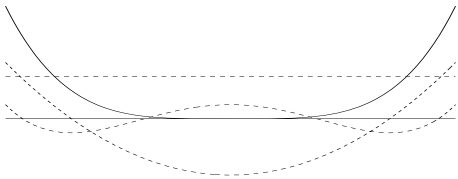


Figure 2.1: x^4 (full curve) and its decomposition into Chebyshev polynomials (broken curves)

where the dash now denotes that the term in $T_0(x)$, if there is one, is to be halved.

EXAMPLE 2.2: Taking $n = 4$ [see Figure 2.1]:

$$\begin{aligned} x^4 &= 2^{-3} \sum_{k=0}^{2}{}' \binom{4}{k} T_{4-2k}(x) \\ &= 2^{-3} \left[T_4(x) + \binom{4}{1} T_2(x) + \tfrac{1}{2} \binom{4}{2} T_0(x) \right] \\ &= \frac{1}{8} T_4(x) + \frac{1}{2} T_2(x) + \frac{3}{8} T_0(x). \end{aligned}$$

2.3.2 $T_n(x)$ in terms of powers of x

It is not quite as simple to derive formulae in the reverse direction. The obvious device to use is de Moivre's Theorem:

$$\cos n\theta + \mathrm{i} \sin n\theta = (\cos\theta + \mathrm{i}\sin\theta)^n.$$

Expanding by the binomial theorem and taking the real part,

$$\cos n\theta = \cos^n\theta - \binom{n}{2} \cos^{n-2}\theta \sin^2\theta + \binom{n}{4} \cos^{n-4}\theta \sin^4\theta + \cdots.$$

If $\sin^2\theta$ is replaced by $1 - \cos^2\theta$ throughout, then a formula is obtained for $\cos n\theta$ in terms of $\cos^n\theta$, $\cos^{n-2}\theta$, $\cos^{n-4}\theta$, ... On transforming to $x = \cos\theta$, this leads to the required formula for $T_n(x)$ in terms of x^n, x^{n-2}, x^{n-4}, ...

We omit the details here, but refer to Rivlin (1974), where the relevant result is obtained in the form

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[(-1)^k \sum_{j=k}^{\lfloor n/2 \rfloor} \binom{n}{2j} \binom{j}{k} \right] x^{n-2k}. \tag{2.15}$$

However, a rather simpler formula is given, for example, by Clenshaw (1962) and Snyder (1966) in the form

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} c_k^{(n)} x^{n-2k} \tag{2.16}$$

where

$$c_k^{(n)} = (-1)^k 2^{n-2k-1} \left[2\binom{n-k}{k} - \binom{n-k-1}{k} \right] \quad (2k < n) \tag{2.17a}$$

and

$$c_k^{(2k)} = (-1)^k \quad (k \geq 0). \tag{2.17b}$$

This formula may be proved by induction, using the three-term recurrence relation (1.3a), and we leave this as an exercise for the reader (Problem 5).

In fact the term in square brackets in (2.17a) may be further simplified, by taking out common ratios, to give

$$c_k^{(n)} = (-1)^k 2^{n-2k-1} \frac{n}{n-k} \binom{n-k}{k}. \tag{2.18}$$

Example 2.3: For $n = 6$ we obtain from (2.17b), (2.18):

$$c_0^{(6)} = 2^5 = 32; \quad c_1^{(6)} = (-1)^1 2^3 \frac{6}{5} \binom{5}{1} = -48;$$

$$c_2^{(6)} = (-1)^2 2^1 \frac{6}{4} \binom{4}{2} = 18; \quad c_3^{(6)} = (-1)^3 2^{-1} \frac{6}{3} \binom{3}{3} = -1.$$

Hence

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1.$$

For an alternative derivation of the results in this section, making use of generating functions, see Chapter 5.

2.3.3 Ratios of coefficients in $T_n(x)$

In applications, recurrence formulae which link pairs of coefficients are often more useful than explicit formulae (such as (2.18) above) for the coefficients themselves since, using such formulae, the whole sequence of coefficients may be assembled rather more simply and efficiently than by working them out one by one.

From (2.18),

$$\begin{aligned} c_k^{(n)} &= (-1)^k 2^{n-2k-1} \frac{n}{k} \frac{(n-k-1)(n-k-2)\cdots(n-2k+1)}{1\cdot 2\cdots(k-1)}; \\ c_{k+1}^{(n)} &= (-1)^{k+1} 2^{n-2k-3} \frac{n}{k+1} \frac{(n-k-2)(n-k-3)\cdots(n-2k-1)}{1\cdot 2\cdots k}. \end{aligned}$$

Hence, on dividing and cancelling common factors,

$$c_{k+1}^{(n)} = -\frac{(n-2k)(n-2k-1)}{4(k+1)(n-k-1)} c_k^{(n)} \tag{2.19}$$

where $c_k^{(n)}$ denotes the coefficient of x^n in $T_n(x)$. Formula (2.19) is valid for $n > 0$ and $k \geq 0$.

2.4 Evaluation of Chebyshev sums, products, integrals and derivatives

A variety of manipulations of Chebyshev polynomials and of sums or series of them can be required in practice. A secret to the efficient and stable execution of these tasks is to avoid rewriting Chebyshev polynomials in terms of powers of x and to operate wherever possible with the Chebyshev polynomials themselves (Clenshaw 1955).

2.4.1 Evaluation of a Chebyshev sum

Suppose that we wish to evaluate the sum

$$S_n = \sum_{r=0}^{n} a_r P_r(x) = a_0 P_0(x) + a_1 P_1(x) + \cdots + a_n P_n(x) \tag{2.20a}$$

where $\{P_r(x)\}$ are Chebyshev polynomials of either the first, second, third or fourth kinds. We may write (2.20a) in vector form as

$$S_n = \mathbf{a}^T \mathbf{p}, \tag{2.20b}$$

where $\mathbf{a}^T$ and $\mathbf{p}$ denote the row- and column-vectors

$$\mathbf{a}^T = (a_0,\ a_1,\ \ldots,\ a_n), \quad \mathbf{p} = \begin{pmatrix} P_0(x) \\ P_1(x) \\ \vdots \\ P_n(x) \end{pmatrix}.$$

In each of the four cases, from (1.3a), (1.6a), (1.12a), (1.12b) above, the recurrence relation between the polynomials takes the same form

$$P_r(x) - 2xP_{r-1}(x) + P_{r-2}(x) = 0,\ r = 2, 3, \ldots. \tag{2.21a}$$

with $P_0(x) = 1$ and, respectively,

$$\begin{gathered} P_1(x) = T_1(x) = x, \quad P_1(x) = U_1(x) = 2x, \\ P_1(x) = V_1(x) = 2x - 1, \quad P_1(x) = W_1(x) = 2x + 1. \end{gathered} \tag{2.21b}$$

Equations (2.21) may be written in matrix notation as

$$\begin{pmatrix} 1 & & & & & & \\ -2x & 1 & & & & & \\ 1 & -2x & 1 & & & & \\ & 1 & -2x & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2x & 1 & \\ & & & & 1 & -2x & 1 \end{pmatrix} \begin{pmatrix} P_0(x) \\ P_1(x) \\ P_2(x) \\ P_3(x) \\ \vdots \\ P_{n-1}(x) \\ P_n(x) \end{pmatrix} = \begin{pmatrix} 1 \\ X \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \tag{2.22a}$$

or (denoting the $(n+1) \times (n+1)$ matrix by $\mathbf{A}$)

$$\mathbf{A}\mathbf{p} = \mathbf{c} \tag{2.22b}$$

where

$$\mathbf{c} = \begin{pmatrix} 1 \\ X \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with $X = -x,\ 0,\ -1,\ 1$, respectively in the four cases.

Let

$$\mathbf{b}^T = (b_0,\ b_1,\ \ldots,\ b_n)$$

be the row vector satisfying the equation

$$(b_0,\ b_1,\ \ldots,\ b_n)\begin{pmatrix} 1 & & & & & & \\ -2x & 1 & & & & & \\ 1 & -2x & 1 & & & & \\ & 1 & -2x & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2x & 1 & \\ & & & & 1 & -2x & 1 \end{pmatrix} = \\ = (a_0,\ a_1,\ \ldots,\ a_n) \tag{2.23a}$$

or

$$\mathbf{b}^T\mathbf{A} = \mathbf{a}^T. \tag{2.23b}$$

Then we have

$$S_n = \mathbf{a}^T\mathbf{p} = \mathbf{b}^T\mathbf{A}\mathbf{p} = \mathbf{b}^T\mathbf{c} = b_0 + b_1X. \tag{2.24}$$

If we write $b_{n+1} = b_{n+2} = 0$, then the matrix equation (2.23a) can be seen to represent the recurrence relation

$$b_r - 2xb_{r+1} + b_{r+2} = a_r,\ r = 0, 1, \ldots, n. \tag{2.25}$$

We can therefore evaluate S_n by starting with $b_{n+1} = b_{n+2} = 0$ and performing the three-term recurrence (2.25) in the reverse direction,

$$b_r = 2xb_{r+1} - b_{r+2} + a_r,\ r = n, \ldots, 1, 0, \tag{2.26}$$

to obtain b_1 and b_0, and finally evaluating the required result S_n as

$$S_n = b_0 + b_1X. \tag{2.27}$$

For the first-kind polynomials $T_r(x)$, it is more usual to need the modified sum

$$S'_n = \sum_{r=0}^{n}{}' a_rT_r(x) = \tfrac{1}{2}a_0T_0(x) + a_1T_1(x) + \cdots + a_nT_n(x),$$

in which the coefficient of T_0 is halved, in which case (2.27) is replaced (remembering that $X = -x$) by

$$\begin{aligned} S'_n &= S_n - \tfrac{1}{2}a_0 \\ &= (b_0 - b_1x) + \tfrac{1}{2}(b_0 - 2xb_1 + b_2), \end{aligned}$$

or

$$S'_n = \tfrac{1}{2}(b_0 - b_2). \tag{2.28}$$

Note that, for a given x, carrying out the recurrence requires only $O(n)$ multiplications, and hence is as efficient as Horner's rule for evaluating a polynomial as a sum of powers using nested multiplication.

In some applications, which we shall refer to later, it is necessary to evaluate Chebyshev sums of a large number of terms at an equally large number of values of x. While the algorithm described above may certainly be used in such cases, one can often gain dramatically in efficiency by making use of the well-known *fast Fourier transform*, as we shall show later in Section 4.7.1.

Sums of even only or odd only polynomials, such as

$$S_n^{(0)} = {\sum_{r=0}^{n}}' \bar{a}_{2r} T_{2r}(x) \text{ and } S_n^{(1)} = \sum_{r=0}^{n} \bar{a}_{2r+1} T_{2r+1}(x)$$

may of course be evaluated by the above method, setting odd or even coefficients (respectively) to zero. However, the sum may be calculated much more efficiently using only the given even/odd coefficients by using a modified algorithm (Clenshaw 1962) which is given in Problem 7 below.

EXAMPLE 2.4: Consider the case $n = 2$ and $x = 1$ with coefficients

$$a_0 = 1, \quad a_1 = 0.1, \quad a_2 = 0.001.$$

Then from (2.21b) we obtain

$$\begin{aligned} b_3 &= b_4 = 0 \\ b_2 &= a_2 = 0.01 \\ b_1 &= 2b_2 - b_3 + 0.1 = 0.12 \\ b_0 &= 2b_1 - b_2 + 1 = 1.23. \end{aligned}$$

Hence

$$\begin{aligned} {\sum_{r=0}^{2}}' a_r T_r(1) &= \tfrac{1}{2}(b_0 - b_2) = 0.61 \\ \sum_{r=0}^{2} a_r U_r(1) &= b_0 = 1.23 \\ \sum_{r=0}^{2} a_r V_r(1) &= b_0 - b_1 = 1.11 \\ \sum_{r=0}^{2} a_r W_r(1) &= b_0 + b_1 = 1.35. \end{aligned}$$

To verify these formulae, we may set $\theta = 0$ (i.e., $x = 1$) in (1.1), (1.4), (1.8), (1.9), giving

$$T_n(1) = 1, \quad U_n(1) = n + 1, \tag{2.29a}$$

$$V_n(1) = 1, \quad W_n(1) = 2n + 1. \tag{2.29b}$$

Hence

$$\begin{aligned}
\sum_{r=0}^{2}{}' a_r T_r(1) &= \tfrac{1}{2}a_0 + a_1 + a_2 = 0.61 \\
\sum_{r=0}^{2} a_r U_r(1) &= a_0 + 2a_1 + 3a_2 = 1.23 \\
\sum_{r=0}^{2} a_r V_r(1) &= a_0 + a_1 + a_2 = 1.11 \\
\sum_{r=0}^{2} a_r W_r(1) &= a_0 + 3a_1 + 5a_2 = 1.35.
\end{aligned}$$

Incidentally, it is also useful to note that, by setting $\theta = \pi, \frac{1}{2}\pi$ in (1.1), (1.4), (1.8), (1.9), we can find further special values of the Chebyshev polynomials at $x = -1$ and $x = 0$, similar to those (2.29) at $x = 1$, namely

$$T_n(-1) = (-1)^n, \quad U_n(-1) = (-1)^n(n+1), \tag{2.30a}$$

$$V_n(-1) = (-1)^n(2n+1), \quad W_n(-1) = (-1)^n, \tag{2.30b}$$

$$T_{2n+1}(0) = U_{2n+1}(0) = 0, \quad T_{2n}(0) = U_{2n}(0) = (-1)^n, \tag{2.30c}$$

$$-V_{2n+1}(0) = W_{2n+1}(0) = (-1)^n, \quad V_{2n}(0) = W_{2n}(0) = (-1)^n. \tag{2.30d}$$

We leave the confirmation of formulae (2.29) and (2.30) as an exercise to the reader (Problem 8 below).

2.4.2 Stability of the evaluation of a Chebyshev sum

It is important to consider the effects of rounding errors when using recurrence relations, and specifically (2.26) above, since it is known that instability can sometimes occur. (By instability, we mean that rounding errors grow unacceptably fast relative to the true solution as the calculation progresses.) Three-term recurrence relations have two families of solutions, and it is possible for contributions from a relatively larger but unwanted solution to appear as rounding errors; so we need to take note of this. A brief discussion is given by Clenshaw (1962); a more detailed discussion is given by Fox & Parker (1968).

In the case of the recurrence (2.26), suppose that each b_s is computed with a local rounding error ϵ_s, which local errors together propagate into errors δ_r in b_r for $r < s$, resulting in an error Δ in S_n or Δ' in S'_n. Writing $\bar{b}_r$ for the computed b_r and $\bar{S}_n$ or $\bar{S}'_n$ for the S_n or S'_n computed without further error from (2.24) or (2.28), then from (2.26) (for fixed x)

$$\bar{b}_r = 2x\bar{b}_{r+1} - \bar{b}_{r+2} + a_r - \epsilon_r \tag{2.31}$$

while

$$b_r - \bar{b}_r = \delta_r. \tag{2.32}$$

Also

$$\begin{aligned} \bar{S}_n &= \bar{b}_0 + \bar{b}_1 X, \\ \bar{S}'_n &= \tfrac{1}{2}(\bar{b}_0 - \bar{b}_2), \end{aligned}$$

and

$$S_n - \bar{S}_n = \Delta, \quad S'_n - \bar{S}'_n = \Delta'.$$

From (2.26), (2.31), (2.32) we deduce that

$$\delta_r = 2x\delta_{r+1} - \delta_{r+2} + \epsilon_r \quad (r < s) \tag{2.33}$$

while

$$\begin{aligned} \Delta &= \delta_0 + \delta_1 X, \\ \Delta' &= \tfrac{1}{2}(\delta_0 - \delta_2). \end{aligned}$$

Now the recurrence (2.33), is identical in form to (2.26), with ϵ_r replacing a_r and δ_r replacing b_r, while obviously $\delta_{n+1} = \delta_{n+2} = 0$. Taking the final steps into account, we deduce that

$$\Delta = \sum_{r=0}^{n} \epsilon_r P_r(x), \tag{2.34}$$

where P_r is T_r, U_r, V_r or W_r, depending on the choice of X, and

$$\Delta' = {\sum_{r=0}^{n}}' \epsilon_r T_r(x). \tag{2.35}$$

Using the well-known inequality

$$\left| \sum_r x_r y_r \right| \le \left(\sum_r |x_r| \right) \max_r |y_r|,$$

we deduce the error bounds

$$|\Delta'| \le \left({\sum_{r=0}^{n}}' |\epsilon_r| \right) \max_{r=0}^{n} |T_r(x)| \le {\sum_{r=0}^{n}}' |\epsilon_r| \tag{2.36}$$

and

$$|\Delta| \le \left(\sum_{r=0}^{n} |\epsilon_r|\right) \max_{r=0}^{n} |P_r(x)| \le C_n \sum_{r=0}^{n} |\epsilon_r| , \tag{2.37}$$

where $C_n = 1,\ n+1,\ 2n+1,\ 2n+1$ when P_r is T_r, U_r, V_r or W_r, respectively. (Note that the ϵ_r in these formulae are the absolute, not relative, errors incurred at each step of the calculation.)

2.4.3 Evaluation of a product

It is frequently necessary to be able to multiply Chebyshev polynomials by each other, as well as by factors such as x, $1-x$ and $1-x^2$, and to re-express the result in terms of Chebyshev polynomials. Such products are much less readily carried out for second-, third- and fourth-kind polynomials, as a consequence of the denominators in their trigonometric definitions. We therefore emphasise $T_n(x)$ and to a lesser extent $U_n(x)$.

Various formulae are readily obtained by using the substitution $x = \cos\theta$ and trigonometric identities, as follows.

$$T_m(x)T_n(x) = \cos m\theta \cos n\theta = \tfrac{1}{2}(\cos(m+n)\theta + \cos|m-n|\,\theta),$$

giving

$$T_m(x)T_n(x) = \tfrac{1}{2}(T_{m+n}(x) + T_{|m-n|}(x)). \tag{2.38}$$

$$\begin{aligned} xT_n(x) &= \cos\theta\cos n\theta = \tfrac{1}{2}(\cos(n+1)\theta + \cos|n-1|\,\theta), \\ xU_n(x)\sin\theta &= \cos\theta\sin(n+1)\theta = \tfrac{1}{2}(\sin(n+2)\theta + \sin n\theta), \end{aligned}$$

giving

$$xT_n(x) = \tfrac{1}{2}(T_{n+1}(x) + T_{|n-1|}(x)) \tag{2.39}$$

and

$$xU_n(x) = \tfrac{1}{2}(U_{n+1}(x) + U_{n-1}(x)), \tag{2.40}$$

(provided that we interpret $U_{-1}(x)$ as $\sin 0/\sin\theta = 0$).

More generally, we may also obtain expressions for $x^m T_n(x)$ (and similarly $x^m U_n(x)$) for any m, by expressing x^m in terms of Chebyshev polynomials by (2.14) and then using (2.38). (See Problem 4 below.)

In a similar vein,

$$\begin{aligned} (1-x^2)T_n(x) &= \sin^2\theta\cos n\theta = \tfrac{1}{2}(1-\cos 2\theta)\cos n\theta \\ &= \tfrac{1}{2}\cos n\theta - \tfrac{1}{4}(\cos(n+2)\theta + \cos|n-2|\,\theta), \\ (1-x^2)U_n(x)\sin\theta &= \sin^2\theta\sin(n+1)\theta = \tfrac{1}{2}(1-\cos 2\theta)\sin(n+1)\theta \\ &= \tfrac{1}{2}\sin(n+1)\theta - \tfrac{1}{4}(\sin(n+3)\theta + \sin(n-1)\theta), \end{aligned}$$

giving

$$(1 - x^2)T_n(x) = -\tfrac{1}{4}T_{n+2}(x) + \tfrac{1}{2}T_n(x) - \tfrac{1}{4}T_{|n-2|}(x) \tag{2.41}$$

and

$$(1 - x^2)U_n(x) = -\tfrac{1}{4}U_{n+2}(x) + \tfrac{1}{2}U_n(x) - \tfrac{1}{4}U_{n-2}(x) \tag{2.42}$$

where we interpret $U_{-1}(x)$ as 0 again, and $U_{-2}(x)$ as $\sin(-\theta)/\sin\theta = -1$.

Note that the particular cases $n = 0$, $n = 1$ are included in the formulae above, so that, specifically

$$\begin{aligned}
xT_0(x) &= T_1(x), \\
xU_0(x) &= \tfrac{1}{2}U_1(x), \\
(1 - x^2)T_0(x) &= \tfrac{1}{2}T_0(x) - \tfrac{1}{2}T_2(x), \\
(1 - x^2)T_1(x) &= \tfrac{1}{4}T_1(x) - \tfrac{1}{4}T_3(x), \\
(1 - x^2)U_0(x) &= \tfrac{3}{4}U_0(x) - \tfrac{1}{4}U_2(x), \\
(1 - x^2)U_1(x) &= \tfrac{1}{2}U_1(x) - \tfrac{1}{4}U_3(x).
\end{aligned}$$

2.4.4 Evaluation of an integral

The indefinite integral of $T_n(x)$ can be expressed in terms of Chebyshev polynomials as follows. By means of the usual substitution $x = \cos\theta$,

$$\begin{aligned}
\int T_n(x)\,\mathrm{d}x &= \int -\cos n\theta \sin\theta \,\mathrm{d}\theta \\
&= -\tfrac{1}{2}\int (\sin(n+1)\theta - \sin(n-1)\theta)\,\mathrm{d}\theta \\
&= \tfrac{1}{2}\left[\frac{\cos(n+1)\theta}{n+1} - \frac{\cos|n-1|\,\theta}{n-1}\right]
\end{aligned}$$

(where the second term in the bracket is to be omitted in the case $n = 1$).

Hence

$$\int T_n(x)\,\mathrm{d}x = \begin{cases} \tfrac{1}{2}\left[\dfrac{T_{n+1}(x)}{n+1} - \dfrac{T_{|n-1|}(x)}{n-1}\right], & n \neq 1; \\ \tfrac{1}{4}T_2(x), & n = 1. \end{cases} \tag{2.43}$$

Clearly this result can be used to integrate the sum

$$S_n(x) = \sum_{r=0}^{n}{}' a_r T_r(x)$$

in the form

$$
\begin{aligned}
I_{n+1}(x) &= \int S_n(x)\,\mathrm{d}x \\
&= \text{constant } + \tfrac{1}{2}a_0T_1(x) + \tfrac{1}{4}a_1T_2(x) + \sum_{r=2}^{n} \frac{a_r}{2}\left[\frac{T_{r+1}(x)}{r+1} - \frac{T_{r-1}(x)}{r-1}\right] \\
&= \sum_{r=0}^{n+1}{}' A_rT_r(x) \qquad (2.44)
\end{aligned}
$$

where A_0 is determined from the constant of integration, and

$$
A_r = \frac{a_{r-1} - a_{r+1}}{2r}, \quad r > 0, \qquad (2.45)
$$

with $a_{n+1} = a_{n+2} = 0$.

EXAMPLE 2.5: Table 2.1 gives 5-decimal values of A_r computed from values of a_r, obtained from an infinite expansion of the function e^x, after each value of a_r had been rounded to 4 decimals (numbers taken from Clenshaw (1962)). Each A_r would be identical to a_r for an exact calculation, but it is interesting to observe that, although there is a possible rounding error of ± 0.00005 in each given a_r, all the computed A_r actually have errors significantly smaller than this.

Table 2.1: Integration of a Chebyshev series

r	a_r	A_r	$\lvert$error in $A_r\rvert$
0	2.53213	—	—
1	1.13032	1.13030	0.00002
2	0.27150	0.27150	0.00000
3	0.04434	0.04433	0.00001
4	0.00547	0.00548	0.00001
5	0.00054	0.00055	0.00001
6	0.00004	0.00004	0.00000

There is an interesting and direct integral relationship between the Chebyshev polynomials of the first and second kinds, namely

$$
\int U_n(x)\,\mathrm{d}x = \frac{1}{n+1}T_{n+1}(x) + \text{ constant} \qquad (2.46)
$$

(which is easily verified by substituting $x = \cos\theta$). Hence, the sum

$$S_n(x) = \sum_{r=1}^{n} b_r U_{r-1}(x)$$

can be integrated immediately to give

$$\int S_n(x)\,\mathrm{d}x = \sum_{r=1}^{n} \frac{b_r}{r} T_r(x) + \text{ constant.} \tag{2.47}$$

2.4.5 Evaluation of a derivative

The formula for the derivative of $T_n(x)$ in terms of first-kind polynomials is not quite as simple as (2.43). From (2.46) we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}x} T_{n+1}(x) = (n+1)U_n(x), \tag{2.48}$$

so that it is easily expressed in terms of a second-kind polynomial. Then from (1.6b) and (1.7) it follows that

$$\frac{\mathrm{d}}{\mathrm{d}x} T_n(x) = 2n \sideset{}{'}\sum_{\substack{r=0 \\ n-r \text{ odd}}}^{n-1} T_r(x). \tag{2.49}$$

However, the derivative of a finite sum of first-kind Chebyshev polynomials is readily expressible as a sum of such polynomials, by reversing the process used in the integration of (2.44). Given the Chebyshev sum (of degree $n+1$, say)

$$I_{n+1}(x) = \sideset{}{'}\sum_{r=0}^{n+1} A_r T_r(x),$$

then

$$S_n(x) = \frac{\mathrm{d}}{\mathrm{d}x} I_{n+1} = \sideset{}{'}\sum_{r=0}^{n} a_r T_r(x) \tag{2.50}$$

where the coefficients $\{a_r\}$ are derived from the given $\{A_r\}$ by using (2.45) in the form

$$a_{r-1} = a_{r+1} + 2rA_r, \quad (r = n+1, n, \ldots, 1) \tag{2.51a}$$

with

$$a_{n+1} = a_{n+2} = 0. \tag{2.51b}$$

Explicitly, if we prefer, we may say that

$$a_r = \sum_{\substack{k=r+1 \\ k-r \text{ odd}}}^{n+1} 2kA_k. \tag{2.52}$$

EXAMPLE 2.6: Table 2.2 shows 4-decimal values of a_r computed from 4-decimal values of A_r, for the same example as in Table 2.1. Each a_r would be identical to A_r in an exact computation, and we see this time that the derivative $S_n(x)$ is less accurate than the original polynomial $I_{n+1}(x)$ by nearly one decimal place. The contrast between these results is consistent with the principle that, in general, numerical integration is a stable process and numerical differentiation an unstable process. The size of the errors in the latter case can be attributed to the propagation, by (2.51a), of the error inherent in the assumptions (2.51b).

Table 2.2: Differentiation of a Chebyshev series

r	A_r	a_r	$\lvert$error in $a_r\rvert$
0	2.53213	2.5314	0.0007
1	1.13032	1.1300	0.0003
2	0.27150	0.2708	0.0007
3	0.04434	0.0440	0.0003
4	0.00547	0.0050	0.0005
5	0.00054	0.0000	0.0005

There is another relatively simple formula for the derivative of $T_n(x)$, which we can obtain as follows.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}T_n(x) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\cos n\theta \Big/ \frac{\mathrm{d}}{\mathrm{d}\theta}\cos\theta \\
&= \frac{n\sin n\theta}{\sin\theta} \\
&= \frac{\frac{1}{2}n(\cos(n-1)\theta - \cos(n+1)\theta)}{\sin^2\theta} \\
&= \frac{\frac{1}{2}n(T_{n-1}(x) - T_{n+1}(x))}{1-x^2}.
\end{aligned}$$

Thus, for $|x| \neq 1$,

$$\frac{\mathrm{d}}{\mathrm{d}x}T_n(x) = \frac{n}{2}\frac{T_{n-1}(x) - T_{n+1}(x)}{1-x^2}. \tag{2.53}$$

Higher derivatives may be obtained by similar formulae (see Problem 17 for the second derivative).

2.5 Problems for Chapter 2

1. Determine the positions of the zeros of the Chebyshev polynomials of the second and third kinds for the general interval $[a, b]$ of x.

2. From numerical values of the cosine function (from table, calculator or computer), determine the zeros of $T_4(x)$, $U_4(x)$, $V_4(x)$, $W_4(x)$ and the extrema of $T_4(x)$.

3. Show that

   (a) $\frac{1}{2}U_{2k}(x) = \frac{1}{2}T_0(x) + T_2(x) + T_4(x) + \cdots + T_{2k}(x)$;

   (b) $\frac{1}{2}U_{2k+1}(x) = T_1(x) + T_3(x) + \cdots + T_{2k+1}(x)$;

   (c) $xU_{2k+1}(x) = T_0(x) + 2T_2(x) + \cdots + 2T_{2k-2}(x) + T_{2k}(x)$.

   [Hint: In (3a), multiply by $\sin\theta$ and use $2\sin A\cos B = \sin(A - B) + \sin(A + B)$. Use similar ideas in (3b), (3c).]

4. Obtain the expression

$$x^m T_n(x) = 2^{-m}\sum_{r=0}^{m}\binom{m}{r}T_{n-m-2r}(x) \quad (m < n)$$

   (a) by applying the formula (2.39) m times;

   (b) by applying the expression (2.14) for x^m in terms of Chebyshev polynomials and the expression (2.38) for products of Chebyshev polynomials.

5. Prove by induction on n that

$$T_n(x) = \sum_{k=0}^{\lfloor n/2\rfloor} c_k^n x^{n-2k},$$

   where

$$c_k^n = (-1)^k 2^{n-2k-1}\left[2\binom{n-k}{k} - \binom{n-k-1}{k}\right] \quad (n, k > 0)$$

$$c_0^n = 2^{n-1} \quad (n > 0)$$

$$c_0^0 = 1.$$

   [Hint: Assume the formulae are true for $n = N - 2$, $N - 1$ and hence derive them for $n = N$, using $T_n = 2xT_{n-1} - T_{n-2}$.]

6. Derive formulae for $T_m^*(x)T_n^*(x)$ and $xT_n^*(x)$ in terms of $\{T_r^*(x)\}$, using the ideas of Section 2.4.3.

7. Suppose

$$S_n^{(0)} = \sum_{r=0}^{n}{}' a_{2r}T_{2r}(x), \quad S_n^{(1)} = \sum_{r=0}^{n} a_{2r+1}T_{2r+1}(x)$$

are sums of even-only/odd-only Chebyshev polynomials.

(a) Show that $S_n^{(0)}$ may be efficiently determined by applying the recurrence (2.26) followed by (2.28), with x replaced by $(2x^2-1)$ and a_r replaced by a_{2r};

(b) Show that $S_n^{(1)}$ may be efficiently determined by applying the recurrence (2.26), with x replaced by $(2x^2 - 1)$ and a_r replaced by a_{2r+1}, and then taking

$$S_n^{(1)} = x(b_0 - b_1).$$

[Hint: From (1.14) and (1.15), we have $T_{2r}(x) = T_r(2x^2 - 1)$ and $T_{2r+1}(x) = xV_r(2x^2 - 1)$.]

8. Derive the formulae (2.29a)–(2.30d) for the values of T_n, U_n, V_n, W_n at $x = -1,\ 0,\ 1$, using only the trigonometric definitions of the Chebyshev polynomials.

9. Use the algorithm (2.21b) to evaluate

$$\sum_{r=0}^{3}{}' c_rT_r(x),\ \sum_{r=0}^{3} c_rU_r(x),\ \sum_{r=0}^{3} c_rV_r(x),\ \sum_{r=0}^{3} c_rW_r(x)$$

at $x = -1,\ 0,\ 1$ for $c_0 = 1$, $c_1 = 0.5$, $c_2 = 0.25$, $c_3 = 0.125$. Check your results using correct values of T_r, U_r, V_r, W_r at 0, 1.

10. Illustrate the algorithms (7a), (7b) of Problem 7 by using them to evaluate at $x = -1$

$$\sum_{0}^{2}{}' c_rT_r(x),\ \sum_{0}^{2}{}' c_rT_{2r}(x),\ \sum_{0}^{2}{}' c_rT_{2r+1}(x),$$

where $c_0 = 1$, $c_1 = 0.1$, $c_2 = 0.001$. Check your results using correct values of T_r at $x = -1$.

11. Discuss the stability of the summation formulae for sums of Chebyshev polynomials U_r, V_r, W_r when the size of each sum is

(a) proportional to unity,

(b) proportional to the largest value in $[-1, 1]$ of U_n, V_n, W_n, respectively (where the sums are from $r = 0$ to $r = n$).

12. Show that

    (a) $2(1-x^2)U_{n-2}(x) = T_n(x) - T_{n-2}(x)$;
    (b) $(1+x)V_{n-1}(x) = T_n(x) + T_{n-1}(x)$;
    (c) $(1-x)W_{n-1}(x) = T_n(x) - T_{n-1}(x)$;
    (d) $(1+x)V_m(x)V_n(x) = T_{|m-n|}(x) + T_{m+n+1}(x)$;
    (e) $(1-x)W_m(x)W_n(x) = T_{|m-n|}(x) - T_{m+n+1}(x)$.

13. Show that $T_m(x)U_{n-1}(x) = \frac{1}{2}\{U_{n+m-1}(x) + U_{n-m-1}(x)\}$, and determine an expression for $x^m U_{n-1}(x)$ in terms of $\{U_k\}$ by a similar procedure to that of Problem 4.

14. Show that (ignoring constants of integration)

    (a) $\int (1-x^2)^{-\frac{1}{2}} T_n(x)\,\mathrm{d}x = n^{-1}(1-x^2)^{\frac{1}{2}} U_{n-1}(x)$;
    (b) $\int (1-x)^{-\frac{1}{2}} V_n(x)\,\mathrm{d}x = (n+\frac{1}{2})^{-1}(1-x)^{\frac{1}{2}} W_n(x)$;
    (c) $\int (1+x)^{-\frac{1}{2}} W_n(x)\,\mathrm{d}x = (n+\frac{1}{2})^{-1}(1+x)^{\frac{1}{2}} V_n(x)$.

15. Show that, for $n > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}x}U_n(x) = \frac{(n+2)U_{n-1}(x) - nU_{n+1}(x)}{2(1-x^2)}.$$

16. Using (2.52), show that if

$$\sum_{r=0}^{n+1}{}' A_r T_r(x) = I_{n+1}(x)$$

and

$$\sum_{r=0}^{n-1}{}' a_r T_r(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} I_{n+1}(x),$$

then

$$a_r = \sum_{\substack{k=r+2 \\ k-r \text{ even}}}^{n+1} (k-r)k(k+r)A_k.$$

Show further that

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}T_n(x) = \sum_{\substack{r=0 \\ n-r \text{ even}}}^{n-2}{}' (n-r)n(n+r)T_r(x).$$

17. Using (2.53) and (1.3a), prove that, for $n > 1$,

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} T_n(x) = \frac{n}{4} \frac{(n+1)T_{n-2}(x) - 2nT_n(x) + (n-1)T_{n+2}(x)}{(1-x^2)^2}.$$

18. Show that

$$\sum_{j=0}^{n}{}' T_j(x) T_j(y) = \tfrac{1}{4} \Big\{ W_n \left(xy + \sqrt{(1-x^2)(1-y^2)} \right) + \\ + W_n \left(xy - \sqrt{(1-x^2)(1-y^2)} \right) \Big\}.$$

19. Show that

$$(1-x^2) \sum_{j=0}^{\infty}{}' c_j T_j(x) = \tfrac{1}{4} \sum_{j=0}^{\infty} (c_j - c_{j+2})(T_j(x) - T_{j+2}(x)).$$

Chapter 3

The Minimax Property and Its Applications

3.1 Approximation — theory and structure

One area above all in which the Chebyshev polynomials have a pivotal role is the minimax approximation of functions by polynomials. It is therefore appropriate at the beginning of this discussion to trace the structure of the subject of approximation and to present some essential theoretical results, concentrating primarily on uniform (or $\mathcal{L}_\infty$) approximation and introducing the minimax property of the Chebyshev polynomials.

It is very useful to be able to replace any given function by a simpler function, such as a polynomial, chosen to have values not identical with but very close to those of the given function, since such an 'approximation' may not only be more compact to represent and store but also more efficient to evaluate or otherwise manipulate. The structure of an 'approximation problem' involves three central components: (i) a *function class* (containing the function to be approximated), (ii) a *form* (for the approximating function) and (iii) a *norm* (of the approximation error), in terms of which the problem may be formally posed. The expert's job is to make appropriate selections of these components, then to pose the approximation problem, and finally to solve it.

By a *function class*, we mean a restricted family $\mathcal{F}$ of functions f to which any function $f(x)$ that we may want to fit is assumed to belong. Unless otherwise stated, we shall be concerned with real functions of a real variable, but the family will generally be narrower than this. For example we may consider amongst others the following alternative families $\mathcal{F}$ of functions defined on the real interval $[a, b]$:

1. $\mathcal{C}[a, b]$: continuous functions on $[a, b]$;

2. $\mathcal{L}_\infty[a, b]$: bounded functions on $[a, b]$;

3. $\mathcal{L}_2[a, b]$: square-integrable functions on $[a, b]$;

4. $\mathcal{L}_p[a, b]$: $\mathcal{L}_p$-integrable functions on $[a, b]$, namely functions $f(x)$ for which is defined

$$\int_a^b w(x)\,|f(x)|^p\,\mathrm{d}x, \tag{3.1}$$

   where $w(x)$ is a given non-negative weight function and $1 \leq p < \infty$. Note that $\mathcal{L}_2[a, b]$ is a special case ($p = 2$) of $\mathcal{L}_p[a, b]$.

The reason for defining such a family of functions, when in practice we may only in fact be interested in one specific function, is that this helps to isolate those properties of the function that are relevant to the theory — moreover, there is a close link between the function class we work in and the norms we can use. In particular, in placing functions in one of the four families listed above, it is implicitly assumed that we neither care how the functions behave nor wish to approximate them outside the given interval $[a, b]$.

By *form of approximation* we mean the specific functional form which is to be adopted, which will always include adjustable coefficients or other parameters. This defines a family $\mathcal{A}$ of possible approximations $f^*(x)$ to the given function $f(x)$. For example, we might draw our approximation from one of the following families:

1. Polynomials of degree n, with

$$\mathcal{A} = \Pi_n = \{f^*(x) = p_n(x) = c_0 + c_1 x + \cdots + c_n x^n\} \quad \text{(parameters } \{c_j\})$$

2. Rational functions of type (p, q), with

$$\mathcal{A} = \left\{ f^*(x) = r_{p,q}(x) = \frac{a_0 + a_1 + \cdots + a_p x^p}{1 + b_1 + \cdots + b_q x^q} \right\} \quad \text{(parameters } \{a_j\}, \{b_j\})$$

For theoretical purposes it is usually desirable to choose the function class $\mathcal{F}$ to be a *vector space* (or linear space). A vector space $\mathcal{V}$ comprises elements u, v, w, ... with the properties (which vectors in the conventional sense are easily shown to possess):

1. (*closure under addition*)
   $u + v \in \mathcal{V}$ for any $u, v \in \mathcal{V}$,

2. (*closure under multiplication by a scalar*)
   $\alpha u \in \mathcal{V}$ for any $u \in \mathcal{V}$ and for any scalar α.

When these elements are functions $f(x)$, with $f + g$ and αf defined as the functions whose values at any point x are $f(x) + g(x)$ and $\alpha f(x)$, we refer to $\mathcal{F}$ as a *function space*. This space $\mathcal{F}$ typically has infinite dimension, the 'vector' in question consisting of the values of $f(x)$ at each of the continuum of points x in $[a, b]$.

The family $\mathcal{A}$ of approximations is normally taken to be a subclass of $\mathcal{F}$:

$$\mathcal{A} \subset \mathcal{F}$$

— in practice, $\mathcal{A}$ is usually also a vector space, and indeed a function space. In contrast to $\mathcal{F}$, $\mathcal{A}$ is a *finite dimensional* function space, its dimension being the number of parameters in the form of approximation. Thus the space Π_n of polynomials $p_n(x)$ of degree n has dimension $n+1$ and is in fact isomorphic (i.e., structurally equivalent) to the space $\mathbb{R}^{n+1}$ of real vectors with $n+1$ components:

$$\{\mathbf{c} = (c_0, c_1, \ldots, c_n)\}.$$

(Note that the family of rational functions $r_{p,q}$ of type (p,q) is *not* a vector space, since the sum of two such functions is in general a rational function of type $(p+q, 2q)$, which is not a member of the same family.)

The *norm* of approximation $\|\cdot\|$ serves to compare the function $f(x)$ with the approximation $f^*(x)$, and gives a single scalar measure of the closeness of f^* to f, namely

$$\|f - f^*\|. \tag{3.2}$$

Definition 3.1 *A norm* $\|\cdot\|$ *is defined as any real scalar measure of elements of a vector space that satisfies the axioms:*

1. $\|u\| \geq 0$, *with equality if and only if* $u \equiv 0$;
2. $\|u+v\| \leq \|u\| + \|v\|$ *(*the 'triangle inequality'*)*;
3. $\|\alpha u\| = |\alpha| \; \|u\|$ *for any scalar* α.

Such a definition encompasses all the key features of distance or, in the case of a function, size. Standard choices of norm for function spaces are the following:

1. $\mathcal{L}_\infty$ norm (or *uniform norm*, *minimax norm*, or *Chebyshev norm*):

$$\|f\| = \|f\|_\infty = \max_{a \leq x \leq b} |f(x)|; \tag{3.3}$$

2. $\mathcal{L}_2$ norm (or *least-squares norm*, or *Euclidean norm*):

$$\|f\| = \|f\|_2 = \sqrt{\int_a^b w(x) \, |f(x)|^2 \, \mathrm{d}x}, \tag{3.4}$$

where $w(x)$ is a non-negative weight function;

3. $\mathcal{L}_1$ norm (or *mean norm*, or *Manhattan norm*):

$$\|f\| = \|f\|_1 = \int_a^b w(x) \, |f(x)| \, \mathrm{d}x; \tag{3.5}$$

4. The above three norms can be collected into the more general $\mathcal{L}_p$ norm (or *Hölder norm*):

$$\|f\| = \|f\|_p = \left[\int_a^b w(x) \, |f(x)|^p \, \mathrm{d}x \right]^{\frac{1}{p}}, \quad (1 \leq p < \infty), \tag{3.6}$$

where $w(x)$ is a non-negative weight function.

With suitable restrictions on f, which are normally satisfied in practice, this $\mathcal{L}_p$ norm corresponds to the $\mathcal{L}_\infty$, $\mathcal{L}_2$ and $\mathcal{L}_1$ norms in the cases $p \to \infty$, $p = 2$, $p = 1$, respectively.

5. The *weighted minimax norm*:

$$\|f\| = \max_{a \le x \le b} w(x)\,|f(x)| \tag{3.7}$$

(which does not fall into the pattern of Hölder norms) also turns out to be appropriate in some circumstances.

The $\mathcal{L}_p$ norm becomes stronger as p increases, as the following lemma indicates.

Lemma 3.1 *If $1 \le p_1 < p_2 \le \infty$, and if a, b and $\int_a^b w(x)\,dx$ are finite, then $\mathcal{L}_{p_2}[a,b]$ is a subspace of $\mathcal{L}_{p_1}[a,b]$, and there is a finite constant $k_{p_1 p_2}$ such that*

$$\|f\|_{p_1} \le k_{p_1 p_2} \|f\|_{p_2} \tag{3.8}$$

for every f in $\mathcal{L}_{p_2}[a,b]$.

This lemma will be deduced from Hölder's inequality in Chapter 5 (see Lemma 5.4 on page 117).

A vector space to which a norm has been attached is termed a *normed linear space*. Hence, once a norm is chosen, the vector spaces $\mathcal{F}$ and $\mathcal{A}$ of functions and approximations become normed linear spaces.

3.1.1 The approximation problem

We defined above a family of functions or function space, $\mathcal{F}$, a family of approximations or approximation (sub)space, $\mathcal{A}$, and a measure $\|f - f^*\|$ of how close a given function $f(x)$ in $\mathcal{F}$ is to a derived approximation $f^*(x)$ in $\mathcal{A}$. How then do we more precisely judge the quality of $f^*(x)$, as an approximation to $f(x)$ in terms of this measure? In practice there are three types of approximation that are commonly aimed for:

Definition 3.2 *Let $\mathcal{F}$ be a normed linear space, let $f(x)$ in $\mathcal{F}$ be given, and let $\mathcal{A}$ be a given subspace of $\mathcal{F}$.*

1. *An approximation $f^*(x)$ in $\mathcal{A}$ is said to be* good *(or* acceptable*) if*

$$\|f - f^*\| \le \epsilon \tag{3.9}$$

where ϵ is a prescribed level of absolute accuracy.

2. *An approximation* $f_B^*(x)$ *in* $\mathcal{A}$ *is a* best *approximation if, for any other approximation* $f^*(x)$ *in* $\mathcal{A}$,

$$\|f - f_B^*\| \le \|f - f^*\| . \qquad (3.10)$$

Note that there will sometimes be more than one best approximation to the same function.

3. *An approximation* $f_N^*(x)$ *in* $\mathcal{A}$ *is said to be* near-best within a relative distance ρ *if*

$$\|f - f_N^*\| \le (1+\rho)\,\|f - f_B^*\| , \qquad (3.11)$$

where ρ *is a specified positive scalar and* $f_B^*(x)$ *is a best approximation.*

In the case of the $\mathcal{L}_\infty$ *norm, we often use the terminology* minimax *and* near-minimax *in place of* best *and* near-best.

The 'approximation problem' is to determine an approximation of one of these types (good, best or near-best). In fact, it is commonly required that both 'good' and 'best', or both 'good' and 'near-best', should be achieved — after all, it cannot be very useful to obtain a best approximation if it is also a very poor approximation.

In defining 'good' in Definition 3.2 above, an absolute error criterion is adopted. It is, however, also possible to adopt a *relative error* criterion, namely

$$\left\| 1 - \frac{f^*}{f} \right\| \le \epsilon. \qquad (3.12)$$

This can be viewed as a problem of weighted approximation in which we require

$$\|w\,(f - f^*)\| \le \epsilon, \qquad (3.13)$$

where, in this case,

$$w(x) = 1/\,|f(x)|\,.$$

In approximating by polynomials on $[a, b]$, it is always possible to obtain a good approximation by taking the degree high enough. This is the conclusion of the following well-known results.

Theorem 3.2 (Weierstrass's theorem) *For any given* f *in* $\mathcal{C}[a, b]$ *and for any given* $\epsilon > 0$, *there exists a polynomial* p_n *for some sufficiently large* n *such that* $\|f - p_n\|_\infty < \epsilon$.

Proof: A proof of this will be given later (see Corollary 5.8A on page 120). ●●

Corollary 3.2A *The same holds for* $\|f - p_n\|_p$ *for any* $p \ge 1$.

Proof: This corollary follows directly by applying Lemma 3.1. ●●

But of course it is a good thing from the point of view of efficiency if we can keep the degree of polynomial as low as possible, which we can do by concentrating on best or near-best approximations.

3.2 Best and minimax approximation

Given a norm $\|\cdot\|$ (such as $\|\cdot\|_\infty$, $\|\cdot\|_2$ or $\|\cdot\|_1$), a best approximation as defined by (3.10) is a solution of the problem:

$$\underset{f^* \in \mathcal{A}}{\text{minimise}} \|f - f^*\| . \tag{3.14}$$

In the case of polynomial approximation:

$$f^*(x) = p_n(x) = c_0 + c_1 x + \cdots + c_n x^n, \tag{3.15}$$

to which we now restrict our attention, we may rewrite (3.14) in terms of the parameters as:

$$\underset{c_0,\ldots,c_n}{\text{minimise}} \|f - p_n\| . \tag{3.16}$$

Can we always find such a p_n? Is there just one?

Theorem 3.3 *For any given p ($1 \leq p \leq \infty$), there exists a unique best polynomial approximation p_n to any function $f \in \mathcal{L}_p[a,b]$ in the $\mathcal{L}_p$ norm, where $w(x)$ is taken to be unity in the case $p \to \infty$.*

We refrain from giving proofs, but refer the reader to Cheney (1966), or other standard texts, for details.

Note that best approximations also exist in $\mathcal{L}_p$ norms on finite point sets, for $1 \leq p \leq \infty$, and are then unique for $p \neq 1$ but not necessarily unique for $p = 1$. Such $\mathcal{L}_p$ norms are defined by:

$$\|f - f^*\|_p = \left[\sum_{i=1}^{m} w_i \left| f(x_i) - f^*(x_i) \right|^p \right]^{\frac{1}{p}}$$

where $\{w_i\}$ are positive scalar weights and $\{x_i\}$ is a discrete set of m fitting points where the value of $f(x_i)$ is known. These are important in data fitting problems; however, this topic is away from our central discussion, and we shall not pursue it here.

It is possible to define forms of approximation other than polynomials, for which existence or uniqueness of best approximation holds — see Cheney (1966) for examples. Since polynomials are the subject of this book, however, we shall again refrain from going into details.

Note that Theorem 3.3 guarantees in particular the existence of a unique best approximation in the $\mathcal{L}_\infty$ or minimax norm. The best $\mathcal{L}_\infty$ or minimax approximation problem, combining (3.3) and (3.15), is (in concise notation)

$$\underset{c_0,\ldots,c_n}{\text{minimise}} \max_{a\le x\le b} |f(x) - p_n(x)| . \tag{3.17}$$

It is clear from (3.17) why the word 'minimax' is often given to this problem, and why the resulting best approximation is often referred to as a 'minimax approximation'.

Theorem 3.3 is not a constructive theorem and does not characterise (i.e. describe how to recognise) a minimax approximation. However, it is possible to do so rather explicitly, as the following powerful theorem asserts.

Theorem 3.4 (Alternation theorem for polynomials) *For any $f(x)$ in $\mathcal{C}[a, b]$ a unique minimax polynomial approximation $p_n(x)$ exists, and is uniquely characterised by the 'alternating property' (or 'equioscillation property') that there are $n + 2$ points (at least) in $[a, b]$ at which $f(x) - p_n(x)$ attains its maximum absolute value (namely $\|f - p_n\|_\infty$) with alternating signs.*

This theorem, often ascribed to Chebyshev but more properly attributed to Borel (1905), asserts that, for p_n to be the best approximation, it is both *necessary and sufficient* that the alternating property should hold, that only one polynomial has this property, and that there is only one best approximation. The reader is referred to Cheney (1966), for example, for a complete proof. The 'sufficient' part of the proof is relatively straightforward and is set as Problem 6 below; the 'necessary' part of the proof is a little more tricky.

Example 3.1: As an example of the alternation theorem, suppose that the function $f(x) = x^2$ is approximated by the first-degree ($n = 1$) polynomial

$$f^*(x) = p_1(x) = x - 0.125 \tag{3.18}$$

on $[0, 1]$. Then the error $f(x) - p_n(x)$, namely

$$x^2 - x + 0.125,$$

has a maximum magnitude of 0.125 which it attains at $x = 0$, 0.5 and 1. At these points it takes the respective values $+0.125$, -0.125 and $+0.125$, which have alternating signs. (See Figure 3.1.) Hence $p_1(x)$, given by (3.18), is the unique minimax approximation.

Define $\mathcal{C}^0_{2\pi}$ to be the space of functions which are continuous and 2π-periodic (so that $f(2\pi+\theta) = f(\theta)$). There is a theorem similar to Theorem 3.4

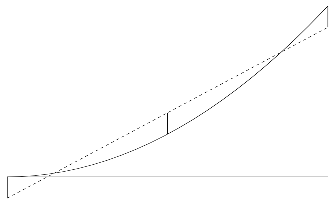


Figure 3.1: Minimax linear approximation to x^2 on range $[0, 1]$

which holds for approximation of a continuous function by a *trigonometric polynomial*, such as

$$q_n(\theta) = a_0 + \sum_{k=1}^{n} (a_k \cos k\theta + b_k \sin k\theta) \tag{3.19}$$

on the range $[-\pi, \pi]$ of θ.

Theorem 3.5 (Alternation theorem for trigonometric polynomials) *For any $f(\theta)$ in $\mathcal{C}^0_{2\pi}$, the minimax approximation $q_n(\theta)$ of form (3.19) exists and is uniquely characterised by an alternating property at $2n + 2$ points of $[-\pi, \pi]$. If $b_1, \ldots, b_n$ (or $a_0, \ldots, a_n$) are set to zero, so that $q_n(\theta)$ is a sum of cosine (or sine) functions alone, and if $f(\theta)$ is an even (or odd) function, then the minimax approximation $q_n(\theta)$ is characterised by an alternating property at $n + 2$ (or respectively $n + 1$) points of $[0, \pi]$.*

Finally, we should mention recent work by Peherstorfer (1997, and elsewhere) on minimax polynomial approximation over collections of non-overlapping intervals.

3.3 The minimax property of the Chebyshev polynomials

We already know, from our discussions of Section 2.2, that the Chebyshev polynomial $T_n(x)$ has $n + 1$ extrema, namely

$$x = y_k = \cos \frac{k\pi}{n} \quad (k = 0, 1, \ldots, n). \tag{3.20}$$

Since $T_n(x) = \cos n\theta$ when $x = \cos\theta$ (by definition), and since $\cos n\theta$ attains its maximum magnitude of unity with alternating signs at its extrema, the following property holds.

Lemma 3.6 (Alternating property of $T_n(x)$) *On $[-1, 1]$, $T_n(x)$ attains its maximum magnitude of 1 with alternating signs at precisely $(n+1)$ points, namely the points (3.20).*

Clearly this property has the flavour of the alternation theorem for minimax polynomial approximation, and indeed we can invoke this theorem as follows. Consider the function

$$f(x) = x^n,$$

and consider its minimax polynomial approximation of degree $n-1$ on $[-1,1]$, $p_{n-1}(x)$, say. Then, by Theorem 3.4, $f(x) - p_{n-1}(x) = x^n - p_{n-1}(x)$ must uniquely have the alternating property on $n+1$ points. But $T_n(x)$ has a leading coefficient (of x^n) equal to 2^{n-1} and hence $2^{1-n}T_n(x)$ is of the same form $x^n - p_{n-1}(x)$ with the same alternating property. It follows that

$$x^n - p_{n-1}(x) = 2^{1-n}T_n(x). \tag{3.21}$$

We say that $2^{1-n}T_n(x)$ is a *monic polynomial*, namely a polynomial with unit leading coefficient. The following two corollaries of the alternation theorem now follow.

Corollary 3.4A (of Theorem 3.4) *The minimax polynomial approximation of degree $n-1$ to the function $f(x) = x^n$ on $[-1,1]$ is*

$$p_{n-1}(x) = x^n - 2^{1-n}T_n(x). \tag{3.22}$$

Corollary 3.4B (The minimax property of T_n) *$2^{1-n}T_n(x)$ is the minimax approximation on $[-1,1]$ to the zero function by a monic polynomial of degree n.*

EXAMPLE 3.2: As a specific example of Corollary 3.4B, the minimax monic polynomial approximation of degree $n = 4$ to zero on $[-1,1]$ is

$$2^{-3}T_4(x) = 2^{-3}(8x^4 - 8x^2 + 1) = x^4 - x^2 + 0.125.$$

This polynomial has the alternating property, taking extreme values $+0.125$, -0.125, $+0.125$, -0.125, $+0.125$, respectively, at the 5 points $y_k = \cos k\pi/4$ $(k = 0, 1, \ldots, 4)$, namely

$$y_k = 1,\ \frac{1}{\sqrt{2}},\ 0,\ -\frac{1}{\sqrt{2}},\ -1. \tag{3.23}$$

Moreover, by Corollary 3.4A, the minimax cubic polynomial approximation to the function $f(x) = x^4$ on $[-1,1]$ is, from (3.22),

$$p_3(x) = x^4 - (x^4 - x^2 + 0.125) = x^2 - 0.125, \tag{3.24}$$

the error $f(x) - p_3(x)$ having the alternating property at the points (3.23). Thus the minimax cubic polynomial approximation in fact reduces to a quadratic polynomial in this case.

It is noteworthy that $x^2 - 0.125$ is also the minimax quadratic polynomial ($n = 2$) approximation to x^4 on $[-1,1]$. The error still has 5 extrema, and so in this case the alternation theorem holds with $n+3$ alternation points. It is thus certainly possible for the number of alternation points to exceed $n+2$.

If the interval of approximation is changed to $[0, 1]$, then a shifted Chebyshev polynomial is required. Thus the minimax monic polynomial approximation of degree n to zero on $[0, 1]$ is

$$2^{1-2n} T_n^*(x). \tag{3.25}$$

For example, for $n = 2$, the minimax monic quadratic is

$$2^{-3} T_2^*(x) = 2^{-3}(8x^2 - 8x + 1) = x^2 - x + 0.125.$$

This is precisely the example (3.18) that was first used to illustrate Theorem 3.4 above.

3.3.1 Weighted Chebyshev polynomials of second, third and fourth kinds

We saw above that the minimax property of $T_n(x)$ depended on the alternating property of $\cos n\theta$. However, an alternating property holds at $n+1$ points θ in $[0, \pi]$ for each of the trigonometric polynomials

$$\begin{array}{lll} \sin(n+1)\theta\,, & \text{at } \theta = \dfrac{(k+\frac{1}{2})\pi}{n+1} & (k = 0, \ldots, n), \\ \cos(n+\frac{1}{2})\theta, & \text{at } \theta = \dfrac{k\pi}{n+\frac{1}{2}} & (k = 0, \ldots, n), \\ \sin(n+\frac{1}{2})\theta, & \text{at } \theta = \dfrac{(k+\frac{1}{2})\pi}{n+\frac{1}{2}} & (k = 0, \ldots, n). \end{array}$$

The following properties may therefore readily be deduced from the definitions (1.4), (1.8) and (1.9) of $U_n(x)$, $V_n(x)$, $W_n(x)$.

Corollary 3.5A (of Theorem 3.5) *(Weighted minimax properties of U_n, V_n, W_n)*

The minimax approximations to zero on $[-1, 1]$, by monic polynomials of degree n weighted respectively by $\sqrt{1-x^2}$, $\sqrt{1+x}$ and $\sqrt{1-x}$, are

$$2^{-n}U_n(x),\ 2^{-n}V_n(x) \text{ and } 2^{-n}W_n(x).$$

The characteristic equioscillation may be seen in Figure 3.2.

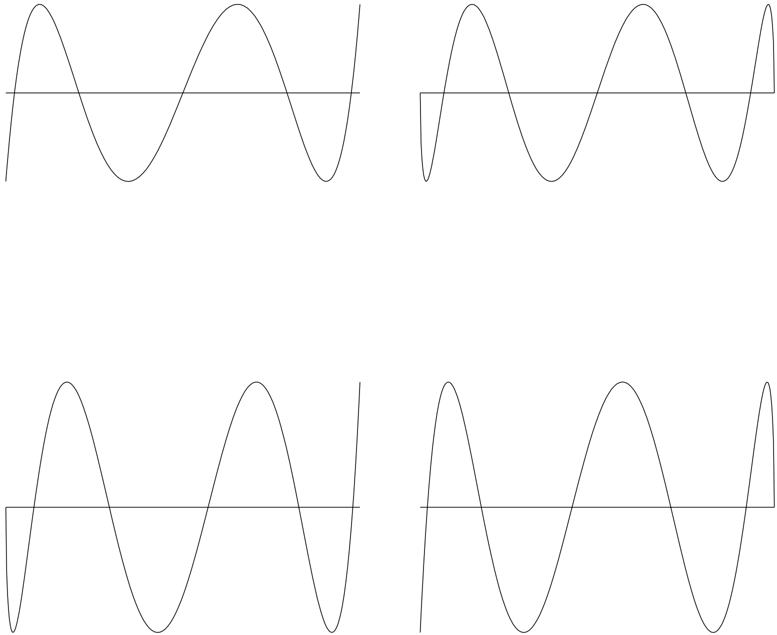


Figure 3.2: Equioscillation on $[-1, 1]$ of $T_5(x)$, $\sqrt{1-x^2}U_5(x)$, $\sqrt{1+x}V_5(x)$ and $\sqrt{1-x}W_5(x)$

3.4 The Chebyshev semi-iterative method for linear equations

The minimax property of the Chebyshev polynomials T_n has been exploited to accelerate the convergence of iterative solutions of linear algebraic equations (Varga 1962, p.138), (Golub & van Loan 1983, p.511).

Let a set of linear equations be written in matrix form as

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{3.26}$$

Then a standard method of solution is to express the square matrix $\mathbf{A}$ in the form $\mathbf{A} = \mathbf{M} - \mathbf{N}$, where the matrix $\mathbf{M}$ is easily inverted (e.g., a diagonal or banded matrix), to select an initial vector $\mathbf{x}_0$, and to perform the iteration

$$\mathbf{M}\mathbf{x}_{k+1} = \mathbf{N}\mathbf{x}_k + \mathbf{b}. \tag{3.27}$$

This iteration will converge to the solution $\mathbf{x}$ of (3.26) if the *spectral radius* $\rho(\mathbf{G})$ of the matrix $\mathbf{G} = \mathbf{M}^{-1}\mathbf{N}$ (absolute value of its largest eigenvalue) is less than unity, converging at a geometric rate proportional to $\rho(\mathbf{G})^k$.

Now suppose that we replace each iterate $\mathbf{x}_k$ by a linear combination of successive iterates:

$$\mathbf{y}_k = \sum_{j=0}^{k} \nu_j(k)\mathbf{x}_j \tag{3.28}$$

where

$$\sum_{j=0}^{k} \nu_j(k) = 1, \tag{3.29}$$

and write

$$p_k(z) := \sum_{j=0}^{k} \nu_j(k) z^j,$$

so that $p_k(1) = 1$.

From (3.26) and (3.27), we have $\mathbf{M}(\mathbf{x}_{j+1} - \mathbf{x}) = \mathbf{N}(\mathbf{x}_j - \mathbf{x})$, so that

$$\mathbf{x}_j - \mathbf{x} = \mathbf{G}^j(\mathbf{x}_0 - \mathbf{x})$$

and, substituting in (3.28) and using (3.29),

$$\mathbf{y}_k - \mathbf{x} = \sum_{j=0}^{k} \nu_j(k)\mathbf{G}^j(\mathbf{x}_0 - \mathbf{x}) = p_k(\mathbf{G})(\mathbf{x}_0 - \mathbf{x}), \tag{3.30}$$

where $p_k(\mathbf{G})$ denotes the matrix $\sum_{j=0}^{k} \nu_j(k)\mathbf{G}^j$.

Assume that the matrix $\mathbf{G} = \mathbf{M}^{-1}\mathbf{N}$ has all of its eigenvalues $\{\lambda_i\}$ real and lying in the range $[\alpha, \beta]$, where $-1 < \alpha < \beta < +1$. Then $p_k(\mathbf{G})$ has eigenvalues $p_k(\lambda_i)$, and

$$\rho(p_k(\mathbf{G})) = \max_i |p_k(\lambda_i)| \le \max_{\alpha \le \lambda \le \beta} |p_k(\lambda)|\,. \tag{3.31}$$

Let F denote the linear mapping of the interval $[\alpha, \beta]$ onto the interval $[-1, 1]$:

$$F(z) = \frac{2z - \alpha - \beta}{\beta - \alpha} \tag{3.32}$$

and write

$$\mu = F(1) = \frac{2 - \alpha - \beta}{\beta - \alpha}. \tag{3.33}$$

Choose the coefficients $\nu_j(k)$ so that

$$p_k(z) = \frac{T_k(F(z))}{T_k(\mu)}. \tag{3.34}$$

Then $p_k(1) = 1$, as required, and

$$\max_{\alpha \le \lambda \le \beta} |p_k(\lambda)| = \frac{1}{|T_k(\mu)|} = \frac{1}{\cosh(k \operatorname{argcosh} \mu)} \sim 2\mathrm{e}^{-k \operatorname{argcosh} \mu}, \tag{3.35}$$

using (1.33a) here, rather than (1.1), since we know that $\mu > 1$. Convergence of $\mathbf{y}_k$ to $\mathbf{x}$ is therefore rapid, provided that μ is large.

It remains to show that $\mathbf{y}_k$ can be computed much more efficiently than by computing $\mathbf{x}_k$ and evaluating the entire summation (3.28) at every step. We can achieve this by making use of the recurrence (1.3a) in the forms

$$\begin{aligned} T_{k-1}(\mu) &= 2\mu\, T_k(\mu) - T_{k+1}(\mu) \\ T_{k+1}(\boldsymbol{\Gamma}) &= 2\boldsymbol{\Gamma}\, T_k(\boldsymbol{\Gamma}) - T_{k-1}(\boldsymbol{\Gamma}) \end{aligned} \tag{3.36}$$

where

$$\boldsymbol{\Gamma} = F(\mathbf{G}) = \frac{2}{\beta - \alpha}\mathbf{G} - \frac{\beta + \alpha}{\beta - \alpha}. \tag{3.37}$$

From (3.30) we have

$$\begin{aligned} \mathbf{y}_{k+1} - \mathbf{y}_{k-1} &= (\mathbf{y}_{k+1} - \mathbf{x}) - (\mathbf{y}_{k-1} - \mathbf{x}) \\ &= p_{k+1}(\mathbf{G})(\mathbf{x}_0 - \mathbf{x}) - p_{k-1}(\mathbf{G})(\mathbf{x}_0 - \mathbf{x}) \\ &= \left(\frac{T_{k+1}(\boldsymbol{\Gamma})}{T_{k+1}(\mu)} - \frac{T_{k-1}(\boldsymbol{\Gamma})}{T_{k-1}(\mu)} \right) (\mathbf{x}_0 - \mathbf{x}); \\ \mathbf{y}_k - \mathbf{y}_{k-1} &= \left(\frac{T_k(\boldsymbol{\Gamma})}{T_k(\mu)} - \frac{T_{k-1}(\boldsymbol{\Gamma})}{T_{k-1}(\mu)} \right) (\mathbf{x}_0 - \mathbf{x}). \end{aligned}$$

Define

$$\omega_{k+1} = 2\mu \frac{T_k(\mu)}{T_{k+1}(\mu)}. \tag{3.38}$$

Then, using (3.36), the expression

$$(\mathbf{y}_{k+1} - \mathbf{y}_{k-1}) - \omega_{k+1}(\mathbf{y}_k - \mathbf{y}_{k-1})$$

simplifies to

$$\begin{aligned} 2(\mathbf{\Gamma}-\mu)\frac{T_k(\mathbf{\Gamma})}{T_{k+1}(\mu)}(\mathbf{x}_0-\mathbf{x}) &= \omega_{k+1}\frac{\mathbf{\Gamma}-\mu}{\mu}(\mathbf{y}_k-\mathbf{x}) \\ &= \omega_{k+1}\gamma(\mathbf{G}-1)(\mathbf{y}_k-\mathbf{x}) = \omega_{k+1}\gamma\mathbf{z}_k \end{aligned}$$

where

$$\gamma = 2/(2-\alpha-\beta) \tag{3.39}$$

and where $\mathbf{z}_k$ satisfies

$$\begin{aligned} \mathbf{M}\mathbf{z}_k &= \mathbf{M}(\mathbf{G}-1)(\mathbf{y}_k-\mathbf{x}) \\ &= (\mathbf{N}-\mathbf{M})(\mathbf{y}_k-\mathbf{x}) = \mathbf{A}(\mathbf{x}-\mathbf{y}_k) = \mathbf{b}-\mathbf{A}\mathbf{y}_k. \end{aligned} \tag{3.40}$$

The successive iterates $\mathbf{y}_k$ can thus be generated by means of the three-term recurrence

$$\mathbf{y}_{k+1} = \omega_{k+1}(\mathbf{y}_k - \mathbf{y}_{k-1} + \gamma\mathbf{z}_k) + \mathbf{y}_{k-1}, \quad k = 1, 2, \ldots, \tag{3.41}$$

starting from

$$\mathbf{y}_0 = \mathbf{x}_0, \quad \mathbf{y}_1 = \mathbf{y}_0 + \gamma\mathbf{z}_0, \tag{3.42}$$

where

$$\omega_{k+1} = 2\mu\frac{T_k(\mu)}{T_{k+1}(\mu)}, \quad \mu = \frac{2-\alpha-\beta}{\beta-\alpha}, \quad \gamma = \frac{2}{2-\alpha-\beta},$$

and $\mathbf{z}_k$ is at each step the solution of the linear system

$$\mathbf{M}\mathbf{z}_k = \mathbf{b} - \mathbf{A}\mathbf{y}_k. \tag{3.43}$$

Using (1.3a) again, we can generate the coefficients ω_k most easily by means of the recurrence

$$\omega_{k+1} = \frac{1}{1-\omega_k/4\mu^2} \tag{3.44}$$

with $\omega_1 = 2$; they converge to a limit $\omega_k \to 2\mu(\mu - \sqrt{\mu^2-1})$ as $k \to \infty$.

In summary, the algorithm is as follows:

Given the system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, with $\mathbf{A} = \mathbf{M}-\mathbf{N}$, where $\mathbf{M}z = \mathbf{b}$ is easily solved and all eigenvalues of $\mathbf{M}^{-1}\mathbf{N}$ lie on the real subinterval $[\alpha, \beta]$ of $[-1, 1]$:

1. Let $\gamma := \dfrac{2}{2-\alpha-\beta}$ and $\mu := \dfrac{2-\alpha-\beta}{\beta-\alpha}$;
2. Take an arbitrary starting vector $\mathbf{y}_0 := \mathbf{x}_0$;
   Take $\omega_1 := 2$;
   Solve $\mathbf{M}\mathbf{z}_0 = \mathbf{b} - \mathbf{A}\mathbf{y}_0$ for $\mathbf{z}_0$;
   Let $\mathbf{y}_1 := \mathbf{x}_0 + \gamma\mathbf{z}_0$ (3.42);

3. For $k = 1, 2, \ldots$:

   Let $\omega_{k+1} := \dfrac{1}{1 - \omega_k/4\mu^2}$ (3.44);

   Solve $\mathbf{M}\mathbf{z}_k = \mathbf{b} - \mathbf{A}\mathbf{y}_k$ for $\mathbf{z}_k$ (3.43);

   Let $\mathbf{y}_{k+1} := \omega_{k+1}(\mathbf{y}_k - \mathbf{y}_{k-1} + \gamma\mathbf{z}_k) + \mathbf{y}_{k-1}$ (3.41).

3.5 Telescoping procedures for power series

If a function $f(x)$ may be expanded in a power series which converges on $[-1, 1]$ (possibly after a suitable transformation of the x variable), then a plausible approximation may clearly be obtained by truncating this power series after $n+1$ terms to a polynomial $p_n(x)$ of degree n. It may be possible, however, to construct an nth degree polynomial approximation better than this, by first truncating the series to a polynomial $p_m(x)$ of some higher degree $m > n$ (which will usually be a better approximation to $f(x)$ than $p_n(x)$) and then exploiting the properties of Chebyshev polynomials to 'economise' $p_m(x)$ to a polynomial of degree n.

The simplest economisation technique is based on the idea of subtracting a constant multiple of a Chebyshev polynomial of the same degree, the constant being chosen so as to reduce the degree of the polynomial.

EXAMPLE 3.3: For $f(x) = \mathrm{e}^x$, the partial sum of degree 7 of the power series expansion is given by

$$\begin{aligned} p_7(x) &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^7}{7!} \\ &= 1 + x + 0.5x^2 + 0.1666667x^3 + 0.0416667x^4 + \\ &\quad + 0.008333x^5 + 0.0013889x^6 + 0.0001984x^7, \end{aligned} \tag{3.45}$$

where a bound on the error in approximating $f(x)$ is given, by the mean value theorem, by

$$|f(x) - p_7(x)| = \left|\frac{x^8}{8!} f^{(8)}(\xi)\right| = \left|\frac{x^8}{8!}\mathrm{e}^\xi\right| \le \frac{\mathrm{e}}{8!} = 0.0000674 \text{ for } x \text{ in } [-1, 1]. \tag{3.46}$$

(The actual maximum error on $[-1, 1]$ in this example is in fact the error at $x = 1$, $|f(1) - p_7(1)| = 0.0000279$.)

Now (3.45) may be economised by forming the degree-6 polynomial

$$\begin{aligned} p_6(x) &= p_7(x) - 0.0001984\ [2^{-6}T_7(x)] \\ &= p_7(x) - 0.0000031\ T_7(x). \end{aligned} \tag{3.47}$$

Since $2^{-6}T_7(x)$ is the minimax monic polynomial of degree 7, this means that p_6 is the minimax 6th degree approximation to p_7 on $[-1, 1]$, and p_7 has been economised in an optimal way.

From (3.45), (3.47) and the coefficients in Table C.2, we obtain

$$\begin{aligned} p_6(x) &= p_7(x) - 0.0001984(64x^7 - 112x^5 + 56x^3 - 7x)/2^6 \\ &= p_7(x) - 0.0001984(x^7 - 1.75x^5 + 0.875x^3 - 0.109375x). \end{aligned}$$

Thus

$$\begin{aligned} p_6(x) &= 1 + 1.0000217x + 0.5x^3 + 0.1664931x^3 + \\ &\quad + 0.0416667x^4 + 0.0086805x^5 + 0.0013889x^6. \end{aligned} \tag{3.48}$$

(Since $T_7(x)$ is an odd function of x, coefficients of even powers of x are unchanged from those in $p_7(x)$.) An error has been committed in replacing p_7 by p_6, and, from (3.47), this error is of magnitude 0.0000031 at most (since $|T_7(x)|$ is bounded by 1 on the interval). Hence, from (3.46), the accumulated error in $f(x)$ satisfies

$$|f(x) - p_6(x)| \le 0.0000674 + 0.0000031 = 0.0000705. \tag{3.49}$$

A further economisation leads to the quintic polynomial

$$\begin{aligned} p_5(x) &= p_6(x) - 0.0013889\ [2^{-5}T_6(x)] \\ &= p_6(x) - 0.0000434\, T_6(x). \end{aligned} \tag{3.50}$$

Here p_5 is the minimax quintic polynomial approximation to p_6. From (3.48), (3.50) and Table C.2, we obtain

$$\begin{aligned} p_5(x) &= p_6(x) - 0.0013889(32x^6 - 48x^2 + 18x^2 - 1)/2^5 \\ &= p_6(x) - 0.0013889(x^6 - 1.5x^4 + 0.5625x^2 - 0.03125). \end{aligned}$$

Thus

$$\begin{aligned} p_5(x) &= 1.0000062 + 1.0000217x + 0.4992188x^2 + \\ &\quad + 0.1664931x^3 + 0.0437500x^4 + 0.0086805x^5 \end{aligned} \tag{3.51}$$

and, since $T_6(x)$ is an even function of x, coefficients of odd powers are unchanged from those in $p_6(x)$. The error in replacing p_6 by p_5 is, from (3.50), at most 0.0000434. Hence, from (3.49), the accumulated error in $f(x)$ now satisfies

$$|f(x) - p_5(x)| \le 0.0000705 + 0.0000434 = 0.0001139. \tag{3.52}$$

Thus the degradation in replacing p_7 (3.45) by p_5 (3.51) is only marginal, increasing the error bound from 0.000067 to 0.000114.

In contrast, the partial sum of degree 5 of the power series (3.45) has a mean-value-theorem error bound of $\left|x^6 e^\xi/6!\right| \le e/6! \sim 0.0038$ on $[-1, 1]$, and the actual maximum error on $[-1, 1]$, attained at $x = 1$, is 0.0016. However, even this is about 15 times as large as (3.52), so that the telescoping procedure based on Chebyshev polynomials is seen to give a greatly superior approximation.

The approximation (3.50) and the 5th degree partial sum of the Taylor series are both too close to e^x for the error to be conveniently shown graphically. However, in Figures 3.3 and 3.4 we show the corresponding approximations of degree 2, where the improved accuracy is clearly visible.

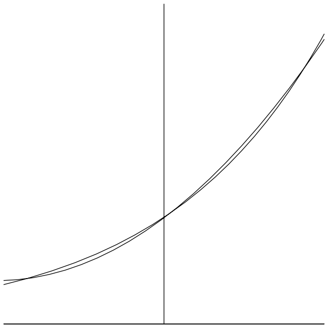


Figure 3.3: The function e^x on $[-1, 1]$ and an economised polynomial approximation of degree 2

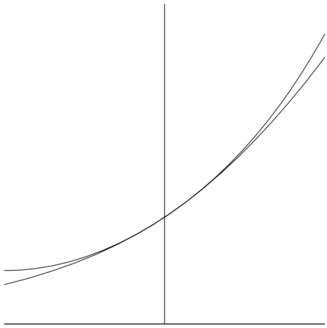


Figure 3.4: The function e^x on $[-1, 1]$ and its Taylor series truncated at the 2nd degree term

An alternative technique which might occur to the reader is to rewrite the polynomial $p_7(x)$, given by (3.45), as a sum of Chebyshev polynomials

$$p_7(x) = \sum_{k=0}^{7} c_k T_k(x), \tag{3.53}$$

where c_k are determined by using, for example, the algorithm of Section 2.3.1 above (powers of x in terms of $\{T_k(x)\}$). Suitable higher order terms, such as those in T_6 and T_7, could then be left out of (3.53) according to the size of their coefficients c_k. However, the telescoping procedure above is exactly equivalent to this, and is in fact a somewhat simpler way of carrying it out. Indeed c_7 and c_6 have been calculated above, in (3.47) and (3.50) respectively, as

$$c_7 = 0.0000031, \quad c_6 = 0.0000434.$$

If the telescoping procedure is continued until a constant approximation $p_0(x)$ is obtained, then all of the Chebyshev polynomial coefficients c_k will be determined.

3.5.1 Shifted Chebyshev polynomials on $[0, 1]$

The telescoping procedure may be adapted to ranges other than $[-1, 1]$, provided that the Chebyshev polynomials are adjusted to the range required. For example, the range $[-c, c]$ involves the use of the polynomials $T_k(x/c)$. A range that is often useful is $[0, 1]$ (or, by scaling, $[0, c]$), and in that case the shifted Chebyshev polynomials $T_k^*(x)$ (or $T_k^*(x/c)$) are used. Since the latter polynomials are neither even nor odd, every surviving coefficient in the polynomial approximation changes at each economisation step.

EXAMPLE 3.4: Suppose that we wish to economise on $[0, 1]$ a quartic approximation to $f(x) = \mathrm{e}^x$:

$$q_4(x) = 1 + x + 0.5x^2 + 0.1666667x^3 + 0.0416667x^4$$

in which the error satisfies

$$|f(x) - q_4(x)| = \frac{x^5}{5!}\mathrm{e}^{\xi} \le \frac{\mathrm{e}}{5!} = 0.0227. \tag{3.54}$$

Then the first economisation step leads to

$$\begin{aligned} q_3(x) &= q_4(x) - 0.0416667\ [2^{-7}T_4^*(x)] \\ &= q_4(x) - 0.0003255\ T_4^*(x). \end{aligned} \tag{3.55}$$

From Table C.2:

$$\begin{aligned} q_3(x) &= q_4(x) - 0.0416667(128x^4 - 256x^3 + 160x^2 - 32x + 1)/2^7 \\ &= q_4(x) - 0.0416667(x^4 - 2x^3 + 1.25x^2 - 0.25x + 0.0078125). \end{aligned}$$

Thus

$$q_3(x) = 0.9996745 + 1.0104166x + 0.4479167x^2 + 0.25x^3. \tag{3.56}$$

Here the maximum additional error due to the economisation is 0.0003255, from (3.55), which is virtually negligible compared with the existing error (3.54) of q_4. In fact, the maximum error of (3.56) on $[0, 1]$ is 0.0103, whereas the maximum error of the power series truncated after the term in x^3 is 0.0516.

The economisation can be continued in a similar way for as many steps as are valid and necessary. It is clear that significantly smaller errors are incurred on $[0, 1]$ by using $T_k^*(x)$ than are incurred on $[-1, 1]$ using $T_k(x)$. This is to be expected, since the range is smaller. Indeed there is always a reduction in error by a factor of 2^m, in economising a polynomial of degree m, since the respective monic polynomials that are adopted are

$$2^{1-m}T_m(x) \text{ and } 2^{1-2m}T_m^*(x).$$

3.5.2 Implementation of efficient algorithms

The telescoping procedures above, based on $T_k(x)$ and $T_k^*(x)$ respectively, are more efficiently carried out in practice by implicitly including the computation of the coefficients of the powers of x in T_k or T_k^* within the procedure (so that Table C.2 does not need to be stored). This is best achieved by using ratios of consecutive coefficients from formula (2.19) of Section 2.3.3 above.

Consider first the use of the shifted polynomial $T_k^*(x/d)$ on a chosen range $[0, d]$. Suppose that $f(x)$ is initially approximated by a polynomial $p_m(x)$ of degree m, where for each $\ell \le m$,

$$p_\ell(x) = \sum_{k=0}^{\ell} a_k^{(\ell)} x^k = a_0^{(\ell)} + a_1^{(\ell)} x + \cdots + a_\ell^{(\ell)} x^\ell. \tag{3.57}$$

Then the first step of the telescoping procedure replaces p_m by

$$p_{m-1}(x) = p_m(x) - a_m^{(m)} 2^{1-2m} d^m T_m^*(x/d). \tag{3.58}$$

(The factor d^m is included, to ensure that $2^{1-2m} d^m T_m^*(x/d)$ is monic.)

Now, write

$$T_m^*(x/d) = 2^{2m-1} d^{-m} \sum d_k^{(m)} x^k d^{m-k} \tag{3.59}$$

where $2^{2m-1} d_k^{(m)}$ is the coefficient of x^k in $T_m^*(x)$. Then, by (3.57), (3.58), (3.59):

$$a_k^{(m-1)} = a_k^{(m)} - a_m^{(m)} d_k^{(m)} \quad (k = m-1, m-2, \ldots, 0). \tag{3.60}$$

The index k has been ordered from $k = m-1$ to $k = 0$ in (3.60), since the coefficients $d_k^{(m)}$ will be calculated in reverse order below.

Now $T_m^*(x) = T_{2m}(x^{\frac{1}{2}})$ and hence, from (2.16),

$$T_m^*(x) = \sum_{k=0}^{m} c_k^{(2m)} x^{m-k} \tag{3.61}$$

where $c_k^{(2m)}$ is defined by (2.17a). Hence, in (3.59),

$$d_k^{(m)} = c_{m-k}^{(2m)} 2^{1-2m}. \tag{3.62}$$

Now, from (2.19)

$$c_{k+1}^{(n)} = -\frac{(n-2k)(n-2k-1)}{4(k+1)(n-k-1)} c_k^{(n)} \tag{3.63}$$

and hence, from (3.62),

$$d_{m-k-1}^{(m)} = -\frac{(2m-2k)(2m-2k-1)}{4(k+1)(2m-k-1)} d_{m-k}^{(m)}.$$

Thus

$$d_{r-1}^{(m)} = \frac{-2r(2r-1)}{4(m-r+1)(m+r-1)} d_r^{(m)} \tag{3.64}$$

where $d_m^{(m)} = 1$.

In summary, the algorithm is as follows:

Given $p_m(x)$ of form (3.57), with coefficients $a_k^{(m)}$:

1. With $d_m^{(m)} = 1$, determine $d_{m-1}^{(m)}, \ldots, d_0^{(m)}$, using (3.64);
2. Determine $a_k^{(m-1)}$, using (3.60), and hence $p_{m-1}(x)$ of form (3.57), with coefficients $a_k^{(m-1)}$.

However, if a telescoping procedure is based on the range $[-d, d]$ and the standard polynomials $T_k(x/d)$, then it is more appropriate to treat even and odd powers of x separately, since each T_k involves only one or the other, and so the algorithm is correspondingly more complicated, but at the same time more efficient.

Suppose $f(x)$ is initially approximated by the polynomial $p_{2M+1}(x)$ of odd degree, where (for each $\ell \leq M$)

$$p_{2\ell+1}(x) = \sum_{k=0}^{\ell} b_k^{(\ell)} x^{2k+1} + \sum_{k=0}^{\ell} c_k^{(\ell)} x^{2k} \tag{3.65a}$$

and

$$p_{2\ell}(x) = \sum_{k=0}^{\ell-1} b_k^{(\ell-1)} x^{2k+1} + \sum_{k=0}^{\ell} c_k^{(\ell)} x^{2k}. \tag{3.65b}$$

Then the first two (odd and even) steps of the telescoping procedure replace $p_{2M+1}(x)$ by $p_{2M}(x)$ and $p_{2M}(x)$ by $p_{2M-1}(x)$, where

$$p_{2M}(x) = p_{2M+1}(x) - b_M^{(M)} 2^{-2M} d^{2M+1} T_{2M+1}(x/d), \tag{3.66a}$$

$$p_{2M-1}(x) = p_{2M}(x) - c_M^{(M)} 2^{1-2M} d^{2M} T_{2M}(x/d). \tag{3.66b}$$

Now, let $2^{2M} e_k^{(M)}$ and $2^{2M-1} f_k^{(M)}$ denote respectively the coefficients of x^{2k+1} in $T_{2M+1}(x)$ and of x^{2k} in $T_{2M}(x)$.

Then, from (2.16),

$$\begin{aligned} T_{2M+1}(x/d) &= 2^{2M} d^{-2M-1} \sum_{k=0}^{M} e_k^{(M)} x^{2k+1} d^{2M-2k} = \\ &= \sum_{k=0}^{M} b_{M-k}^{(M)} x^{2k+1} d^{-2k-1} \end{aligned} \tag{3.67a}$$

$$\begin{aligned} T_{2M}(x/d) &= 2^{2M-1} d^{-2M} \sum_{k=0}^{M} f_k^{(M)} x^{2k} d^{2M-2k} = \\ &= \sum_{k=0}^{M} c_{M-k}^{(M)} x^{2k} d^{-2k} \end{aligned} \tag{3.67b}$$

Hence, from (3.65)–(3.67),

$$b_k^{(M-1)} = b_k^{(M)} - b_M^{(M)} e_k^{(M)} \quad (k = M-1, M-2, \ldots, 0), \tag{3.68a}$$

$$c_k^{(M-1)} = c_k^{(M)} - c_M^{(M)} f_k^{(M)} \quad (k = M-1, M-2, \ldots, 0). \tag{3.68b}$$

Formulae for generating the scaled Chebyshev coefficients $e_k^{(M)}$ and $f_k^{(M)}$ may be determined from (3.63) and (3.67) (by replacing n by $2M+1$, $2M$, respectively) in the form

$$e_{M-k-1}^{(M)} = -\frac{(2M-2k+1)(2M-2k)}{4(k+1)(2M-k)} e_{M-k}^{(M)},$$

$$f_{M-k-1}^{(M)} = -\frac{(2M-2k)(2M-2k-1)}{4(k+1)(2M-k-1)} f_{M-k}^{(M)}.$$

Thus $e_M^{(M)} = f_M^{(M)} = 1$, and

$$e_{r-1}^{(M)} = -\frac{(2r+1)(2r)}{4(M-r+1)(M+r)} e_r^{(M)}, \tag{3.69a}$$

$$f_{r-1}^{(M)} = -\frac{(2r)(2r-1)}{4(M-r+1)(M+r-1)} f_r^{(M)}. \tag{3.69b}$$

In summary, the algorithm is as follows:

Given $p_{2M+1}(x)$ of form (3.65a), with coefficients $b_k^{(M)}$ and $c_k^{(M)}$:

1. With $e_M^{(M)} = 1$, determine $e_{M-1}^{(M)}, \ldots, e_0^{(M)}$, using (3.69a);
2. Determine $b_k^{(M-1)}$, using (3.68a), and hence $p_{2M}(x)$ of form (3.65b), with coefficients $b_k^{(M-1)}$ and $c_k^{(M)}$;
3. With $f_M^{(M)} = 1$, determine $f_{M-1}^{(M)}, \ldots, f_0^{(M)}$, using (3.69b);
4. Determine $c_k^{(M-1)}$, using (3.68b), and hence $p_{2M-1}(x)$ of form (3.65a), with coefficients $b_k^{(M-1)}$ and $c_k^{(M-1)}$.

We should add as a postscript that Gutknecht & Trefethen (1982) have succeeded in implementing an alternative economisation method due to Carathéodory and Fejér, which yields a Chebyshev sum giving a much closer approximation to the original polynomial.

3.6 The tau method for series and rational functions

Sometimes a power series converges very slowly at a point of interest, or even diverges, so that we cannot find a suitable partial sum to provide an initial approximation for the above telescoping procedure. However, in some cases other approaches are useful, one of which is the 'tau' (τ) method[1] of Lanczos (1957).

Consider for example the function

$$y(x) = \frac{1}{1+x}$$

which has the power series expansion

$$1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + \cdots.$$

This series has radius of convergence 1, and since it does not converge for $|x| \geq 1$, cannot be used on $[0, 1]$ or wider ranges. However, $y(x)$ is the solution of the functional equation

$$(1 + x)y(x) = 1 \tag{3.70}$$

and may be approximated on $[0, 1]$ by a polynomial $p_n(x)$ of degree n in the form

$$p_n(x) = {\sum_{k=0}^{n}}{}' c_k T_k^*(x) \tag{3.71}$$

(where, as previously, the dash denotes that the first term in the sum is halved), by choosing the coefficients $\{c_k\}$ so that p_n approximately satisfies the equation

$$(1 + x)p_n(x) = 1. \tag{3.72}$$

Equation (3.72) can be perturbed slightly into one that can be satisfied exactly, by adding to the right-hand side an undetermined multiple τ (say) of a shifted Chebyshev polynomial of degree $n + 1$:

$$(1 + x)p_n(x) = 1 + \tau T_{n+1}^*(x). \tag{3.73}$$

Since there are $n + 2$ free parameters in (3.71) and (3.73), namely c_k ($k = 0, 1, \ldots, n$) and τ, it should be possible to determine them by equating coefficients of powers of x in (3.73) (since there are $n+2$ coefficients in a polynomial of degree $n+1$). Equivalently, we may equate coefficients of Chebyshev polynomials after writing the two sides of (3.73) as Chebyshev summations; this can be done if we note from (2.39) that

$$(2x - 1)T_k(2x - 1) = \tfrac{1}{2}[T_{k+1}(2x - 1) + T_{|k-1|}(2x - 1)]$$

[1] A slightly different but related approach, also known as the 'tau method', is applied to solve differential equations in a later chapter (see Chapter 10).

and hence, since $T_k^*(x) = T_k(2x-1)$,

$$(1+x)T_k^*(x) = \tfrac{1}{4}[T_{k+1}^*(x) + 6T_k^*(x) + T_{|k-1|}^*(x)]. \tag{3.74}$$

Substituting (3.74) into (3.71) and (3.73),

$$\sum_{k=0}^{n}{}' \tfrac{1}{4}c_k[T_{|k-1|}^*(x) + 6T_k^*(x) + T_{k+1}^*(x)] = T_0^*(x) + \tau T_{n+1}^*(x).$$

On equating coefficients of $T_0^*, \ldots, T_{n+1}^*$, we obtain

$$\begin{aligned}
\tfrac{1}{4}(3c_0 + c_1) &= 1,\\
\tfrac{1}{4}(c_{k-1} + 6c_k + c_{k+1}) &= 0 \quad (k = 1, \ldots, n-1),\\
\tfrac{1}{4}(c_{n-1} + 6c_n) &= 0,\\
\tfrac{1}{4}c_n &= \tau.
\end{aligned}$$

These are $n+2$ equations for $c_0, c_1, \ldots, c_n$ and τ, which may be readily solved by back-substituting for c_n in terms of τ, hence (working backwards) determining c_{n-1} , $c_{n-2}, \ldots, c_0$ in terms of τ, leaving the first equation to determine the value of τ.

Example 3.5: For $n = 3$, we obtain (in this order)

$$\begin{aligned}
c_3 &= 4\tau,\\
c_2 &= -6c_3 = -24\tau,\\
c_1 &= -6c_2 - c_3 = 140\tau,\\
c_0 &= -6c_1 - c_2 = -816\tau,\\
3c_0 + c_1 &= -2308\tau = 4.
\end{aligned}$$

Hence $\tau = -1/577$ and, from (3.71),

$$\begin{aligned}
y(x) &\simeq p_3(x)\\
&= \tfrac{1}{577}[408T_0^*(x) - 140T_1^*(x) + 24T_2^*(x) - 4T_3^*(x)]\\
&= 0.707106T_0^*(x) - 0.242634T_1^*(x) +\\
&\quad + 0.041594T_2^*(x) - 0.006932T_3^*(x).
\end{aligned} \tag{3.75}$$

The error $\epsilon(x)$ in (3.75) is known from (3.70) and (3.73) to be

$$\epsilon(x) = y(x) - p_3(x) = \frac{\tau T_4^*(x)}{1+x}.$$

Since $1/(1+x)$ and $T_4^*(x)$ are both bounded by 1 in magnitude, we deduce the bound

$$|\epsilon(x)| \le |\tau| = 0.001704 \simeq 0.002 \text{ on } [0,1]. \tag{3.76}$$

This upper bound is attained at $x = 0$, and we would expect the resulting approximation $p_3(x)$ to be reasonably close to a minimax approximation.

3.6.1 The extended tau method

Essentially the same approach has been proposed by Fox & Parker (1968) for the approximation on $[-1, 1]$ of a rational function $a(x)/b(x)$ of degrees (p, q). They introduce a perturbation polynomial

$$e(x) = \sum_{m=n+1}^{n+q} \tau_{m-n} T_m(x), \tag{3.77}$$

in place of the single term $\tau T_{n+1}(x)$ used above, to give

$$a(x) + e(x) = b(x) \sum_{k=0}^{n}{}' c_k T_k(x). \tag{3.78}$$

The number and degrees of the terms in (3.77) are chosen so that (3.78) is uniquely solvable for $\{c_k\}$ and $\{\tau_m\}$.

For example, for

$$\frac{a(x)}{b(x)} = \frac{1 - x + x^2}{1 + x + x^2}$$

we need two tau terms and (3.78) becomes

$$(1 - x + x^2) + \sum_{m=n+1}^{n+2} \tau_{m-n} T_m(x) = (1 + x + x^2) \sum_{k=0}^{n}{}' c_k T_k(x). \tag{3.79}$$

Both sides of (3.79) are then written in terms of Chebyshev polynomials, and on equating coefficients, a set of equations is obtained for c_k and τ_m. Back-substitution in terms of τ_1 and τ_2 leads to a pair of simultaneous equations for τ_1 and τ_2; hence c_k are found.

EXAMPLE 3.6: For $n = 2$, (3.79) becomes, using (2.38) to transform products into sums,

$$\begin{aligned}
(\tfrac{3}{2}T_0(x) \; - \; & T_1(x) + \tfrac{1}{2}T_2(x)) + \tau_1 T_3(x) + \tau_2 T_4(x) \\
&= (\tfrac{3}{2}T_0(x) + T_1(x) + \tfrac{1}{2}T_2(x))(\tfrac{1}{2}c_0 T_0(x) + c_1 T_1(x) + c_2 T_2(x)) \\
&= (\tfrac{3}{4}c_0 + \tfrac{1}{2}c_1 + \tfrac{1}{4}c_2)T_0(x) + (\tfrac{1}{2}c_0 + \tfrac{7}{4}c_1)T_1(x) + \\
&\qquad + (\tfrac{1}{4}c_0 + \tfrac{1}{2}c_1 + \tfrac{3}{2}c_2)T_2(x) + (\tfrac{1}{4}c_1 + \tfrac{1}{2}c_2)T_3(x) + \tfrac{1}{4}T_4(x).
\end{aligned}$$

Equating coefficients of the Chebyshev polynomials $T_0(x)$, ..., $T_4(x)$ yields the equations

$$\left.\begin{aligned} c_2 &= 4\tau_2 \\ c_1 + 2c_2 &= 4\tau_1 \\ c_0 + 2c_1 + 6c_2 &= 2 \end{aligned}\right\} \tag{3.80}$$

$$\left.\begin{aligned} 2c_0 + 7c_1 &= -4 \\ 3c_0 + 2c_1 + c_2 &= 6 \end{aligned}\right\} \tag{3.81}$$

Back-substituting in (3.80):

$$c_2 = 4\tau_2,\ c_1 = 4\tau_1 - 8\tau_2,\ c_0 = 2 - 8\tau_1 - 8\tau_2.$$

Now (3.81) gives

$$\begin{aligned} 3\tau_1 - 16\tau_2 &= -2 \\ 9\tau_1 + 4\tau_2 &= 0 \end{aligned}$$

and hence

$$\tau_1 = -18/91,\ \tau_2 = 8/91$$

so that

$$c_0 = 262/91,\ c_1 = -136/91,\ c_2 = 32/91.$$

Thus

$$y(x) = \frac{a(x)}{b(x)} = \frac{1 - x + x^2}{1 + x + x^2} \simeq p_3(x) = 1.439\, T_0(x) - 1.494\, T_1(x) + 0.352\, T_2(x)$$

and the error is given by

$$\begin{aligned} \epsilon(x) = y(x) - p_3(x) &= -\frac{\tau_1 T_3(x) + \tau_2 T_4(x)}{1 + x + x^2} \\ &= \frac{0.198\, T_3(x) - 0.088\, T_4(x)}{1 + x + x^2}. \end{aligned}$$

On $[-1, 1]$, $1/(1 + x + x^2)$ is bounded by $\frac{4}{3}$ and $|T_3|$ and $|T_4|$ are bounded by 1. Hence we have the bound (which is not far from the actual maximum error)

$$|\epsilon(x)| < 1.333\ (0.198 + 0.088) = 0.381.$$

With an error bound of 0.381, the approximation found in this example is not particularly accurate, and indeed a much higher degree of polynomial is needed to represent such a rational function at all reasonably, but the method does give credible and measurable results even in this simple case (see Figure 3.5).

We may note that an alternative approach to the whole calculation is to use the power form for the polynomial approximation

$$p_n(x) = \sum_{k=0}^{n} a_k x^k \tag{3.82}$$

and then to replace (3.79) by

$$(1 - x + x^2) + \sum_{m=n+1}^{n+2} \tau_{m-n} T_m(x) = (1 + x + x^2) \sum_{k=0}^{n} a_k x^k. \tag{3.83}$$

We then equate coefficients of powers of x and solve for τ_1 and τ_2.

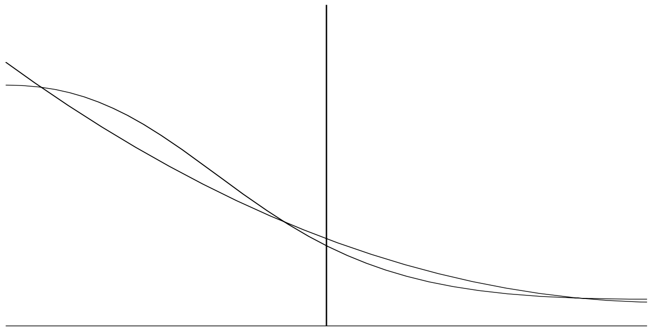


Figure 3.5: Rational function and a quadratic approximation obtained by the extended τ method

---

Example 3.7: For $n = 2$, equation (3.83) takes the form

$$(1 - x + x^2) + \tau_1(4x^3 - 3x) + \tau_2(8x^4 - 8x^2 + 1) = (1 + x + x^2)(a_0 + a_1 x + a_2 x^2),$$

and on equating coefficients of $1, x, \ldots, x^4$,

$$\left.\begin{aligned} a_0 &= 1 + \tau_2 \\ a_1 + a_0 &= -1 - 3\tau_1 \end{aligned}\right\} \tag{3.84}$$

$$\left.\begin{aligned} a_2 + a_1 + a_0 &= 1 - 8\tau_2 \\ a_2 + a_1 &= 4\tau_1 \\ a_2 &= 8\tau_2 \end{aligned}\right\} \tag{3.85}$$

Back-substituting in (3.85):

$$a_2 = 8\tau_2,\ a_1 = 4\tau_1 - 8\tau_2,\ a_0 = 1 - 4\tau_1 - 8\tau_2.$$

Now (3.84) gives

$$4\tau_1 + 9\tau_2 = 0$$

$$3\tau_1 - 16\tau_2 = -2$$

and hence

$$\tau_1 = -18/91,\ \tau_2 = 8/91$$

(the same values as before) so that

$$a_0 = 99/91,\ a_1 = -136/91,\ a_2 = 64/91.$$

Thus

$$y(x) = \frac{a(x)}{b(x)} = \frac{1 - x + x^2}{1 + x + x^2} \simeq p_3(x) = 1.088 - 1.495x + 0.703x^2.$$

It is easily verified that this is precisely the same approximation as was obtained previously, but expressed explicitly as a sum of powers of x.

For the degrees n of polynomial likely to be required in practice, it is not advisable to use the power representation (3.82), even though the algebra appears simpler, since the coefficients a_k tend to become large as n increases, whereas the Chebyshev coefficients c_k in the form (3.71) typically tend to converge with n to the true coefficients of an infinite Chebyshev series expansion (see Chapter 4).

3.7 Problems for Chapter 3

1. Verify the axioms of a vector space for the following families of functions or data:

   (a) $\mathcal{F} = \mathcal{C}[a, b]$;

   (b) $\mathcal{F} = \{\{f(x_k),\ k = 1, \ldots, m\}\}$ (values of a function at discrete points).

   What are the dimensions of these spaces?

2. Verify, from the definition of a norm, that the following is a norm:

$$\|f\| = \|f\|_p = \left[\int_a^b |f(x)|^p \,\mathrm{d}x\right]^{\frac{1}{p}} \qquad (1 \le p < \infty),$$

   by assuming Minkowski's continuous inequality:

$$\left(\int |f+g|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \le \left(\int |f|^p \,\mathrm{d}x\right)^{\frac{1}{p}} + \left(\int |g|^p \,\mathrm{d}x\right)^{\frac{1}{p}}.$$

   Prove the latter inequality for $p = 1, 2$, and show, for $p = 2$, that equality does not occur unless $f(x) = \lambda g(x)$ ('almost everywhere'), where λ is some constant.

3. For what values of p does the function $f(x) = (1 - x^2)^{-1/2}$ belong to the function space $\mathcal{L}_p[-1, 1]$, and what is its norm?

4. Prove Minkowski's discrete inequality:

$$\left(\sum_k |u_k + v_k|^p\right)^{\frac{1}{p}} \le \left(\sum_k |u_k|^p\right)^{\frac{1}{p}} + \left(\sum_k |v_k|^p\right)^{\frac{1}{p}}$$

   in the case $p = 2$ by first showing that

$$\left(\sum u_k v_k\right)^2 \le \sum (u_k)^2 \sum (v_k)^2.$$

Deduce that

$$\|f\|_p = \left[\sum_{k=1}^{m} |f(x_k)|^p\right]^{\frac{1}{p}} \qquad (1 \le p < \infty)$$

is a norm for space (b) of Problem 1.

Find proofs in the literature (Hardy et al. 1952, for example) of both continuous and discrete Minkowski inequalities for general p. Can equality occur for $p = 1$?

5. Find the minimax constant (i.e., polynomial of degree zero) approximation to e^x on $[-1, 1]$, by assuming that its error has the alternating property at $-1, +1$. Deduce that the minimax error in this case is $\sinh 1$.

   Generalise the above approach to determine a minimax constant approximation to any monotonic continuous function $f(x)$.

6. Prove the *sufficiency* of the characterisation of the error in Theorem 3.4, namely that, for a polynomial approximation p_n of degree n to a continuous f to be minimax, it is sufficient that it should have the alternating property at $n + 2$ points $x_1 < \cdots < x_{n+2}$.

   [Hint: Assume that an approximation p'_n exists with smaller error norm than p_n, show that $p_n - p'_n$ changes sign between each pair x_i and x_{i+1}, and hence obtain the result.]

7. Consider the function

$$f(x) = \sum_{i=0}^{\infty} c_i T_{b^i}(x), \qquad (*)$$

   where $\{c_i\}$ are so defined that the series is uniformly convergent and where b is an odd integer not less than 2. Show that, for every $i > n$ with n fixed, T_{b^i} has the alternating property on a set of $b^n + 1$ consecutive points of $[-1, 1]$. Deduce that the partial sum of degree b^n of (*) (namely the sum from $i = 0$ to n) is the minimax polynomial approximation of degree b^n to $f(x)$.

   [Note: A series in $\{T_k(x)\}$ such as (*) in which terms occur progressively more rarely (in this case for $k = 0, b, b^2, b^3, \ldots$) is called *lacunary*; see Section 5.9 below for a fuller discussion.]

8. For $f(x) = \arctan x$, show that $(1 + x^2)f'(x) = 1$, and hence that

$$(1+x^2)f^{(n)}(x)+2x(n-1)f^{(n-1)}(x)+(n+1)(n+2)f^{(n-2)}(x) = 0 \quad (n \ge 2).$$

   Deduce the Taylor–Maclaurin expansion

$$f(x) \sim x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots. \qquad (**)$$

Estimate the error in the partial sum $P_7(x)$ of degree 7 of (**) for x in $[-0.3, 0.3]$.

Telescope P_7, into polynomials P_5 of degree 5 and P_3 of degree 3 by using Chebyshev polynomials normalised to $[-0.3, 0.3]$, and estimate the accumulated errors in P_5 and P_3.

9. Given

$$f(x) = \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots, \qquad (***)$$

use the mean value theorem to give a bound on the error on $[0, 0.1]$ of the partial sum P_n of degree n of (***). Telescope P_4 into polynomials P_3 of degree 3 and P_2 of degree 2, respectively, using a Chebyshev polynomial adjusted to $[0, 0.1]$, and estimate the accumulated errors in each case.

10. (Programming Exercise) Write a computer program (in a programming language of your own choice) to implement the telescoping algorithm of Section 3.5, either

    (a) based on $T_k^*(x/d)$ and using (3.60)–(3.64) or

    (b) based on $T_k(x/d)$ and using (3.68)–(3.69).

11. Apply the tau method of Section 3.6 to determine a polynomial approximation of degree 3 to $x/(1+x)$ on $[0, 1]$ based on the equation

$$(1+x)y = x$$

and determine a bound on the resulting error.

12. Apply the extended tau method of Section 3.6.1 to determine a polynomial approximation of degree 2 to $(1+x+x^2)^{-1}$ on $[-1, 1]$ and determine a bound on the resulting error.

13. Show that $2^{-n}\sqrt{1-x^2}U_n(x)$, $2^{-n}\sqrt{1+x}V_n(x)$ and $2^{-n}\sqrt{1-x}W_n(x)$ equioscillate on $(n+2)$, $(n+1)$ and $(n+1)$ points, respectively, of $[-1, 1]$, and find the positions of their extrema. Deduce that these are minimax approximations to zero by monic polynomials of degree n with respective weight functions $\sqrt{1-x^2}$, $\sqrt{1+x}$, $\sqrt{1-x}$. Why are there more equioscillation points in the first case?

Chapter 4

Orthogonality and Least-Squares Approximation

4.1 Introduction — from minimax to least squares

The Chebyshev polynomials have been shown in Chapter 3 to be unique among all polynomials in possessing a minimax property (Corollaries 3.4B, 3.5A), earning them a central role in the study of uniform (or $\mathcal{L}_\infty$) approximation. This property is remarkable enough, but the four families of Chebyshev polynomials have a second and equally important property, in that each is a family of orthogonal polynomials. Thus, the Chebyshev polynomials have an important role in $\mathcal{L}_2$ or least-squares approximation, too. This link with $\mathcal{L}_2$ approximation is important in itself but, in addition, it enables ideas of orthogonality to be exploited in such areas as Chebyshev series expansions and Galerkin methods for differential equations.

Orthogonal polynomials have a great variety and wealth of properties, many of which are noted in this chapter. Indeed, some of these properties take a very concise form in the case of the Chebyshev polynomials, making Chebyshev polynomials of leading importance among orthogonal polynomials — second perhaps to Legendre polynomials (which have a unit weight function), but having the advantage over the Legendre polynomials that the locations of their zeros are known analytically. Moreover, along with the Legendre polynomials, the Chebyshev polynomials belong to an exclusive band of orthogonal polynomials, known as *Jacobi polynomials*, which correspond to weight functions of the form $(1-x)^\alpha(1+x)^\beta$ and which are solutions of Sturm–Liouville equations.

The Chebyshev polynomials have further properties, which are peculiar to them and have a trigonometric origin, namely various kinds of discrete orthogonality over the zeros of Chebyshev polynomials of higher degree. In consequence, interpolation at Chebyshev zeros can be achieved exceptionally inexpensively (Chapter 6) and Gauss quadrature methods based on Chebyshev zeros are extremely convenient (Chapter 8).

The continuous and discrete orthogonality of the Chebyshev polynomials may be viewed as a direct consequence of the orthogonality of sine and cosine functions of multiple angles, a central feature in the study of Fourier series. It is likely, therefore, that a great deal may be learned about Chebyshev series by studying their links with Fourier series (or, in the complex plane, Laurent series); this is considered in Chapter 5.

Finally, the Chebyshev polynomials are orthogonal not only as polynomials in the real variable x on the real interval $[-1, 1]$ but also as polynomials in a complex variable z on elliptical contours and domains of the complex plane (the foci of the ellipses being at -1 and $+1$). This property is exploited in fields such as crack problems in fracture mechanics (Gladwell & England 1977) and two-dimensional aerodynamics (Fromme & Golberg 1979, Fromme & Golberg 1981), which rely on complex-variable techniques. More generally, however, many real functions may be extended into analytic functions, and Chebyshev polynomials are remarkably robust in approximating on $[-1, 1]$ functions which have complex poles close to that interval. This is a consequence of the fact that the interval $[-1, 1]$ may be enclosed in an arbitrarily thin ellipse which excludes nearby singularities.

4.2 Orthogonality of Chebyshev polynomials

4.2.1 Orthogonal polynomials and weight functions

Definition 4.1 *Two functions $f(x)$ and $g(x)$ in $\mathcal{L}_2[a, b]$ are said to be* orthogonal *on the interval $[a, b]$ with respect to a given continuous and non-negative weight function $w(x)$ if*

$$\int_a^b w(x)f(x)g(x)\,\mathrm{d}x = 0. \tag{4.1}$$

If, for convenience, we use the 'inner product' notation

$$\langle f\,,\,g\rangle = \int_a^b w(x)f(x)g(x)\,\mathrm{d}x, \tag{4.2}$$

where w, f and g are functions of x on $[a, b]$, then the orthogonality condition (4.1) is equivalent to saying that f is orthogonal to g if

$$\langle f\,,\,g\rangle = 0. \tag{4.3}$$

The formal definition of an inner product (in the context of real functions of a real variable — see Definition 4.3 for the complex case) is as follows:

Definition 4.2 *An* inner product $\langle\cdot\,,\,\cdot\rangle$ *is a bilinear function of elements $f, g, h, \ldots$ of a vector space that satisfies the axioms:*

1. $\langle f\,,\,f\rangle \geq 0$ *with equality if and only if* $f \equiv 0$;
2. $\langle f\,,\,g\rangle = \langle g\,,\,f\rangle$;
3. $\langle f + g\,,\,h\rangle = \langle f\,,\,h\rangle + \langle g\,,\,h\rangle$;
4. $\langle \alpha f\,,\,g\rangle = \alpha\,\langle f\,,\,g\rangle$ *for any scalar* α.

An inner product defines an $\mathcal{L}_2$-type norm

$$\|f\| = \|f\|_2 := \sqrt{\langle f\,,\,f\rangle}. \tag{4.4}$$

We shall adopt the inner product (4.2) (with various weight functions) and the associated $\mathcal{L}_2$ norm (4.4), which is identical to that defined in Chapter 3 (3.4), through most of the remainder of this chapter.

Here we shall in particular be concerned with families of orthogonal polynomials $\{\phi_i(x),\ i = 0, 1, 2, \ldots\}$ where ϕ_i is of degree i exactly, defined so that

$$\langle \phi_i\,,\,\phi_j\rangle = 0 \quad (i \neq j). \tag{4.5}$$

Clearly, since $w(x)$ is non-negative,

$$\langle \phi_i\,,\,\phi_i\rangle = \|\phi_i\|^2 > 0. \tag{4.6}$$

The requirement that ϕ_i should be of exact degree i, together with the orthogonality condition (4.5), defines each polynomial ϕ_i uniquely apart from a multiplicative constant (see Problem 3). The definition may be made unique by fixing the value of $\langle \phi_i\,,\,\phi_i\rangle$ or of its square root $\|\phi_i\|$. In particular, we say that the family is *orthonormal* if, in addition to (4.5), the functions $\{\phi_i(x)\}$ satisfy

$$\|\phi_i\| = 1 \text{ for all } i. \tag{4.7}$$

4.2.2 Chebyshev polynomials as orthogonal polynomials

If we define the inner product (4.2) using the interval and weight function

$$[a, b] = [-1, 1], \quad w(x) = (1 - x^2)^{-\frac{1}{2}}, \tag{4.8}$$

then we find that the first kind Chebyshev polynomials satisfy

$$\begin{aligned} \langle T_i\,,\,T_j\rangle &= \int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}}\,\mathrm{d}x \\ &= \int_0^{\pi} \cos i\theta\ \cos j\theta\,\mathrm{d}\theta \end{aligned} \tag{4.9}$$

(shown by setting $x = \cos\theta$ and using the relations $T_i(x) = \cos i\theta$ and $\mathrm{d}x = -\sin\theta\,\mathrm{d}\theta = -\sqrt{1-x^2}\,\mathrm{d}\theta$).

Now, for $i \neq j$,

$$\begin{aligned} \int_0^{\pi} \cos i\theta\ \cos j\theta\,\mathrm{d}\theta &= \tfrac{1}{2}\int_0^{\pi} [\cos(i+j)\theta + \cos(i-j)\theta]\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\frac{\sin(i+j)\theta}{i+j} + \frac{\sin(i-j)\theta}{i-j}\right]_0^{\pi} = 0. \end{aligned}$$

Hence

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j), \tag{4.10}$$

and $\{T_i(x),\ i = 0, 1, \ldots\}$ forms an orthogonal polynomial system on $[-1, 1]$ with respect to the weight $(1 - x^2)^{-\frac{1}{2}}$.

The norm of T_i is given by

$$\begin{aligned} \|T_i\|^2 &= \langle T_i\,,\,T_i\rangle \\ &= \int_0^\pi (\cos i\theta)^2\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\int_0^\pi (1 + \cos 2i\theta)\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\theta + \frac{\sin 2i\theta}{2i}\right]_0^\pi \quad (i \neq 0) \\ &= \tfrac{1}{2}\pi, \end{aligned} \tag{4.11a}$$

while

$$\|T_0\|^2 = \langle T_0\,,\,T_0\rangle = \langle 1\,,\,1\rangle = \pi. \tag{4.11b}$$

The system $\{T_i\}$ is therefore not orthonormal. We could, if we wished, scale the polynomials to derive the orthonormal system

$$\sqrt{1/\pi}\,T_0(x), \left\{\sqrt{2/\pi}\,T_i(x),\ i = 1, 2, \ldots\right\},$$

but the resulting irrational coefficients usually make this inconvenient. It is simpler in practice to adopt the $\{T_i\}$ we defined initially, taking note of the values of their norms (4.11).

The second, third and fourth kind Chebyshev polynomials are also orthogonal systems on $[-1, 1]$, with respect to appropriate weight functions:

- $U_i(x)$ are orthogonal with respect to $w(x) = (1 - x^2)^{\frac{1}{2}}$;
- $V_i(x)$ are orthogonal with respect to $w(x) = (1 + x)^{\frac{1}{2}}(1 - x)^{-\frac{1}{2}}$;
- $W_i(x)$ are orthogonal with respect to $w(x) = (1 + x)^{-\frac{1}{2}}(1 - x)^{\frac{1}{2}}$.

These results are obtained from trigonometric relations as follows (using the appropriate definition of $\langle\cdot\,,\,\cdot\rangle$ in each case):

$$\begin{aligned} \langle U_i\,,\,U_j\rangle &= \int_{-1}^1 (1 - x^2)^{\frac{1}{2}}\,U_i(x)U_j(x)\,\mathrm{d}x \\ &= \int_{-1}^1 (1 - x^2)^{-\frac{1}{2}}\,(1 - x^2)^{\frac{1}{2}}U_i(x)\,(1 - x^2)^{\frac{1}{2}}U_j(x)\,\mathrm{d}x \end{aligned}$$

$$= \int_0^\pi \sin(i+1)\theta \, \sin(j+1)\theta \, d\theta$$

$$\text{(since } \sin\theta \, U_i(x) = \sin(i+1)\theta)$$

$$= \tfrac{1}{2} \int_0^\pi [\cos(i-j)\theta - \cos(i+j+2)\theta] \, d\theta$$

$$= 0 \quad (i \neq j).$$

$$\begin{aligned}
\langle V_i \,,\, V_j \rangle &= \int_{-1}^1 (1-x)^{-\frac{1}{2}} (1+x)^{\frac{1}{2}} \, V_i(x) V_j(x) \, dx \\
&= \int_{-1}^1 (1-x^2)^{-\frac{1}{2}} \, (1+x)^{\frac{1}{2}} V_i(x) \, (1+x)^{\frac{1}{2}} V_j(x) \, dx \\
&= 2 \int_0^\pi \cos(i+\tfrac{1}{2})\theta \, \cos(j+\tfrac{1}{2})\theta \, d\theta \\
&\qquad \text{(since } (1+x)^{\frac{1}{2}} = (1+\cos\theta)^{\frac{1}{2}} = (2\cos^2 \tfrac{1}{2}\theta)^{\frac{1}{2}} = \sqrt{2}\cos\tfrac{1}{2}\theta \\
&\qquad\quad \text{and } (1+x)^{\frac{1}{2}} V_i(x) = \sqrt{2}\cos(i+\tfrac{1}{2})\theta) \\
&= \int_0^\pi [\cos(i+j+1)\theta + \cos(i-j)\theta] \, d\theta \\
&= 0 \quad (i \neq j).
\end{aligned}$$

$$\begin{aligned}
\langle W_i \,,\, W_j \rangle &= \int_{-1}^1 (1+x)^{-\frac{1}{2}} (1-x)^{\frac{1}{2}} \, W_i(x) W_j(x) \, dx \\
&= \int_{-1}^1 (1-x^2)^{-\frac{1}{2}} \, (1-x)^{\frac{1}{2}} W_i(x) \, (1-x)^{\frac{1}{2}} W_j(x) \, dx \\
&= 2 \int_0^\pi \sin(i+\tfrac{1}{2})\theta \, \sin(j+\tfrac{1}{2})\theta \, d\theta \\
&\qquad \text{(since } (1-x)^{\frac{1}{2}} = (1-\cos\theta)^{\frac{1}{2}} = (2\sin^2 \tfrac{1}{2}\theta)^{\frac{1}{2}} = \sqrt{2}\sin\tfrac{1}{2}\theta \\
&\qquad\quad \text{and } (1-x)^{\frac{1}{2}} W_i(x) = \sqrt{2}\sin(i+\tfrac{1}{2})\theta) \\
&= \int_0^\pi [\cos(i-j)\theta - \cos(i+j+1)\theta] \, d\theta \\
&= 0 \quad (i \neq j).
\end{aligned}$$

The normalisations that correspond to these polynomials are as follows (for all $i \geq 0$):

$$\langle U_i \,,\, U_i \rangle = \|U_i\|^2 = \int_0^\pi \sin^2(i+1)\theta \, d\theta = \tfrac{1}{2}\pi; \tag{4.12}$$

$$\langle V_i\,,\,V_i\rangle = \|V_i\|^2 = 2\int_0^\pi \cos^2(i+\tfrac{1}{2})\theta\,\mathrm{d}\theta = \pi; \tag{4.13}$$

$$\langle W_i\,,\,W_i\rangle = \|W_i\|^2 = 2\int_0^\pi \sin^2(i+\tfrac{1}{2})\theta\,\mathrm{d}\theta = \pi. \tag{4.14}$$

(Remember that each of these three identities uses a different definition of the inner product $\langle\cdot\,,\,\cdot\rangle$, since the weights $w(x)$ differ.)

4.3 Orthogonal polynomials and best $\mathcal{L}_2$ approximations

In Chapter 3, we characterised a best $\mathcal{L}_\infty$ (minimax) polynomial approximation, by way of Chebyshev's theorem, and this led us to an equioscillation property. Now we consider the best $\mathcal{L}_2$ polynomial approximation of a given degree, which leads us to an orthogonality property.

The theorems in this section are valid not only for the inner product (4.2), but for any inner product $\langle\cdot\,,\,\cdot\rangle$ as defined by Definition 4.2.

Theorem 4.1 *The best $\mathcal{L}_2$ polynomial approximation $p_n^B(x)$ of degree n (or less) to a given ($\mathcal{L}_2$-integrable) function $f(x)$ is unique and is characterised by the (necessary and sufficient) property that*

$$\left\langle f - p_n^B\,,\,p_n\right\rangle = 0 \tag{4.15}$$

for any other polynomial p_n of degree n.

Proof: Write

$$e_n^B := f - p_n^B.$$

1. (Necessity) Suppose that, for some polynomial p_n,

$$\left\langle e_n^B\,,\,p_n\right\rangle \neq 0.$$

Then, for any real scalar multiplier λ,

$$\begin{aligned}
\left\| f - (p_n^B + \lambda p_n)\right\|^2 &= \left\| e_n^B - \lambda p_n\right\|^2 \\
&= \left\langle e_n^B - \lambda p_n\,,\,e_n^B - \lambda p_n\right\rangle \\
&= \left\langle e_n^B\,,\,e_n^B\right\rangle - 2\lambda\left\langle e_n^B\,,\,p_n\right\rangle + \lambda^2\left\langle p_n\,,\,p_n\right\rangle \\
&= \left\| e_n^B\right\|^2 - 2\lambda\left\langle e_n^B\,,\,p_n\right\rangle + \lambda^2\left\|p_n\right\|^2 \\
&< \left\| e_n^B\right\|^2 \quad \text{for some small } \lambda \text{ of the same sign as } \left\langle e_n^B\,,\,p_n\right\rangle.
\end{aligned}$$

Hence $p_n^B + \lambda p_n$ is a better approximation than p_n^B for this value of λ, contradicting the assertion that p_n^B is a best approximation.

2. **(Sufficiency)** Suppose that (4.15) holds and that q_n is any specified polynomial of degree n, not identical to p_n^B. Then

$$\begin{aligned}
\|f - q_n\|^2 - \left\|f - p_n^B\right\|^2 &= \left\|e_n^B + (p_n^B - q_n)\right\|^2 - \left\|e_n^B\right\|^2 \\
&= \left\langle e_n^B + (p_n^B - q_n)\,,\, e_n^B + (p_n^B - q_n)\right\rangle - \left\langle e_n^B\,,\, e_n^B\right\rangle \\
&= \left\langle p_n^B - q_n\,,\, p_n^B - q_n\right\rangle + 2\left\langle e_n^B\,,\, p_n^B - q_n\right\rangle \\
&= \left\|p_n^B - q_n\right\|^2 + 0, \text{ from (4.15)} \\
&> 0.
\end{aligned}$$

Therefore $\|f - q_n\|^2 > \|f - p_n^B\|^2$.
Since q_n is arbitrary, p_n^B must be a best $\mathcal{L}_2$ approximation. It must also be unique, since otherwise we could have taken q_n to be another best approximation and obtained the last inequality as a contradiction. ●●

Corollary 4.1A *If $\{\phi_n\}$ (ϕ_i being of exact degree i) is an orthogonal polynomial system on $[a, b]$, then:*

1. *the zero function is the best $\mathcal{L}_2$ polynomial approximation of degree $(n-1)$ to ϕ_n on $[a, b]$;*

2. *ϕ_n is the best $\mathcal{L}_2$ approximation to zero on $[a, b]$ among polynomials of degree n with the same leading coefficient.*

Proof:

1. Any polynomial p_{n-1} of degree $n-1$ can be written in the form

$$p_{n-1} = \sum_{i=0}^{n-1} c_i \phi_i.$$

Then

$$\begin{aligned}
\langle \phi_n - 0\,,\, p_{n-1}\rangle &= \left\langle \phi_n\,,\, \sum_{i=0}^{n-1} c_i\phi_i\right\rangle \\
&= \sum_{i=0}^{n-1} c_i \langle \phi_n\,,\, \phi_i\rangle \\
&= 0 \text{ by the orthogonality of } \{\phi_i\}.
\end{aligned}$$

The result follows from Theorem 4.1.

2. Let q_n be any other polynomial of degree n having the same leading coefficient as ϕ_n. Then $q_n - \phi_n$ is a polynomial of degree $n-1$. We can therefore write

$$q_n - \phi_n = \sum_{i=0}^{n-1} c_i \phi_i$$

and deduce from the orthogonality of $\{\phi_i\}$ that

$$\langle \phi_n \, , \, q_n - \phi_n \rangle = 0. \tag{4.16}$$

Now we have

$$\begin{aligned} \|q_n\|^2 - \|\phi_n\|^2 &= \langle q_n \, , \, q_n \rangle - \langle \phi_n \, , \, \phi_n \rangle \\ &= \langle q_n - \phi_n \, , \, q_n - \phi_n \rangle - 2 \langle \phi_n \, , \, q_n - \phi_n \rangle \\ &= \|q_n - \phi_n\|^2 , \text{ using (4.16)} \\ &> 0. \end{aligned}$$

Therefore ϕ_n is the best approximation to zero. ●●

The interesting observation that follows from Corollary 4.1A is that every polynomial in an orthogonal system has a minimal $\mathcal{L}_2$ property — analogous to the minimax property of the Chebyshev polynomials. Indeed, the four kinds of Chebyshev polynomials T_n, U_n, V_n, W_n, being orthogonal polynomials, each have a minimal property on $[-1, 1]$ with respect to their respective weight functions

$$\frac{1}{\sqrt{1-x^2}}, \ \sqrt{1-x^2}, \ \sqrt{\frac{1+x}{1-x}}, \ \sqrt{\frac{1-x}{1+x}}$$

over all polynomials with the same leading coefficients.

The main result above, namely Theorem 4.1, is essentially a generalisation of the statement that *the shortest distance from a point to a plane is in the direction of a vector perpendicular to all vectors in that plane.*

Theorem 4.1 is important in that it leads to a very direct algorithm for determining the best $\mathcal{L}_2$ polynomial approximation p_n^B to f:

Corollary 4.1B *The best $\mathcal{L}_2$ polynomial approximation p_n^B of degree n to f may be expressed in terms of the orthogonal polynomial family $\{\phi_i\}$ in the form*

$$p_n^B = \sum_{i=0}^{n} c_i \phi_i,$$

where

$$c_i = \frac{\langle f \, , \, \phi_i \rangle}{\langle \phi_i \, , \, \phi_i \rangle}.$$

Proof: For $k = 0, 1, \ldots, n$

$$\left\langle f - p_n^B\,,\,\phi_k \right\rangle = \left\langle f - \sum_{i=0}^{n} c_i\phi_i\,,\,\phi_k \right\rangle$$
$$= \langle f\,,\,\phi_k\rangle - \sum_{i=0}^{n} c_i\,\langle \phi_i\,,\,\phi_k\rangle$$
$$= \langle f\,,\,\phi_k\rangle - c_k\,\langle \phi_k\,,\,\phi_k\rangle$$
$$= 0,\ \text{by definition of } c_k. \tag{4.17}$$

Now, any polynomial p_n can be written as

$$p_n = \sum_{i=0}^{n} d_i\phi_i,$$

and hence

$$\left\langle f - p_n^B\,,\,p_n \right\rangle = \sum_{i=0}^{n} d_i \left\langle f - p_n^B\,,\,\phi_i \right\rangle$$
$$= 0 \text{ by (4.17).}$$

Thus p_n^B is the best approximation by Theorem 4.1. ●●

EXAMPLE 4.1: To illustrate Corollary 4.1B, suppose that we wish to determine the best $\mathcal{L}_2$ linear approximation p_1^B to $f(x) = 1 - x^2$ on $[-1, 1]$, with respect to the weight $w(x) = (1 - x^2)^{-\frac{1}{2}}$. In this case $\{T_i(x)\}$ is the appropriate orthogonal system and hence

$$p_1^B = c_0T_0(x) + c_1T_1(x)$$

where, by (4.17),

$$c_0 = \frac{\langle f\,,\,T_0\rangle}{\langle T_0\,,\,T_0\rangle} = \frac{\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}(1-x^2)\,\mathrm{d}x}{\pi},$$
$$c_1 = \frac{\langle f\,,\,T_0\rangle}{\langle T_0\,,\,T_0\rangle} = \frac{\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}(1-x^2)x\,\mathrm{d}x}{\frac{1}{2}\pi}.$$

Substituting $x = \cos\theta$,

$$c_0 = \frac{1}{\pi}\int_0^{\pi}\sin^2\theta\,\mathrm{d}\theta = \frac{1}{2\pi}\int_0^{\pi}(1-\cos 2\theta)\,\mathrm{d}\theta = \tfrac{1}{2},$$
$$c_1 = \frac{2}{\pi}\int_0^{\pi}\sin^2\theta\,\cos\theta\,\mathrm{d}\theta = \frac{2}{\pi}\left[\tfrac{1}{3}\sin^3\theta\right]_0^{\pi} = 0$$

and therefore

$$p_1^B = \tfrac{1}{2}T_0(x) + 0\,T_1(x) = \tfrac{1}{2},$$

so that the linear approximation reduces to a constant in this case.

4.3.1 Orthogonal polynomial expansions

On the assumption that it is possible to expand a given function $f(x)$ in a (suitably convergent) series based on a system $\{\phi_k\}$ of polynomials orthogonal over the interval $[a, b]$, ϕ_k being of exact degree k, we may write

$$f(x) = \sum_{k=0}^{\infty} c_i \phi_i(x), \quad x \in [a, b]. \tag{4.18}$$

It follows, by taking inner products with ϕ_k, that

$$\langle f\,,\,\phi_k\rangle = \sum_{k=0}^{\infty} c_i \langle \phi_i\,,\,\phi_k\rangle = c_k \langle \phi_k\,,\,\phi_k\rangle\,,$$

since $\langle \phi_i\,,\,\phi_k\rangle = 0$ for $i \neq k$. This is identical to the formula for c_k given in Corollary 4.1B. Thus (applying the same corollary) an orthogonal expansion has the property that its partial sum of degree n is the best $\mathcal{L}_2$ approximation of degree n to its infinite sum. Hence it is an ideal expansion to use in the $\mathcal{L}_2$ context. In particular, the four Chebyshev series expansions have this property on $[-1, 1]$ with respect to their respective weight functions $(1 + x)^{\pm\frac{1}{2}}(1 - x)^{\pm\frac{1}{2}}$.

We shall have much more to say on this topic in Chapter 5.

4.3.2 Convergence in $\mathcal{L}_2$ of orthogonal expansions

Convergence questions will be considered in detail in Chapter 5, where we shall restrict attention to Chebyshev polynomials and use Fourier series theory. However, we may easily make some deductions from general orthogonal polynomial properties.

In particular, if f is continuous, then we know (Theorem 3.2) that arbitrarily accurate polynomial approximations exist in $\mathcal{C}[a, b]$, and it follows from Lemma 3.1 that these are also arbitrarily accurate in $\mathcal{L}_2[a, b]$. However, we have shown in Section 4.3.1 that the nth degree polynomial, $P_n(x)$ say, obtained by truncating an orthogonal polynomial expansion is a best $\mathcal{L}_2$ approximation. Hence (*a fortiori*) P_n must also achieve an arbitrarily small $\mathcal{L}_2$ error $\|f - P_n\|_2$ for sufficiently large n. This gives the following result.

Theorem 4.2 *If f is in $\mathcal{C}[a, b]$, then its expansion in orthogonal polynomials converges in $\mathcal{L}_2$ (with respect to the appropriate weight function).*

In Chapter 5, we obtain much more powerful convergence results for Chebyshev series, ensuring $\mathcal{L}_2$ convergence of the series itself for f in $L_2[a, b]$ and L_∞ convergence of Cesàro sums of the series for f in $\mathcal{C}[a, b]$.

4.4 Recurrence relations

Using the inner product (4.2), namely

$$\langle f\,,\,g\rangle = \int_a^b w(x)f(x)g(x)\,\mathrm{d}x,$$

we note that

$$\begin{aligned} \langle f\,,\,g\rangle &= \langle g\,,\,f\rangle\,, && (4.19)\\ \langle xf\,,\,g\rangle &= \langle f\,,\,xg\rangle\,. && (4.20) \end{aligned}$$

The following formulae uniquely define an orthogonal polynomial system $\{\phi_i\}$, in which ϕ_i is a monic polynomial (i.e., a polynomial with a leading coefficient of unity) of exact degree i.

Theorem 4.3 *The unique system of monic polynomials $\{\phi_i\}$, with ϕ_i of exact degree i, which are orthogonal on $[a,b]$ with respect to $w(x)$ are defined by*

$$\begin{aligned} \phi_0(x) &= 1,\\ \phi_1(x) &= x - a_1,\\ \phi_n(x) &= (x-a_n)\phi_{n-1}(x) - b_n\phi_{n-2}(x), \end{aligned} \qquad (4.21)$$

where

$$a_n = \frac{\langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}, \qquad b_n = \frac{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle}.$$

Proof: This is readily shown by induction on n. It is easy to show that the polynomials ϕ_n generated by (4.21) are all monic. We assume that the polynomials $\phi_0, \phi_1, \ldots, \phi_{n-1}$ are orthogonal, and we then need to test that ϕ_n, as given by (4.21), is orthogonal to ϕ_k $(k = 0, 1, \ldots, n-1)$.

The polynomial $x\phi_k$ is a monic polynomial of degree $k+1$, expressible in the form

$$x\phi_k(x) = \phi_{k+1}(x) + \sum_{i=1}^{k} c_i\phi_i(x),$$

so that, using (4.20),

$$\begin{aligned} \langle x\phi_{n-1}\,,\,\phi_k\rangle &= \langle \phi_{n-1}\,,\,x\phi_k\rangle = 0 \ (k < n-2),\\ \langle x\phi_{n-1}\,,\,\phi_{n-2}\rangle &= \langle \phi_{n-1}\,,\,x\phi_{n-2}\rangle = \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle\,. \end{aligned}$$

For $k < n-2$, then, we have

$$\langle \phi_n\,,\,\phi_k\rangle = \langle x\phi_{n-1}\,,\,\phi_k\rangle - a_n\langle \phi_{n-1}\,,\,\phi_k\rangle - b_n\langle \phi_{n-2}\,,\,\phi_k\rangle = 0,$$

while

$$\begin{aligned}
\langle \phi_n\,,\,\phi_{n-2}\rangle &= \langle x\phi_{n-1}\,,\,\phi_{n-2}\rangle - a_n\,\langle \phi_{n-1}\,,\,\phi_{n-2}\rangle - b_n\,\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle \\
&= \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle - 0 - \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle = 0, \\
\langle \phi_n\,,\,\phi_{n-1}\rangle &= \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - a_n\,\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle - b_n\,\langle \phi_{n-2}\,,\,\phi_{n-1}\rangle \\
&= \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - 0 = 0.
\end{aligned}$$

Starting the induction is easy, and the result follows. ●●

We have already established a recurrence relation for each of the four kinds of Chebyshev polynomials. We can verify that (4.21) leads to the same recurrences.

Consider the case of the polynomials of the first kind. We convert $T_n(x)$ to a monic polynomial by writing $\phi_0 = T_0$, $\phi_n = 2^{1-n}T_n$ $(n > 0)$. Then we can find the inner products:

$$\begin{aligned}
\langle T_0\,,\,T_0\rangle &= \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \mathrm{d}\theta = \pi, \\
\langle xT_0\,,\,T_0\rangle &= \int_{-1}^{1} \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos\theta\,\mathrm{d}\theta = 0, \\
\langle T_n\,,\,T_n\rangle &= \int_{-1}^{1} \frac{T_n(x)^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos^2 n\theta\,\mathrm{d}\theta = \tfrac{1}{2}\pi, \\
\langle xT_n\,,\,T_n\rangle &= \int_{-1}^{1} \frac{xT_n(x)^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos\theta\,\cos^2 n\theta\,\mathrm{d}\theta = 0.
\end{aligned}$$

Therefore $a_1 = 0$, $a_n = 0$ $(n > 1)$, and

$$\begin{aligned}
b_2 &= \frac{\langle \phi_1\,,\,\phi_1\rangle}{\langle \phi_0\,,\,\phi_0\rangle} = \frac{\langle T_1\,,\,T_1\rangle}{\langle T_0\,,\,T_0\rangle} = \tfrac{1}{2}, \\
b_n &= \frac{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle} = \frac{\left\langle 2^{2-n}T_{n-1}\,,\,2^{2-n}T_{n-1}\right\rangle}{\left\langle 2^{3-n}T_{n-2}\,,\,2^{3-n}T_{n-2}\right\rangle} = \tfrac{1}{4} \quad (n > 2).
\end{aligned}$$

So

$$\begin{aligned}
\phi_0 &= 1, \\
\phi_1 &= x, \\
\phi_2 &= x\phi_1 - \tfrac{1}{2}\phi_0, \\
\phi_n &= x\phi_{n-1} - \tfrac{1}{4}\phi_{n-2}\ (n > 2).
\end{aligned}$$

Hence the recurrence (1.3) for T_n.

We may similarly derive the recurrences (1.6) for U_n and (1.12) for V_n and W_n, by using their respective weight functions to obtain the appropriate a_n and b_n (see Problem 5).

4.5 Rodrigues' formulae and differential equations

If $\{\phi_i\}$ is a set of polynomials orthogonal on $[-1, 1]$ with respect to $w(x)$, with ϕ_i of degree i, then

$$\int_{-1}^{1} w(x)\phi_n(x)q_{n-1}(x)\,\mathrm{d}x = 0 \tag{4.22}$$

for any polynomial q_{n-1} of degree $n-1$.

Now suppose that $r_n(x)$ is an nth integral of $w(x)\phi_n(x)$, so that

$$r_n^{(n)}(x) = w(x)\phi_n(x). \tag{4.23}$$

Then (4.22) gives, on integration by parts,

$$\begin{aligned} 0 &= \int_{-1}^{1} r_n^{(n)}(x)q_{n-1}(x)\,\mathrm{d}x \\ &= \left[r^{(n-1)}(x)q_{n-1}(x)\right]_{-1}^{1} - \int_{-1}^{1} r_n^{(n-1)}(x)q_{n-1}'(x)\,\mathrm{d}x \\ &= \left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q_{n-1}'(x)\right]_{-1}^{1} + \int_{-1}^{1} r_n^{(n-2)}(x)q_{n-1}''(x)\,\mathrm{d}x \\ &= \cdots \\ &= \left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q_{n-1}'(x) + \cdots + (-1)^{n-1}r_n(x)q_{n-1}^{(n-1)}(x)\right]_{-1}^{1} + \\ &\qquad + (-1)^n \int_{-1}^{1} r_n(x)q_{n-1}^{(n)}(x)\,\mathrm{d}x. \end{aligned}$$

Hence, since $q_{n-1}^{(n)}(x) \equiv 0$, it follows that

$$\left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q_{n-1}'(x) + \cdots + (-1)^{n-1}r_n(x)q_{n-1}^{(n-1)}(x)\right]_{-1}^{1} = 0 \tag{4.24}$$

for any polynomial q_{n-1} of degree $n-1$.

Now $\phi_n^{(n+1)}(x) \equiv 0$, since ϕ_n is of degree n; hence, because of (4.23), r_n is a solution of the $(2n+1)$st order homogeneous differential equation

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}\left(w(x)^{-1}\frac{\mathrm{d}^n}{\mathrm{d}x^n}r_n(x)\right) = 0. \tag{4.25}$$

An arbitrary polynomial of degree $n-1$ may be added to r_n, without affecting the truth of (4.23) and (4.25). Hence we may without loss of generality arrange that $r_n(-1) = r_n'(-1) = \cdots = r_n^{(n-1)}(-1) = 0$, when the fact that (4.24) is valid for all q_{n-1} implies that $r_n(+1) = r'_n(+1) = \cdots = r_n^{(n-1)}(+1) = 0$, so that r_n satisfies the $2n$ homogeneous boundary conditions

$$r_n(\pm 1) = r'_n(\pm 1) = \cdots = r_n^{(n-1)}(\pm 1) = 0. \tag{4.26}$$

One function satisfying (4.26), for any real $\alpha > -1$, is

$$r_n(x) = (1 - x^2)^{n+\alpha}. \tag{4.27}$$

If we then choose

$$w(x) = (1 - x^2)^{\alpha} \tag{4.28}$$

then (4.25) is satisfied, and $r_n^{(n)}$ is of the form (4.23) with $\phi_n(x)$ a polynomial of degree n.

Since ϕ_n, as defined by (4.22), is unique apart from a multiplicative constant, it follows from (4.25), (4.27) and (4.28) that (for $\alpha > -1$)

$$\phi_n(x) = P_n^{(\alpha)}(x) := c_n \frac{1}{(1 - x^2)^{\alpha}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n+\alpha}, \tag{4.29}$$

where c_n is a constant, defines a system of polynomials $\{P_n^{(\alpha)}(x)\}$ orthogonal with respect to $w(x) = (1 - x^2)^{\alpha}$ on $[-1, 1]$. These polynomials are known as the *ultraspherical* (or *Gegenbauer*) *polynomials*, and the formula (4.29) for them is known as *Rodrigues' formula*.

It immediately follows that the Chebyshev polynomials of the first and second kinds are ultraspherical polynomials and, by comparing their leading coefficients with those in (4.29), we may readily deduce (see Problem 12) that, taking $\alpha = -\frac{1}{2}$ and $\alpha = +\frac{1}{2}$,

$$T_n(x) = \frac{(-1)^n \, 2^n \, n!}{(2n)!}(1 - x^2)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n-\frac{1}{2}}, \tag{4.30}$$

$$U_n(x) = \frac{(-1)^n \, 2^n \, (n+1)!}{(2n+1)!}(1 - x^2)^{-\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n+\frac{1}{2}}. \tag{4.31}$$

(In the standard notation for Gegenbauer polynomials, as in Abramowitz and Stegun's *Handbook of Mathematical Functions* (1964) for example, $P_n^{(\alpha)}(x)$ is written as $C_n^{\alpha+\frac{1}{2}}(x)$, so that $T_n(x)$ is proportional to $C_n^{(0)}(x)$ and $U_n(x)$ to $C_n^{(1)}(x)$.)

The well-known Legendre polynomials $P_n(x)$, which are orthogonal with weight unity, are ultraspherical polynomials for $\alpha = 0$ and are given by

$$P_n(x) = \frac{(-1)^n 2^{-n}}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^n. \tag{4.32}$$

Note that the Chebyshev polynomials of third and fourth kinds are not ultraspherical polynomials, but only Jacobi polynomials. Their Rodrigues' formulae are

$$V_n(x) = \frac{(-1)^n\, 2^n\, n!}{(2n)!} \left(\frac{1-x}{1+x}\right)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left\{ (1-x^2)^n \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \right\}, \quad (4.33a)$$

$$W_n(x) = \frac{(-1)^n\, 2^n\, n!}{(2n)!} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left\{ (1-x^2)^n \left(\frac{1-x}{1+x}\right)^{\frac{1}{2}} \right\}. \quad (4.33b)$$

From the general formula (4.29) it can be verified by substitution (see Problem 13) that $P_n^{(\alpha)}(x)$ is a solution of the second-order differential equation

$$(1-x^2)y'' - 2(\alpha+1)xy' + n(n+2\alpha+1)y = 0. \quad (4.34)$$

Thus $T_n(x)$, $U_n(x)$, $P_n(x)$ are solutions of

$$(1-x^2)y'' - xy' + n^2 y = 0 \quad (\alpha = -\tfrac{1}{2}), \quad (4.35a)$$

$$(1-x^2)y'' - 3xy' + n(n+2)y = 0 \quad (\alpha = \tfrac{1}{2}), \quad (4.35b)$$

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0 \quad (\alpha = 0), \quad (4.35c)$$

respectively.

The differential equations satisfied by $V_n(x)$ and $W_n(x)$ are, respectively,

$$(1-x^2)y'' - (2x-1)y' + n(n+1)y = 0, \quad (4.36a)$$

$$(1-x^2)y'' - (2x+1)y' + n(n+1)y = 0. \quad (4.36b)$$

4.6 Discrete orthogonality of Chebyshev polynomials

It is always possible to convert a (continuous) orthogonality relationship, as defined in Definition 4.1, into a discrete orthogonality relationship simply by replacing the integral with a summation. In general, of course, the result is only approximately true. However, where trigonometric functions or Chebyshev polynomials are involved, there are many cases in which the discrete orthogonality can be shown to hold exactly. We give here a few of the discrete orthogonality relations that exist between the four kinds of Chebyshev polynomials. Further relations are given by Mason & Venturino (1996) (see also Problem 14).

4.6.1 First-kind polynomials

Consider the sum

$$s_n^{(1)}(\theta) = \sum_{k=1}^{n+1} \cos(k-\tfrac{1}{2})\theta = \cos\tfrac{1}{2}\theta + \cos\tfrac{3}{2}\theta + \cdots + \cos(n+\tfrac{1}{2})\theta. \quad (4.37)$$

By summing the arithmetic series

$$z^{\frac{1}{2}}(1 + z + z^2 + \cdots + z^n),$$

substituting $z = \mathrm{e}^{\mathrm{i}\theta}$ and taking the real part of the result, it is easily verified (see Problem 4) that

$$s_n^{(1)}(\theta) = \frac{\sin(n+1)\theta}{2\sin\frac{1}{2}\theta} \tag{4.38}$$

and hence that $s_n^{(1)}(\theta)$ vanishes when $\theta = \dfrac{r\pi}{n+1}$, for integers r in the range $0 < r < 2(n+1)$. Further, we can see directly from (4.37) that

$$s_n^{(1)}(0) = n+1, \quad s_n^{(1)}(2\pi) = -(n+1). \tag{4.39}$$

Now consider

$$a_{ij} = \sum_{k=1}^{n+1} T_i(x_k)T_j(x_k) \quad (0 \le i, j \le n) \tag{4.40}$$

where x_k are the zeros of $T_{n+1}(x)$, namely

$$x_k = \cos\theta_k, \quad \theta_k = \frac{(k-\frac{1}{2})\pi}{n+1}. \tag{4.41}$$

Then

$$\begin{aligned} a_{ij} &= \sum_{k=1}^{n+1} \cos i\theta_k \cos j\theta_k \\ &= \tfrac{1}{2}\sum_{k=1}^{n+1}[\cos(i+j)\theta_k + \cos(i-j)\theta_k] \\ &= \tfrac{1}{2}\left[s_n^{(1)}\left(\frac{(i+j)\pi}{n+1}\right) + s_n^{(1)}\left(\frac{(i-j)\pi}{n+1}\right)\right]. \end{aligned}$$

Hence

$$a_{ij} = 0 \quad (i \neq j;\ i, j \le n), \tag{4.42a}$$

while, using (4.39),

$$a_{ii} = \tfrac{1}{2}(n+1) \quad (i \neq 0;\ i \le n) \tag{4.42b}$$

and

$$a_{00} = n+1. \tag{4.42c}$$

It follows from (4.42a) that the polynomials $\{T_i(x),\ i = 0, 1, \ldots, n\}$ are orthogonal over the discrete point set $\{x_k\}$ consisting of the zeros of $T_{n+1}(x)$. Specifically, the orthogonality is defined for the discrete inner product

$$\langle u\,,\ v\rangle = \sum_{k=1}^{n+1} u(x_k)v(x_k) \tag{4.43}$$

in the form

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j;\ i,j \leq n),$$

with $\langle T_0\,,\,T_0\rangle = n+1$ and $\langle T_i\,,\,T_i\rangle = \frac{1}{2}(n+1)$ $(0 < i \leq n)$.

This is not the only discrete orthogonality property of $\{T_i\}$. Indeed, by considering instead of (4.37) the sum

$$s_n^{(2)}(\theta) = \sum_{k=0}^{n}{}'' \cos k\theta = \tfrac{1}{2}\sin n\theta\,\cot\tfrac{1}{2}\theta \quad (n > 0)$$

(see Problem 4), where the double dash in $\sum''$ denotes that both first and last terms in the sum are to be halved, we deduce that

$$s_n^{(2)}(r\pi/n) = 0$$

for $0 < r < 2n$, while

$$s_n^{(2)}(0) = s_n^{(2)}(2\pi) = n.$$

If we now consider the extrema y_k of $T_n(x)$, namely

$$y_k = \cos\phi_k, \quad \phi_k = \frac{k\pi}{n} \ (k = 0, 1, \ldots, n) \tag{4.44}$$

(note that these $\{y_k\}$ are also the zeros of $U_{n-1}(x)$ together with the end points ± 1), and define

$$b_{ij} = \sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k), \tag{4.45}$$

then we have

$$b_{ij} = 0 \quad (i \neq j;\ i,j \leq n) \tag{4.46a}$$

$$b_{ii} = \tfrac{1}{2}n \quad (0 < i < n) \tag{4.46b}$$

$$b_{00} = b_{nn} = n. \tag{4.46c}$$

In this case the inner product is

$$\langle u\,,\,v\rangle = \sum_{k=0}^{n}{}'' u(y_k)v(y_k) \tag{4.47}$$

and we have again

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j;\ i,j \leq n),$$

but this time with $\langle T_0\,,\,T_0\rangle = \langle T_n\,,\,T_n\rangle = n$ and $\langle T_i\,,\,T_i\rangle = \frac{1}{2}n$, $(0 < i < n)$.

4.6.2 Second-kind polynomials

In a similar way, we can establish a pair of discrete orthogonality relationships for the weighted second-kind polynomials $\{\sqrt{1-x^2}\,U_i(x)\}$ corresponding to the point sets $\{x_k\}$ and $\{y_k\}$ defined in (4.41) and (4.44)).

Define

$$a_{ij}^{(2)} = \sum_{k=1}^{n+1} (1 - x_k^2) U_i(x_k) U_j(x_k) \quad (0 \le i, j \le n) \tag{4.48}$$

where $\{x_k\}$ are zeros of $T_{n+1}(x)$. Then we note that

$$\begin{aligned} a_{ij}^{(2)} &= \sum_{k=1}^{n+1} \sin(i+1)\theta_k \, \sin(j+1)\theta_k \\ &= \tfrac{1}{2} \sum_{k=1}^{n+1} [\cos(i-j)\theta_k - \cos(i+j+2)\theta_k] \\ &= \tfrac{1}{2} \left[s_n^{(1)} \left(\frac{(i-j)\pi}{n+1} \right) - s_n^{(1)} \left(\frac{(i+j+2)\pi}{n+1} \right) \right]. \end{aligned}$$

Hence

$$a_{ij}^{(2)} = 0 \quad (i \ne j;\ 0 \le i, j \le n) \tag{4.49a}$$

and

$$a_{ii}^{(2)} = \tfrac{1}{2}(n+1) \quad (0 \le i < n), \tag{4.49b}$$

while

$$a_{nn}^{(2)} = n + 1. \tag{4.49c}$$

Thus $\{\sqrt{1-x^2}\,U_i(x),\ i = 0, 1, \ldots, n\}$ are orthogonal for the inner product (4.43).

Similarly, considering the zeros $\{y_k\}$ of $(1-x^2)U_{n-1}(x)$,

$$\begin{aligned} b_{ij}^{(2)} &= \sum_{k=0}^{n}{}'' (1 - y_k^2) U_i(y_k) U_j(y_k) \\ &= \sum_{k=0}^{n}{}'' \sin(i+1)\phi_k \, \sin(j+1)\phi_k. \end{aligned} \tag{4.50}$$

Then

$$b_{ij}^{(2)} = 0 \quad (i \ne j;\ i, j \le n-1) \tag{4.51a}$$

$$b_{ii}^{(2)} = \tfrac{1}{2} n \quad (0 \le i < n-1) \tag{4.51b}$$

$$b_{n-1,n-1}^{(2)} = 0 \tag{4.51c}$$

and $\{\sqrt{1-x^2}\,U_i(x),\ i = 0, 1, \ldots, n-1\}$ are orthogonal for the inner product (4.47).

4.6.3 Third- and fourth-kind polynomials

Surprisingly, perhaps, the same discrete abscissae and inner products (4.43) and (4.47) provide orthogonality for the weighted third- and fourth-kind polynomials

$$\{\sqrt{1+x}\,V_i(x)\},\ \{\sqrt{1-x}\,W_i(x)\}.$$

For we have

$$\begin{aligned}
a_{ij}^{(3)} &= \sum_{k=1}^{n+1}(1+x_k)V_i(x_k)V_j(x_k) \quad (0 \le i,j \le n) \\
&= 2\sum_{k=1}^{n+1} \cos(i+\tfrac{1}{2})\theta_k \ \cos(j+\tfrac{1}{2})\theta_k \\
&= \sum_{k=1}^{n+1}[\cos(i+j+1)\theta_k + \cos(i-j)\theta_k],
\end{aligned}$$

giving us

$$a_{ij}^{(3)} = 0 \quad (i \ne j;\ i,j \le n) \tag{4.52a}$$

$$a_{ii}^{(3)} = n+1 \quad (0 \le i \le n), \tag{4.52b}$$

while

$$\begin{aligned}
b_{ij}^{(3)} &= \sideset{}{''}\sum_{k=0}^{n}(1+y_k)V_i(y_k)V_j(y_k) \quad (0 \le i,j \le n) \\
&= 2\sideset{}{''}\sum_{k=0}^{n} \cos(i+\tfrac{1}{2})\phi_k \ \cos(j+\tfrac{1}{2})\phi_k \\
&= \sideset{}{''}\sum_{k=0}^{n}[\cos(i+j+1)\phi_k + \cos(i-j)\phi_k],
\end{aligned}$$

giving

$$b_{ij}^{(3)} = 0 \quad (i \ne j;\ i,j \le n) \tag{4.53a}$$

$$b_{ii}^{(3)} = n \quad (0 \le i \le n). \tag{4.53b}$$

The same formulae (4.52)–(4.53) hold for $a_{ij}^{(4)}$ and $b_{ij}^{(4)}$, where

$$a_{ij}^{(4)} = \sum_{k=1}^{n+1}(1-x_k)W_i(x_k)W_j(x_k) \text{ and } b_{ij}^{(4)} = \sideset{}{''}\sum_{k=0}^{n}(1-y_k)W_i(y_k)W_j(y_k). \tag{4.54}$$

4.7 Discrete Chebyshev transforms and the fast Fourier transform

Using the values of a function $f(x)$ at the extrema $\{y_k\}$ of $T_n(x)$, which are also the zeros of $(1-x^2)U_{n-1}(x)$, given as in (4.44) by

$$y_k = \cos\frac{k\pi}{n} \quad (k = 0, \ldots, n), \tag{4.55}$$

we can define a *discrete Chebyshev transform* $\hat{f}(x)$, defined at these same points only, by the formula

$$\hat{f}(y_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n}{}'' T_k(y_j) f(y_j) \quad (k = 0, \ldots, n). \tag{4.56}$$

These values $\hat{f}(y_k)$ are in fact proportional to the coefficients in the interpolant of $f(y_k)$ by a sum of Chebyshev polynomials — see Section 6.3.2.

Using the discrete orthogonality relation (4.45), namely

$$\sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k) = \begin{cases} 0 & (i \neq j;\ i, j \leq n), \\ \frac{1}{2}n & (0 < i = j < n), \\ n & (i = j = 0 \text{ or } n), \end{cases} \tag{4.57}$$

we can easily deduce that the inverse transform is given by

$$f(y_j) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n}{}'' T_k(y_j)\hat{f}(y_k) \quad (j = 0, \ldots, n). \tag{4.58}$$

In fact, since

$$T_k(y_j) = \cos\frac{jk\pi}{n} = T_j(y_k),$$

which is symmetric in j and k, it is clear that the discrete Chebyshev transform is self-inverse.

It is possible to define other forms of discrete Chebyshev transform, based on any of the other discrete orthogonality relations detailed in Section 4.6.

The discrete Chebyshev transform defined here is intimately connected with the discrete Fourier (cosine) transform. Defining

$$\phi_k = \frac{k\pi}{n}$$

(the zeros of $\sin n\theta$) and

$$g(\theta) = f(\cos\theta), \quad \hat{g}(\theta) = \hat{f}(\cos\theta),$$

the formula (4.56) converts to the form

$$\hat{g}(\phi_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n}{}'' \cos\frac{jk\pi}{n} g(\phi_j) \quad (k = 0, \ldots, n). \tag{4.59}$$

Since $\cos\theta$ and therefore $g(\theta)$ are even and 2π-periodic functions of θ, (4.59) has alternative equivalent expressions

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=-n}^{n}{}'' \cos\frac{jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = -n, \ldots, n) \tag{4.60}$$

or

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=-n}^{n}{}'' \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = -n, \ldots, n) \tag{4.61}$$

or

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{2n-1} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = 0, \ldots, 2n-1). \tag{4.62}$$

The last formulae (4.61) and (4.62) in fact define the general discrete Fourier transform, applicable to functions $g(\theta)$ that are periodic but not necessarily even, whose inverse is the complex conjugate transform

$$g\left(\frac{j\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{k=-n}^{n}{}'' \exp\frac{-\mathrm{i}jk\pi}{n}\, \hat{g}\left(\frac{k\pi}{n}\right) \quad (j = -n, \ldots, n) \tag{4.63}$$

or

$$g\left(\frac{j\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{k=0}^{2n-1} \exp\frac{-\mathrm{i}jk\pi}{n}\, \hat{g}\left(\frac{k\pi}{n}\right) \quad (j = 0, \ldots, 2n-1). \tag{4.64}$$

4.7.1 The fast Fourier transform

Evaluation of (4.56) or (4.58) for a particular value of k or j, respectively, requires a number $O(n)$ of arithmetic operations; the algorithm described in Section 2.4.1 is probably the most efficient. If we require their values to be calculated for all values of k or j, however, use of this scheme would call for $O(n^2)$ operations in all, whereas it is possible to achieve the same results in $O(n \log n)$ operations (at the slight cost of working in complex arithmetic rather than real arithmetic, even though the final result is known to be real) by converting the Chebyshev transform to the equivalent Fourier transform (4.62) or (4.64), and then computing its $2n$ values simultaneously by means of the so-called *fast Fourier transform* (FFT) algorithm (Cooley & Tukey 1965, Gentleman & Sande 1966). The required $n+1$ values of the Chebyshev transform may then be extracted. (The remaining $n-1$ computed results will be redundant, by reason of the symmetry of g and $\hat{g}$.)

While there are versions of this algorithm that apply when n is a product of any small prime factors (Kolba & Parks 1977, Burrus & Eschenbacher 1981, for instance), it is easiest to describe it for the original and most useful case where n is a power of 2; say $n = 2^m$. Then, separating the even and odd terms of the summation, (4.62) becomes

$$\begin{aligned}
\hat{g}\left(\frac{k\pi}{n}\right) &= \sqrt{\tfrac{1}{2n}} \sum_{\substack{j=0 \\ j \text{ even}}}^{2n-2} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) + \sqrt{\tfrac{1}{2n}} \sum_{\substack{j=1 \\ j \text{ odd}}}^{2n-1} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \\
&= \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{2j\pi}{n}\right) + \\
&\qquad + \sqrt{\tfrac{1}{2n}} \exp\frac{\mathrm{i}k\pi}{n} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{(2j+1)\pi}{n}\right). \qquad (4.65\text{a})
\end{aligned}$$

while

$$\begin{aligned}
\hat{g}\left(\frac{(k+n)\pi}{n}\right) &= \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{2j\pi}{n}\right) - \\
&\qquad - \sqrt{\tfrac{1}{2n}} \exp\frac{\mathrm{i}k\pi}{n} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{(2j+1)\pi}{n}\right). \qquad (4.65\text{b})
\end{aligned}$$

Now, if for $j = 0, \ldots, n-1$ we define

$$g_1\left(\frac{2j\pi}{n}\right) := g\left(\frac{2j\pi}{n}\right) \text{ and } g_2\left(\frac{2j\pi}{n}\right) := g\left(\frac{(2j+1)\pi}{n}\right),$$

we can further rewrite (4.65a) and (4.65b) as

$$\begin{aligned}
\hat{g}\left(\frac{k\pi}{n}\right) &= \tfrac{1}{\sqrt{2}}\, \hat{g}_1\left(\frac{2k\pi}{n}\right) + \tfrac{1}{\sqrt{2}} \exp\frac{\mathrm{i}k\pi}{n}\, \hat{g}_2\left(\frac{2k\pi}{n}\right), \qquad (4.66\text{a}) \\
\hat{g}\left(\frac{(k+n)\pi}{n}\right) &= \tfrac{1}{\sqrt{2}}\, \hat{g}_1\left(\frac{2k\pi}{n}\right) - \tfrac{1}{\sqrt{2}} \exp\frac{\mathrm{i}k\pi}{n}\, \hat{g}_2\left(\frac{2k\pi}{n}\right), \qquad (4.66\text{b}) \\
&(k = 0, \ldots, n-1)
\end{aligned}$$

where the discrete Fourier transforms from g_1 to $\hat{g}_1$ and from g_2 to $\hat{g}_2$ each take a set of n values into another set of n values, whereas that from g to $\hat{g}$ takes $2n$ values into $2n$.

Thus, once we have performed the two transforms of order n, it requires fewer than Kn further arithmetic operations (where K denotes a small fixed integer) to compute the transform of order $2n$. Similarly it requires fewer than $2 \times K\frac{n}{2} = Kn$ operations to derive these two transforms of order n from four transforms of order $n/2$, fewer than $4 \times K\frac{n}{4} = Kn$ operations to derive these four transforms of order $n/2$ from eight transforms of order $n/4$, and so on. If $n = 2^m$, therefore, (a transform of order 1 being just the identity and therefore trivial) the discrete Fourier transform of order $2n$ may be performed in $(m+1)$ stages, each requiring fewer than Kn operations, so that the total number of operations is less than $(m+1)Kn = O(n \log n)$, as claimed above.

We do not propose to discuss in detail how this computation is best organised, but refer the reader to the extensive published literature (Canuto et al. 1988, van Loan 1992, for instance). Reliable off-the-peg implementations of a fast Fourier transform algorithm can be found in any comprehensive numerical subroutine library.

4.8 Discrete data fitting by orthogonal polynomials: the Forsythe–Clenshaw method

In this section we consider a least-squares/orthogonal polynomial method, in which Chebyshev polynomials fulfil what is essentially a supporting role. However, this is one of the most versatile polynomial approximation algorithms available, and the use of Chebyshev polynomials makes the resulting approximations much easier to use and compute. Moreover, the algorithm, in its Chebyshev polynomial form, is an essential tool in the solution of multivariate data-fitting problems for data on families of lines or curves.

We saw in Section 4.6 that an inner product may be defined on a discrete data set just as well as on a continuum, and in (4.43) we defined such an inner product based on Chebyshev polynomial zeros. However, we are frequently given a set of arbitrarily spaced data abscissae

$$x = x_k \quad (k = 1, \ldots, m), \tag{4.67}$$

and asked to approximate in a least-squares sense the corresponding ordinates

$$y = y_k$$

by a polynomial of degree n, where the number $(n+1)$ of free parameters is no more than the number m of given data — typically much smaller.

Now

$$\langle u\,,\, v\rangle = \sum_{k=1}^{m} w_k u(x_k) v(x_k) \tag{4.68}$$

defines an inner product over the points (4.67), where $\{w_k\}$ is a specified set of positive weights to be applied to the data. From Corollary 4.1B, the best polynomial approximation of degree n in the least-squares sense on the point set (4.67) is therefore

$$p_n^B = \sum_{i=0}^{n} c_i \phi_i(x) \tag{4.69}$$

where $\{\phi_i\}$ are orthogonal polynomials defined by the recurrence (4.21) with the inner product (4.68) and where

$$\begin{aligned} c_i &= \frac{\langle y \, , \, \phi_i \rangle}{\langle \phi_i \, , \, \phi_i \rangle} \\ &= \frac{\displaystyle\sum_{k=1}^{m} w_k y_k \phi_i(x_k)}{\displaystyle\sum_{k=1}^{m} w_k [\phi_i(x_k)]^2}. \end{aligned} \tag{4.70}$$

This is precisely the algorithm proposed by Forsythe (1957) for approximating discrete data y_k at arbitrary points x_k $(k = 1, \ldots, n)$.

The Forsythe algorithm, as we have described it so far, does not explicitly involve Chebyshev polynomials (or, for that matter, any other well-known set of orthogonal polynomials). However, if the data are distributed uniformly and very densely over an interval of x, $[-1, 1]$ say, then we expect the resulting polynomials to be very similar to conventional orthogonal polynomials defined on the continuum $[-1, 1]$. For example, if all the w_k are equal to unity, then $\{\phi_k\}$ should closely resemble the Legendre polynomials (orthogonal with respect to $w(x) = 1$), and if

$$w_k = (1 - x_k^2)^{-\frac{1}{2}}$$

then $\{\phi_k\}$ should resemble the Chebyshev polynomials of the first kind. In spite of this resemblance, we cannot simply use the Legendre or Chebyshev polynomials in place of the polynomials ϕ_k in (4.69) and (4.70), since on these points they are only approximately orthogonal, not exactly, and so we have to consider some other approach.

The goal is a formula for p_n^B based on Chebyshev polynomials of the first kind, say, in the form

$$p_n^B(x) = {\sum_{i=0}^{n}}' d_i^{(n)} T_i(x), \tag{4.71}$$

the new set of coefficients $d_i^{(n)}$ being chosen so that (4.71) is identical to (4.69). This form has the advantage over (4.69) that the basis $\{T_i(x)\}$ is independent of the abscissae (4.67) and therefore more convenient for repeated computation of $p_n^B(x)$. (This is a very useful step in the development of multivariate polynomial approximations on lines or curves of data.) An efficient algorithm for deriving the coefficients $d_i^{(n)}$ from the coefficients c_i is due to Clenshaw (1959/1960); it makes use of the recurrence relations for Chebyshev polynomials as well as those for the discretely orthogonal polynomials $\{\phi_i\}$. We shall not give more details here.

4.8.1 Bivariate discrete data fitting on or near a family of lines or curves

Formula (4.71), which gives an approximation p_n^B to data unequally spaced along one line, may readily be extended to much more general situations in two or more dimensions (Clenshaw & Hayes 1965). In two dimensions, for example, suppose that data are given at unequally-spaced and different locations on each of a family of lines parallel to the x-axis, say at the points

$$(x_{k\ell}, y_\ell), \quad k = 1, \ldots, m_{1\ell};\ \ell = 1, \ldots, m_2.$$

We suppose too that all of these points lie within the square $[-1, 1] \times [-1, 1]$. Then the data on each line $y = y_\ell$ may be approximated, using Clenshaw's algorithm, in the form (4.71), giving us a set of approximations[1]

$$p_{n_1,\ell}^B(x) = \sum_{i=0}^{n_1}{}' d_{i\ell}^{(n_1)} T_i(x), \quad \ell = 1, \ldots, m_2. \tag{4.72}$$

The set of ith coefficients $d_{i\ell}^{(n_1)}$, $\ell = 1, \ldots, m_2$, may then be treated as data on a line parallel to the y-axis and may be approximated in a similar manner, for each i from 0 to n_1, giving approximations

$$d_i(y) = \sum_{j=0}^{n_2}{}' d_{ij}^{(n_1,n_2)} T_j(y). \tag{4.73}$$

We thus arrive at the overall approximation

$$\sum_{i=0}^{n_1}{}' \sum_{j=0}^{n_2}{}' d_{ij}^{(n_1,n_2)} T_i(x) T_j(y). \tag{4.74}$$

If m_1, m_2, n_1, n_2 all $= O(n)$, then the algorithm involves $O(n^4)$ operations—compared with $O(n^3)$ for a meshed data polynomial (tensor product) algorithm. It is important (Clenshaw & Hayes 1965) to ensure that there are data located close to $x = \pm 1$ and $y = \pm 1$, if necessary by changing variables to transform boundary curves into straight lines.

[1] If the number of data points on any line is less than the number of degrees of freedom, $m_{1\ell} \leq n_1$, then instead of approximating we can interpolate with a polynomial of minimal degree by requiring that $d_{i\ell}^{(n_1)} = 0$ for $i \geq m_{1\ell}$.

This algorithm has been extended further by Bennell & Mason (1991) to data on a family of curves. The procedure is to run secant lines across the family of curved lines and interpolate the data on each curve to give values at the intersections, which are then approximated by using the method just described. The algorithm involves only about twice as many operations as that for data on lines, which would appear very satisfactory.

Other writers have developed similar algorithms. For example, Anderson et al. (1995) fit data lying 'near' a family of lines, using an iteration based on estimating values on the lines from the neighbouring data.

Algorithms such as these have considerable potential in higher dimensions. An application to modelling the surfaces of human teeth has been successfully carried out by Jovanovski (1999).

4.9 Orthogonality in the complex plane

Formulae for Chebyshev polynomials in terms of a complex variable z have been given in Section 1.4; we repeat them here for convenience.

Given any complex number z, we define the related complex number w to be such that

$$z = \tfrac{1}{2}(w + w^{-1}). \tag{4.75}$$

Unless z lies on the real interval $[-1, 1]$, this equation for w has two solutions

$$w = z \pm \sqrt{z^2 - 1}, \tag{4.76}$$

one of which has $|w| > 1$ and one has $|w| < 1$; we choose the one with $|w| > 1$. Then we have:

$$T_n(z) = \tfrac{1}{2}(w^n + w^{-n}); \tag{4.77}$$

$$U_n(z) = \frac{w^{n+1} - w^{-n-1}}{w - w^{-1}}; \tag{4.78}$$

$$V_n(z) = \frac{w^{n+\frac{1}{2}} + w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} + w^{-\frac{1}{2}}} = \frac{w^{n+1} + w^{-n}}{w + 1}; \tag{4.79}$$

$$W_n(z) = \frac{w^{n+\frac{1}{2}} - w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} - w^{-\frac{1}{2}}} = \frac{w^{n+1} - w^{-n}}{w - 1}. \tag{4.80}$$

For any $r > 1$, the elliptical contour E_r given by

$$E_r := \left\{ z : \left| z + \sqrt{z^2 - 1} \right| = r \right\} \quad (r > 1) \tag{4.81}$$

has foci at $z = \pm 1$, and is the image under (4.75) of the circle

$$C_r := \{ w : |w| = r \} \tag{4.82}$$

of radius r.

The Chebyshev polynomials have useful orthogonality properties on the ellipse E_r. In order to describe them, we need first to extend our definition (Definition 4.2) of the idea of an inner product to allow for functions that take complex values. Let $\overline{z}$ denote the complex conjugate of z.

Definition 4.3 *An* inner product $\langle \cdot\,,\,\cdot \rangle$ *is defined as a bilinear function of elements $f, g, h, \ldots$ of a vector space that satisfies the axioms:*

1. $\langle f\,,\,f \rangle$ *is real and* ≥ 0*, with equality if and only if* $f \equiv 0$*;*
2. $\langle f\,,\,g \rangle = \overline{\langle g\,,\,f \rangle}$ *(note the complex conjugate);*
3. $\langle f+g\,,\,h \rangle = \langle f\,,\,h \rangle + \langle g\,,\,h \rangle$*;*
4. $\langle \alpha f\,,\,g \rangle = \alpha \langle f\,,\,g \rangle$ *for any scalar* α*.*
   (Hence $\langle f\,,\,\alpha g \rangle = \overline{\alpha} \langle f\,,\,g \rangle$*.)*

This definition agrees with the earlier one if everything is real.

Now define an inner product

$$\langle f\,,\,g \rangle := \oint_{E_r} f(z)\overline{g(z)}\,|\mu(z)|\,|\mathrm{d}z|\,, \tag{4.83}$$

where $\mu(z)$ is a weighting function ($|\mu(z)|$ is real and positive) and $\oint \cdots |\mathrm{d}z|$ denotes integration with respect to arc length around the ellipse in an anticlockwise direction. This inner product corresponds to a norm $\|\cdot\|_2$ on E_r defined by

$$\|f\|_2^2 := \langle f\,,\,f \rangle = \oint_{E_r} |f(z)|^2\,|\mu(z)|\,|\mathrm{d}z|\,. \tag{4.84}$$

Then we can show, using this inner product, that

$$\langle T_m\,,\,T_n \rangle = \begin{cases} 0 & (m \neq n) \\ 2\pi & (m = n = 0) \\ \frac{1}{2}\pi(r^{2n} + r^{-2n}) & (m = n > 0) \end{cases}$$

$$\text{if } \mu(z) = \frac{1}{\sqrt{1-z^2}}, \tag{4.85a}$$

$$\langle U_m\,,\,U_n \rangle = \begin{cases} 0 & (m \neq n) \\ \frac{1}{2}\pi(r^{2n+2} + r^{-2n-2}) & (m = n \geq 0) \end{cases}$$

$$\text{if } \mu(z) = \sqrt{1-z^2}, \tag{4.85b}$$

$$\langle V_m\,,\,V_n \rangle = \begin{cases} 0 & (m \neq n) \\ \pi(r^{2n+1} + r^{-2n-1}) & (m = n \geq 0) \end{cases}$$

$$\text{if } \mu(z) = \sqrt{\frac{1+z}{1-z}}, \tag{4.85c}$$

$$\langle W_m\,,\,W_n\rangle = \begin{cases} 0 & (m \neq n) \\ \pi(r^{2n+1}+r^{-2n-1}) & (m = n \geq 0) \end{cases}$$

$$\text{if } \mu(z) = \sqrt{\frac{1-z}{1+z}}. \tag{4.85d}$$

Proof: Taking the first of these orthogonalities, for example, we have

$$\begin{aligned}
z &= \tfrac{1}{2}(w+w^{-1}) = \tfrac{1}{2}(r\mathrm{e}^{\mathrm{i}\theta}+r^{-1}\mathrm{e}^{-\mathrm{i}\theta}), \\
\mathrm{d}z &= \tfrac{1}{2}(1-w^{-2})\,\mathrm{d}w = \tfrac{1}{2}\mathrm{i}(r\mathrm{e}^{\mathrm{i}\theta}-r^{-1}\mathrm{e}^{-\mathrm{i}\theta})\,\mathrm{d}\theta, \\
T_m(z) &= \tfrac{1}{2}(w^m+w^{-m}) = \tfrac{1}{2}(r^m\mathrm{e}^{\mathrm{i}m\theta}+r^{-m}\mathrm{e}^{-\mathrm{i}m\theta}), \\
\mu(z) &= \frac{2}{\sqrt{2-w^2-w^{-2}}} = \frac{2}{\pm\mathrm{i}(w-w^{-1})} = \pm\frac{2\mathrm{i}}{r\mathrm{e}^{\mathrm{i}\theta}-r^{-1}\mathrm{e}^{-\mathrm{i}\theta}}, \\
|\mu(z)|\,|\,\mathrm{d}z| &= \mathrm{d}\theta, \\
T_m(z)\overline{T_n(z)} &= \tfrac{1}{4}(r^{m+n}\mathrm{e}^{\mathrm{i}(m-n)\theta}+r^{m-n}\mathrm{e}^{\mathrm{i}(m+n)\theta}+ \\
&\quad +r^{-m+n}\mathrm{e}^{-\mathrm{i}(m+n)\theta}+r^{-m-n}\mathrm{e}^{-\mathrm{i}(m-n)\theta}).
\end{aligned}$$

Now

$$\int_0^{2\pi}\mathrm{e}^{\mathrm{i}m\theta}\,\mathrm{d}\theta = 0 \quad (m\neq 0), \qquad = 2\pi \quad (m=0).$$

Hence we easily show that $\langle T_m\,,\,T_n\rangle = 0$ for $m \neq n$ and is as stated for $m = n$. ••

The other results (4.85) may be proved similarly (Problem 19), after noting that

$$\begin{aligned}
\left|1-z^2\right| &= \frac{1}{4\,|w|^2}\left|1-w^2\right|^2, \\
|1+z| &= \frac{1}{2\,|w|}\,|1+w|^2, \\
|1-z| &= \frac{1}{2\,|w|}\,|1-w|^2.
\end{aligned}$$

Note that $\|f\|_2$, as defined in (4.84), as well as being a norm on the space of functions square integrable round the contour E_r, can be used as a norm on the space of functions which are continuous on this contour and analytic throughout its interior. This may be shown using the maximum-modulus theorem (Problem 21).

Alternatively, we may define an inner product (and a corresponding norm) over the whole of the interior D_r of the ellipse E_r

$$\langle f\,,\,g\rangle = \iint_{D_r} f(z)\overline{g(z)}\,|\mu(z)|\;\mathrm{d}x\;\mathrm{d}y, \tag{4.86}$$

where $z = x + \mathrm{i}y$. Remarkably, the Chebyshev polynomials are orthogonal with respect to this inner product too, defining $\mu(z)$ for each kind of polynomial as in equations (4.85).

Proof: Take the first-kind polynomials, for example, when $\mu(z) = 1/\sqrt{1-z^2}$. If $z = \frac{1}{2}(w + w^{-1})$ and $w = s\mathrm{e}^{\mathrm{i}\theta}$, then z runs over the whole ellipse D_r when s runs from 1 to r and θ runs from 0 to 2π. We have

$$\begin{aligned} x &= \tfrac{1}{2}(s + s^{-1})\cos\theta, \\ y &= \tfrac{1}{2}(s - s^{-1})\sin\theta, \end{aligned}$$

so that

$$\mathrm{d}x\ \mathrm{d}y = \frac{\partial(x,y)}{\partial(s,\theta)}\,\mathrm{d}s\ \mathrm{d}\theta$$

with

$$\begin{aligned} \frac{\partial(x,y)}{\partial(s,\theta)} &= \det\begin{pmatrix} \frac{1}{2}(1-s^{-2})\cos\theta & -\frac{1}{2}(s+s^{-1})\sin\theta \\ +\frac{1}{2}(1+s^{-2})\sin\theta & \frac{1}{2}(s-s^{-1})\cos\theta \end{pmatrix} \\ &= \frac{1}{4}\frac{(s^2+2s\cos\theta+1)(s^2-2s\cos\theta+1)}{s^3} \end{aligned}$$

while

$$|\mu(z)| = \frac{4s^2}{(s^2+2s\cos\theta+1)(s^2-2s\cos\theta+1)}$$

Thus,

$$\begin{aligned} \langle f\,,\,g\rangle &= \iint_{D_r} f(z)\overline{g(z)}\,|\mu(z)|\ \mathrm{d}x\ \mathrm{d}y \\ &= \int_{s=1}^{r}\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}\,|\mu(z)|\,\frac{\partial(x,y)}{\partial(s,\theta)}\,\mathrm{d}s\ \mathrm{d}\theta \\ &= \int_{s=1}^{r}\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}s^{-1}\,\mathrm{d}s\ \mathrm{d}\theta \\ &= \int_{s=1}^{r} s^{-1}\,\mathrm{d}s\left\{\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}\,\mathrm{d}\theta\right\}. \end{aligned}$$

But the inner integral is simply the inner product around the ellipse E_s, which we have already shown to vanish if $f(z) = T_m(z)$ and $g(z) = T_n(z)$, with $m \neq n$. Therefore, the whole double integral vanishes, and $\langle T_m\,,\,T_n\rangle = 0$ for $m \neq n$. ●●

Orthogonality of the other three kinds of polynomial may be proved in the same way (Problem 20).

4.10 Problems for Chapter 4

1. Verify that the inner product (4.2) satisfies the axioms of Definition 4.2.

2. Using only the properties of an inner product listed in Definition 4.2, show that the norm defined by (4.4) satisfies the axioms of Definition 3.1.

3. If ϕ_n and ψ_n are two polynomials of degree n, each of which is orthogonal to every polynomial of degree less than n (over the same interval and with respect to the same weight function), show that $\phi_n(x)$ and $\psi_n(x)$ are proportional.

4. Derive the summations

$$\sum_{k=1}^{n+1} \cos(k - \tfrac{1}{2})\theta = \frac{\sin(n+1)\theta}{2\sin\frac{1}{2}\theta}$$

$$\sum_{k=1}^{n+1} \sin(k - \tfrac{1}{2})\theta = \frac{1 - \cos(n+1)\theta}{2\sin\frac{1}{2}\theta}$$

$$\sideset{}{''}\sum_{k=0}^{n} \cos k\theta = \tfrac{1}{2}\sin n\theta \cot \tfrac{1}{2}\theta$$

$$\sideset{}{''}\sum_{k=0}^{n} \sin k\theta = \tfrac{1}{2}(1 - \cos n\theta)\cot \tfrac{1}{2}\theta.$$

5. Using a similar analysis to that in Section 4.4, derive from (4.21) and the trigonometric formulae for $U_n(x)$, $V_n(x)$ and $W_n(x)$ the recurrence relations which are satisfied by these polynomials. Show that these relations coincide.

6. Using the recurrence (4.21), obtain formulae for the monic (Legendre) polynomials of degrees 0, 1, 2, 3, 4, which are orthogonal on $[-1, 1]$ with respect to $w(x) = 1$.

7. If $\{\phi_r\}$ is an orthogonal system on $[-1, 1]$, with ϕ_r a polynomial of exact degree r, prove that the zeros of ϕ_{r-1} separate those of ϕ_r; that is to say, between any two consecutive zeros of ϕ_r there lies a zero of ϕ_{r-1}. [Hint: Consider the signs of ϕ_r as $x \to +\infty$ and at the zeros of ϕ_{r-1}, using the recurrence (4.21).]

8. A simple alternative to the recurrence (4.21) for the generation of a system of monic orthogonal polynomials is the Gram–Schmidt orthogonalisation procedure:

   Given monic orthogonal polynomials $\phi_0, \phi_1, \ldots, \phi_{n-1}$, define ϕ_n in the form

$$\phi_n(x) = x^n + \sum_{k=0}^{n-1} c_k \phi_k(x)$$

and determine values of c_k such that ϕ_n is orthogonal to ϕ_0, ϕ_1, ..., ϕ_{n-1}.

Use this recurrence to generate monic polynomials of degrees 0, 1, 2 orthogonal on $[-1, 1]$ with respect to $(1+x)^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}$. What is the key disadvantage (in efficiency) of this algorithm, compared with the recurrence (4.21)?

9. By using the trigonometric formulae for $T_n(x)$ and $U_n(x)$, under the transformation $x = \cos\theta$, verify that these Chebyshev polynomials satisfy the respective differential equations (4.35a), (4.35b).

   Show similarly that $V_n(x)$ and $W_n(x)$ satisfy the differential equations (4.36a), (4.36b).

10. The second order differential equation (4.35a)

$$(1-x^2)y'' - xy' + n^2 y = 0$$

has $T_n(x)$ as one solution. Show that a second solution is $\sqrt{1-x^2}U_{n-1}(x)$. Find a second solution to (4.35b)

$$(1-x^2)y'' - 3xy' + n(n+2)y = 0,$$

one solution of which is $U_n(x)$.

11. By substituting $T_n(x) = t_0 + t_1 x + \cdots + t_n x^n$ into the differential equation that it satisfies, namely

$$(1-x^2)y'' - xy' + n^2 y = 0,$$

and equating coefficients of powers of x, show that $t_{n-1} = 0$ and

$$t_k(n^2 - k^2) + t_{k+2}(k+2)(k+1) = 0, \quad k = 0, \ldots, n-2.$$

Deduce that

$$t_{n-2m} = (-1)^m \frac{n}{n-m} \binom{n-m}{m} 2^{n-2m-1}$$

where

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

12. Writing

$$\frac{\mathrm{d}^r}{\mathrm{d}x^r}(1-x^2)^{n+\alpha} = (1-x^2)^{n-r+\alpha}(A_r x^r + \text{ lower degree terms}),$$

show that $A_{r+1} = -(2n - r + 2\alpha)A_r$ and deduce that

$$A_n = (-1)^n (2n+2\alpha)(2n+2\alpha-1)\cdots(n+2\alpha).$$

Hence, verify the formulae (4.30), (4.31) for $T_n(x)$, $U_n(x)$, determining the respective values of c_n in (4.29) by equating coefficients of x^n.

13. Verify that $P_n^{(\alpha)}(x)$, given by (4.29), is a solution of the second order equation

$$(1-x^2)y'' - 2(\alpha+1)xy' + n(n+2\alpha+1)y = 0.$$

[Hint: Write

$$\psi_n(x) := c_n^{-1}(1-x^2)^\alpha P_n^{(\alpha)}(x) = \mathrm{D}^n(1-x^2)^{n+\alpha},$$

where D stands for $\mathrm{d}/\mathrm{d}x$. Then derive two expressions for $\psi'_{n+1}(x)$:

$$\begin{aligned}
\psi'_{n+1}(x) &= \mathrm{D}^{n+2}(1-x^2)^{n+\alpha+1} \\
&= \mathrm{D}^{n+2}[(1-x^2)(1-x^2)^{n+\alpha}], \\
\psi'_{n+1}(x) &= \mathrm{D}^{n+1}\mathrm{D}(1-x^2)^{n+\alpha+1} \\
&= -2(n+\alpha+1)\mathrm{D}^{n+1}[x(1-x^2)^{n+\alpha}],
\end{aligned}$$

differentiate the two products by applying Leibniz's theorem, and equate the results. This should give a second-order differential equation for $\psi_n(x)$, from which the result follows.]

14. Determine which of the following four systems of $n+1$ weighted polynomials,

$$\{T_i(x)\}, \quad \{\sqrt{1-x^2}\,U_i(x)\}, \quad \{\sqrt{1+x}\,V_i(x)\}, \quad \{\sqrt{1-x}\,W_i(x)\}$$

$(0 \le i \le n)$ is discretely orthogonal with respect to which of the four following summations

$$\sum_{\text{zeros of } T_{n+1}(x)}, \qquad {\sum}''_{\text{zeros of } (1-x^2)U_{n-1}(x)},$$

$${\sum}^*_{\text{zeros of } (1+x)V_n(x)}, \qquad {\sum}'_{\text{zeros of } (1-x)W_n(x)}.$$

(Pay particular attention to the cases $i=0$ and $i=n$.) Find the values of $\langle T_i, T_i\rangle$, and similar inner products, in each case, noting that the result may not be the same for all values of i.

15. Using the discrete inner product $\langle u\,,\,v\rangle = \sum_k u(x_k)v(x_k)$, where $\{x_k\}$ are the zeros of $T_3(x)$, determine monic orthogonal polynomials of degrees 0, 1, 2 using the recurrence (4.21), and verify that they are identical to $\{2^{1-i}T_i(x)\}$.

16. If

$$f_n(x) = \sum_{i=1}^{n} c_i T_{i-1}(x), \qquad (*)$$

where

$$c_i = \frac{1}{n} \sum_{k=1}^{n}{}' f(x_k) T_{i-1}(x_k),$$

and x_k are the zeros of $T_n(x)$, show that $f_n(x_k) = f(x_k)$.

What does the formula (*) provide?

What can we say about the convergence in norm of f_n to f as $n \to \infty$?

17. Using the values of a function $f(x)$ at the zeros $\{x_k\}$ of $T_n(x)$, namely

$$x_k = \cos \frac{(2k+1)\pi}{2n} \quad (k = 0, \ldots, n-1),$$

define another form of discrete Chebyshev transform by

$$\hat{f}(x_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n-1} T_k(x_j) f(x_j) \quad (k = 0, \ldots, n-1).$$

Use discrete orthogonality to deduce that

$$f(x_j) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n-1} T_k(x_j) \hat{f}(x_k) \quad (j = 0, \ldots, n-1).$$

[See Canuto et al. (1988, p.503) for a fast computation procedure based on sets of alternate f values.]

18. Using the values of a function $f(x)$ at the positive zeros $\{x_k\}$ of $T_{2n}(x)$, namely

$$x_k = \cos \frac{(2k+1)\pi}{4n} \quad (k = 0, \ldots, n-1),$$

define another (odd) form of discrete Chebyshev transform by

$$\hat{f}(x_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n-1} T_{2k+1}(x_j) f(x_j) \quad (k = 0, \ldots, n-1).$$

Deduce that

$$f(x_j) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n-1} T_{2k}(x_j) \hat{f}(x_k) \quad (j = 0, \ldots, n-1),$$

and that this transform is self-inverse. [See Canuto et al. (1988, p.504) for a fast computation procedure.]

19. Verify the orthogonality properties (4.85).

20. Show that $\{U_n\}$, $\{V_n\}$ and $\{W_n\}$ are orthogonal over the ellipse E_r with respect to the inner product (4.86) and the appropriate weights.

    Evaluate $\langle T_n \, , \, T_n \rangle$, $\langle U_n \, , \, U_n \rangle$, $\langle V_n \, , \, V_n \rangle$ and $\langle W_n \, , \, W_n \rangle$ for this inner product.

21. Prove that if $\mathcal{A}_r$ denotes the linear space of functions that are analytic throughout the domain $\{z : \left| z + \sqrt{z^2 - 1} \right| \leq r\}$ $(r > 1)$, then $\|\cdot\|_2$, as defined by (4.84), has all of the properties of a norm required by Definition 3.1.

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

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

where

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

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

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

Also

$$c_0 = 4/\pi.$$

Hence,

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

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,

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

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

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

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

$\delta(x) = 0$ for $x \neq 0$,

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

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

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 $|u| \leq \frac{1}{2}$ either at $u = \frac{1}{2}x$ (local maximum) or at one or other of $u = \pm\frac{1}{2}$. It follows that

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

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 $|u| \leq \frac{1}{2}$. 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 $|x|\leq 1$, 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

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

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

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

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

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

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

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

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

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

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

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

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

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

(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*:

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

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 $\int w(x)\,dx$ 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

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

that

$$\|S_n^{FC}\|_\infty \le \|S_n^F\|_\infty \le \lambda_n. \tag{5.74}$$

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

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

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 $n \le 500$. 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 '$\leq$' 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

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

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

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

where

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

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

$$f(x) - (S_n^T f)(x) =$$

$$= \sum_{k=n+1}^{\infty} \frac{1}{4i\pi} \oint_{C_1} g(\eta)(\eta^k + \eta^{-k})(\xi^k + \xi^{-k}) \frac{d\eta}{\eta}$$

$$= \sum_{k=n+1}^{\infty} \frac{1}{4i\pi} \left[\oint_{C_r} g(\eta)\eta^{-k}(\xi^k + \xi^{-k}) \frac{d\eta}{\eta} + \oint_{C_{1/r}} g(\eta)\eta^{k}(\xi^k + \xi^{-k}) \frac{d\eta}{\eta} \right]$$

— since all parts of the integrand are analytic between C_r and $C_{1/r}$

$$= \sum_{k=n+1}^{\infty} \frac{1}{2i\pi} \oint_{C_r} g(\eta)\eta^{-k}(\xi^k + \xi^{-k}) \frac{d\eta}{\eta}$$

— replacing η by η^{-1} in the second integral, and using $g(\eta) = g(\eta^{-1})$

$$= \frac{1}{2i\pi} \oint_{C_r} g(\eta) \left(\frac{\xi^{n+1}\eta^{-n-1}}{1 - \xi\eta^{-1}} + \frac{\xi^{-n-1}\eta^{-n-1}}{1 - \xi^{-1}\eta^{-1}} \right)) \frac{d\eta}{\eta}, \tag{5.105}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

Chapter 6

Chebyshev Interpolation

6.1 Polynomial interpolation

One of the simplest ways of obtaining a polynomial approximation of degree n to a given continuous function $f(x)$ on $[-1, 1]$ is to interpolate between the values of $f(x)$ at $n+1$ suitably selected distinct points in the interval. For example, to interpolate at

$$x_1, x_2, \dots, x_{n+1}$$

by the polynomial

$$p_n(x) = c_0 + c_1 x + \cdots + c_n x^n,$$

we require that

$$c_0 + c_1 x_k + \cdots + c_n x_k^n = f(x_k) \quad (k = 1, \dots, n+1). \tag{6.1}$$

The equations (6.1) are a set of $n+1$ linear equations for the $n+1$ coefficients $c_0, \dots, c_n$ that define $p_n(x)$.

Whatever the values of $f(x_k)$, the interpolating polynomial $p_n(x)$ exists and is unique, since the determinant of the linear system (6.1) is non-zero. Specifically

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n+1} & x_{n+1}^2 & \cdots & x_{n+1}^n \end{pmatrix} = \prod_{i>j} (x_i - x_j) \neq 0.$$

It is generally not only rather time-consuming, but also numerically unstable, to determine $p_n(x)$ by solving (6.1) as it stands, and indeed many more efficient and reliable formulae for interpolation have been devised.

Some interpolation formulae are tailored to equally spaced points x_1, x_2, …, x_{n+1}, such as those based on finite differences and bearing the names of Newton and Stirling (Atkinson 1989, for example). Surprisingly however, if we have a free choice of interpolation points, it is not necessarily a good idea to choose them equally spaced. An obvious equally-spaced set for the interval $[-1, 1]$ is given for each value of n by

$$x_k = -1 + \frac{2k+1}{n+1} \quad (k = 0, \dots, n); \tag{6.2}$$

these points are spaced a distance $2/(n+1)$ apart, with half spacings of $1/(n+1)$ between the first and last points and the end points of the interval.

(This set would provide equally spaced interpolation on $(-\infty, \infty)$ if $f(x)$ were periodic with period 2.) However, the following example demonstrates that the points (6.2) are not appropriate for all continuous functions $f(x)$ when n becomes large.

Theorem 6.1 (Runge phenomenon) *If x_k are chosen to be the points (6.2) for each $n \geq 0$, then the interpolating polynomial $p_n(x)$ does* not *converge uniformly on $[-1, 1]$ as $n \to \infty$ for the function $f(x) = 1/(1 + 25x^2)$.*

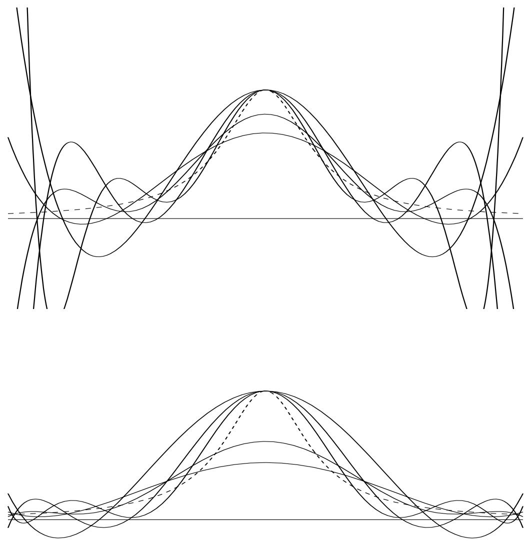


Figure 6.1: Interpolation to $f(x) = 1/(1 + 25x^2)$ by polynomials of degrees 4 to 8 at evenly-spaced points (above) and at Chebyshev polynomial zeros (below)

Proof: We refer the reader to (Mayers 1966) for a full discussion. The function $f(z)$ has complex poles at $z = \pm\frac{1}{5}\mathrm{i}$, which are close to the relevant part of the real axis, and it emerges that such nearby poles are sufficient to prevent uniform convergence. In fact the error $f(x) - p_n(x)$ oscillates wildly close to $x = \pm 1$, for large n. This is illustrated in the upper half of Figure 6.1.

See also (Trefethen & Weideman 1991), where it is noted that Turetskii (1940) showed that the Lebesgue constant for interpolation at evenly-spaced points is asymptotically $2^{n+1}/(\mathrm{e}\, n\, \log n)$. ●●

However, formulae are also available for unequally spaced interpolation, notably Neville's divided-difference algorithm or Aitken's algorithm (Atkinson 1989) and the general formula of Lagrange quoted in Lemma 6.3 below.

A better choice of interpolation points to ensure uniform convergence, though still not necessarily for every continuous function, is the set of zeros of the Chebyshev polynomial $T_{n+1}(x)$, namely (as given in Section 2.2)

$$x = x_k = \cos \frac{(k - \frac{1}{2})\pi}{n+1} \quad (k = 1, \ldots, n+1). \tag{6.3}$$

This choice of points does in fact ensure convergence for the function of Theorem 6.1, and indeed for any continuous $f(x)$ that satisfies a Dini–Lipschitz condition. Thus only a very slight restriction of $f(x)$ is required. This is illustrated in the lower half of Fig. 6.1. See Cheney (1966) or Mason (1982) for a proof of this. We note also from Theorem 6.5 that convergence in a weighted L_2 norm occurs for any continuous $f(x)$.

By expressing the polynomial in terms of Chebyshev polynomials, this choice of interpolation points (6.3) can be made far more efficient and stable from a computational point of view than the equally-spaced set (6.2). So we gain not only from improved convergence but also from efficiency and reliability. We show this in Section 6.3.

Finally, we shall find that we obtain a near-minimax approximation by interpolation at Chebyshev zeros, just as we could by truncating the Chebyshev series expansion — but in this case by a much simpler procedure.

6.2 Orthogonal interpolation

If $\{\phi_i\}$ is any orthogonal polynomial system with ϕ_i of exact degree i then, rather than by going to the trouble of computing an orthogonal polynomial expansion (which requires us to evaluate the inner-product integrals $\langle f\,,\,\phi_i\rangle$), an easier way to form a polynomial approximation $P_n(x)$ of degree n to a given function $f(x)$ is by interpolating $f(x)$ at the $(n+1)$ zeros of $\phi_{n+1}(x)$. In fact, the resulting approximation is often just as good.

The following theorem establishes for general orthogonal polynomials what we already know in the case of Chebyshev polynomials, namely that $\phi_{n+1}(x)$ does indeed have the required $(n+1)$ distinct zeros in the chosen interval.

Theorem 6.2 *If the system $\{\phi_i\}$, with ϕ_i a polynomial of exact degree i, is orthogonal on $[a,b]$ with respect to a non-negative weight $w(x)$, then ϕ_n has exactly n distinct real zeros in $[a,b]$, for every $n \geq 0$.*

Proof: (Snyder 1966, p.7, for example) Suppose that ϕ_n has fewer than n real zeros, or that some of its zeros coincide. Then there are m points $t_1, t_2, \ldots, t_m$ in $[a,b]$, with $0 \leq m < n$, where $\phi_n(x)$ changes sign. Let

$$\Pi_m(x) := \prod_{i=1}^{m}(x - t_i), \quad m \geq 1; \quad \Pi_0(x) := 1.$$

Then Π_m is a polynomial of degree $m < n$, and so must be orthogonal to ϕ_n. But

$$\langle \Pi_m\,,\ \phi_n \rangle = \int_a^b w(x)\Pi_m(x)\phi_n(x)\,\mathrm{d}x \neq 0,$$

since this integrand $w(x)\Pi_m(x)\phi_n(x)$ must have the same sign throughout the interval (except at the m points where it vanishes). We thus arrive at a contradiction. ●●

Since an interpolant samples the values of the function in a discrete set of points only, it is usual to require the function to be in $\mathcal{C}[a,b]$ (i.e., to be continuous), even if we wish to measure the goodness of the approximation in a weaker norm such as $\mathcal{L}_2$.

Some basic facts regarding polynomial interpolation are given by the following lemmas.

Lemma 6.3 *The polynomial of degree n interpolating the continuous function $f(x)$ at the $n+1$ distinct points $x_1, \ldots, x_{n+1}$ can be written as*

$$p_n(x) = \sum_{i=1}^{n+1} \ell_i(x) f(x_i) \tag{6.4}$$

where $\ell_i(x)$ are the usual Lagrange polynomials

$$\ell_i(x) = \prod_{\substack{k=1\\k\neq i}}^{n+1} \left(\frac{x - x_k}{x_i - x_k} \right). \tag{6.5}$$

Lemma 6.4 *If $x_1, \ldots, x_{n+1}$ are the zeros of the polynomial $\phi_{n+1}(x)$, then the Lagrange polynomials (6.5) may be written in the form*

$$\ell_i(x) = \frac{\phi_{n+1}(x)}{(x - x_i)\phi'_{n+1}(x_i)}, \tag{6.6}$$

where $\phi'(x)$ denotes the first derivative of $\phi(x)$.

In the special case of the first-kind Chebyshev polynomials, the preceding lemma gives the following specific result.

Corollary 6.4A *For polynomial interpolation at the zeros of the Chebyshev polynomial $T_{n+1}(x)$, the Lagrange polynomials are*

$$\ell_i(x) = \frac{T_{n+1}(x)}{(n+1)(x - x_i)U_n(x_i)},$$

or

$$\begin{aligned} \ell_i(\cos\theta) &= \frac{\cos(n+1)\theta\ \sin\theta_i}{(n+1)(\cos\theta - \cos\theta_i)\sin(n+1)\theta_i} \\ &= -\frac{\sin(n+1)(\theta - \theta_i)\ \sin\theta_i}{(n+1)(\cos\theta - \cos\theta_i)}. \end{aligned} \tag{6.7}$$

The following general result establishes $\mathcal{L}_2$ convergence in this framework of interpolation at orthogonal zeros.

Theorem 6.5 *(Erdös & Turán 1937) If $f(x)$ is in $\mathcal{C}[a,b]$, if $\{\phi_i(x),\ i = 0,1,\ldots\}$ is a system of polynomials (with ϕ_i of exact degree i) orthogonal with respect to $w(x)$ on $[a,b]$ and if $p_n(x)$ interpolates $f(x)$ in the zeros of $\phi_{n+1}(x)$, then*

$$\lim_{n\to\infty} (\|f - p_n(x)\|_2)^2 = \lim_{n\to\infty} \int_a^b w(x)(f(x) - p_n(x))^2 \, dx = 0.$$

Proof: The proof is elegant and subtle, and a version for Chebyshev polynomials is given by Rivlin (1974). We give a sketched version.

It is not difficult to show that $\{\ell_i\}$ are orthogonal. By ordering the factors appropriately, we can use (6.6) to write

$$\ell_i(x)\ell_j(x) = \phi_{n+1}(x)\psi_{n-1}(x) \quad (i \neq j)$$

where ψ_{n-1} is a polynomial of degree $n-1$. This must be orthogonal to ϕ_{n+1} and hence

$$\langle \ell_i\, ,\, \ell_j \rangle = \langle \phi_{n+1}\, ,\, \psi_{n-1} \rangle = 0.$$

Therefore

$$\langle \ell_i\, ,\, \ell_j \rangle = 0 \quad (i \neq j). \tag{6.8}$$

Now

$$\|f - p_n\|_2 \leq \left\| f - p_n^B \right\|_2 + \left\| p_n^B - p_n \right\|_2$$

where p_n^B is the best $\mathcal{L}_2$ approximation. Therefore, in view of Theorem 4.2, it suffices to prove that

$$\lim_{n\to\infty} \left\| p_n^B - p_n \right\|_2 = 0.$$

Since

$$p_n^B(x) = \sum_{i=1}^{n+1} \ell_i(x) p_n^B(x_i)$$

it follows from (6.4) and (6.8) that

$$\left(\left\| p_n^B - p_n \right\|_2 \right)^2 = \sum_{i=1}^{n+1} \langle \ell_i\, ,\, \ell_i \rangle \left[f(x_i) - p_n^B(x_i) \right]^2 .$$

Provided that $\langle \ell_i\, ,\, \ell_i \rangle$ can be shown to be uniformly bounded for all i, the right-hand side of this equality tends to zero by Theorem 4.2. This certainly holds in the case of Chebyshev polynomials, where $\langle \ell_i\, ,\, \ell_i \rangle = \dfrac{\pi}{n+1}$. ●●

In the cases $w(x) = (1+x)^{\pm\frac{1}{2}}(1-x)^{\pm\frac{1}{2}}$, Theorem 6.5 gives $\mathcal{L}_2$ convergence properties of polynomial interpolation at Chebyshev polynomial zeros. For example, if x_i are taken to be the zeros of $T_{n+1}(x)$ then

$$\lim_{n\to\infty} \int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} (f(x) - p_n(x))^2 \, dx = 0.$$

This result can be extended, and indeed Erdös & Feldheim (1936) have established $\mathcal{L}_p$ convergence for all $p > 1$:

$$\lim_{n\to\infty} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \left| f(x) - p_n(x) \right|^p \, \mathrm{d}x = 0.$$

In the case of Chebyshev zeros we are able to make more precise comparisons with best approximations (see Section 6.5).

If the function $f(x)$ has an analytic extension into the complex plane, then it may be possible to use the calculus of residues to obtain the following further results.

Lemma 6.6 *If the function $f(x)$ extends to a function $f(z)$ of the complex variable z, which is analytic within a simple closed contour C that encloses the point x and all the zeros $x_1, \ldots, x_{n+1}$ of the polynomial $\phi_{n+1}(x)$, then the polynomial of degree n interpolating $f(x)$ at these zeros can be written as*

$$p_n(x) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{\{\phi_{n+1}(z) - \phi_{n+1}(x)\} f(z)}{\phi_{n+1}(z)(z-x)} \, \mathrm{d}z \tag{6.9}$$

and its error is

$$f(x) - p_n(x) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{\phi_{n+1}(x) f(z)}{\phi_{n+1}(z)(z-x)} \, \mathrm{d}z. \tag{6.10}$$

In particular, if $f(x)$ extends to a function analytic within the elliptical contour E_r of Figure 1.5, then we can get a bound on the error of interpolation using the zeros of $T_{n+1}(x)$, implying uniform convergence in this case.

Corollary 6.6A *If the contour C in Lemma 6.6 is the ellipse E_r of (1.34), the locus of the points $\frac{1}{2}(r\mathrm{e}^{\mathrm{i}\theta} + r^{-1}\mathrm{e}^{-\mathrm{i}\theta})$ as θ varies (with $r > 1$, and if $|f(z)| \le M$ at every point z on E_r, then for every real x on $[-1, 1]$ we can show (see Problem 2) from (6.10), using (1.50) and the fact that $|T_{n+1}(x)| \le 1$, that*

$$|f(x) - p_n(x)| \le \frac{(r + r^{-1})M}{(r^{n+1} - r^{-n-1})(r + r^{-1} - 2)}, \quad x \text{ real}, \ -1 \le x \le 1. \tag{6.11}$$

6.3 Chebyshev interpolation formulae

We showed in Section 4.6 that the Chebyshev polynomials $\{T_i(x)\}$ of degrees up to n are orthogonal in a discrete sense on the set (6.3) of zeros $\{x_k\}$ of $T_{n+1}(x)$. Specifically

$$\sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) = \begin{cases} 0 & i \ne j \ (\le n) \\ n+1 & i = j = 0 \\ \frac{1}{2}(n+1) & 0 < i = j \le n. \end{cases} \tag{6.12}$$

This discrete orthogonality property leads us to a very efficient interpolation formula. Write the nth degree polynomial $p_n(x)$, interpolating $f(x)$ in the points (6.3), as a sum of Chebyshev polynomials in the form

$$p_n(x) = {\sum_{i=0}^{n}}' c_i T_i(x). \tag{6.13}$$

Theorem 6.7 *The coefficients c_i in (6.13) are given by the explicit formula*

$$c_i = \frac{2}{n+1} \sum_{k=1}^{n+1} f(x_k) T_i(x_k). \tag{6.14}$$

Proof: If we set $f(x)$ equal to $p_n(x)$ at the points $\{x_k\}$, then it follows that

$$f(x_k) = {\sum_{i=0}^{n}}' c_i T_i(x_k).$$

Hence, multiplying by $\dfrac{2}{n+1} T_j(x_k)$ and summing,

$$\begin{aligned} \frac{2}{n+1} \sum_{k=1}^{n+1} f(x_k) T_j(x_k) &= {\sum_{i=0}^{n}}' c_i \left\{ \frac{2}{n+1} \sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) \right\} \\ &= c_j, \end{aligned}$$

from (6.12), giving the formula (6.14). ●●

Corollary 6.7A *Formula (6.14) is equivalent to a 'discrete Fourier transform' of the transformed function*

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

Proof: We have

$$p_n(\cos\theta) = {\sum_{i=0}^{n}}' c_i \cos i\theta$$

with

$$c_i = \frac{2}{n+1} \sum_{k=1}^{n+1} g(\theta_k) \cos i\theta_k, \tag{6.15}$$

with

$$\theta_k = \frac{(k - \frac{1}{2})\pi}{n+1}. \tag{6.16}$$

Thus $\{c_i\}$ are discrete approximations to the true Fourier cosine series coefficients

$$c_i^S = \frac{1}{\pi} \int_{-\pi}^{\pi} g(\theta) \cos i\theta \, \mathrm{d}\theta, \tag{6.17}$$

obtained by applying a trapezoidal quadrature rule to the (periodic) function $g(\theta)$ with equal intervals $\pi/(n+1)$ between the points θ_k. Indeed, a trapezoidal rule approximation to (6.17), valid for any periodic function $g(\theta)$, is

$$c_i^S \simeq \frac{1}{\pi} \frac{\pi}{n+1} \sum_{k=-n}^{n+1} g\left(\frac{(k-\frac{1}{2})\pi}{n+1}\right) \cos \frac{i(k-\frac{1}{2})\pi}{n+1},$$

which gives exactly the formula (6.15) for c_i (when we note that the fact that both $g(\theta)$ and $\cos i\theta$ are even functions implies that the kth and $(1-k)$th terms in the summation are identical). ●●

Thus, Chebyshev interpolation has precisely the same effect as taking the partial sum of an approximate Chebyshev series expansion, obtained by approximating the integrals in the coefficients of the exact expansion by changing the independent variable from x to θ and applying the trapezoidal rule — thus effectively replacing the Fourier transforms c_i^S by discrete Fourier transforms c_i. It is well known among practical mathematicians and engineers that the discrete Fourier transform is a very good substitute for the continuous Fourier transform for periodic functions, and this therefore suggests that Chebyshev interpolation should be a very good substitute for a (truncated) Chebyshev series expansion.

In Sections 4.6.2 and 4.6.3 we obtained analogous discrete orthogonality properties to (6.12), based on the *same* abscissae x_k (zeros of T_{n+1}) but weighted, for the second, third and fourth kind polynomials. However, it is more natural to interpolate a Chebyshev polynomial approximation at the zeros of a polynomial of the same kind, namely the zeros of U_{n+1}, V_{n+1}, W_{n+1} in the case of second, third and fourth kind polynomials. We shall therefore show that analogous discrete orthogonality properties also follow for these new abscissae, and develop corresponding fast interpolation formulae.

6.3.1 Aliasing

We have already seen (Section 6.1) that polynomial interpolation at Chebyshev polynomial zeros is safer than polynomial interpolation at evenly distributed points. Even the former, however, is unreliable if too small a number of points (and so too low a degree of polynomial) is used, in relation to the properties of the function being interpolated.

One mathematical explanation of this remark, particularly as it applies to Chebyshev interpolation, is through the phenomenon of *aliasing*, described as follows.

Suppose that we have a function $f(x)$, having an expansion

$$f(x) = {\sum_{j=0}^{\infty}}' c_j T_j(x) \tag{6.18}$$

in Chebyshev polynomials, which is to be interpolated between its values at the zeros $\{x_k\}$ of $T_{n+1}(x)$ by the finite sum

$$f_n(x) = \sum_{j=0}^{n}{}' \hat{c}_j T_j(x). \tag{6.19}$$

The only information we can use, in order to perform such interpolation, is the set of values of each Chebyshev polynomial at the interpolation points. However, we have the following identity (where $x = \cos\theta$):

$$\begin{aligned} T_j(x) + T_{2n+2\pm j}(x) &= \cos j\theta + \cos(2n+2\pm j)\theta \\ &= \tfrac{1}{2}\cos(n+1)\theta\,\cos(n+1\pm j)\theta \\ &= \tfrac{1}{2}T_{n+1}(x)\,T_{n+1\pm j}(x), \end{aligned} \tag{6.20}$$

so that

$$T_j(x_k) + T_{2n+2\pm j}(x_k) = 0, \quad k = 1, \ldots, n+1. \tag{6.21}$$

Thus $T_{2n+2\pm j}$ is indistinguishable from $-T_j$ over the zeros of T_{n+1}. Figure 6.2 illustrates this in the case $n = 9$, $j = 4$ ($2n+2-j = 16$).

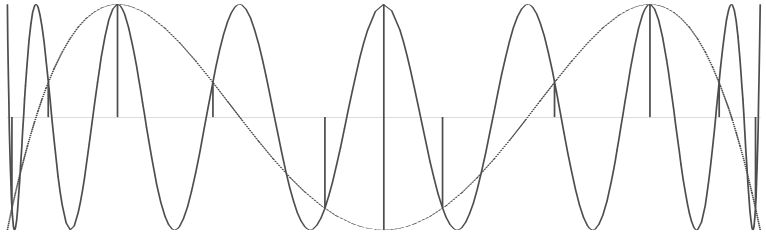


Figure 6.2: $T_{16}(x) = -T_4(x)$ at zeros of $T_{10}(x)$

In consequence, we can say that $f_n(x)$ as in (6.19) interpolates $f(x)$ as in (6.18) between the zeros of $T_{n+1}(x)$ when

$$\hat{c}_j = c_j - c_{2n+2-j} - c_{2n+2+j} + c_{4n+4-j} + c_{4n+4+j} - \cdots, \quad j = 0, \ldots, n. \tag{6.22}$$

(Note that the coefficients c_{n+1}, c_{3n+3}, ...do not figure in (6.22), as they correspond to terms in the expansion that vanish at every interpolation point.)

In effect, the process of interpolation removes certain terms of the expansion (6.18) entirely, while replacing the Chebyshev polynomial in each term after that in $T_n(x)$ by ($\pm 1\times$) a Chebyshev polynomial (its '*alias*') of lower degree. Since the coefficients $\{c_j\}$ tend rapidly to zero for well-behaved functions, the difference between c_j and $\hat{c}_j$ will therefore usually be small, but only if n is taken large enough for the function concerned.

Aliasing can cause problems to the unwary, for instance in working with nonlinear equations. Suppose, for instance that one has a differential equation involving $f(x)$ and $f(x)^3$, and one represents the (unknown) function $f(x)$ in the form $\sum_{j=0}^{\prime n} \hat{c}_j T_j(x)$ as in (6.19). Then one might be tempted to collocate the equation at the zeros of $T_{n+1}(x)$ — effectively carrying out a polynomial interpolation between these points. Instances such as the following, however, cast doubt on the wisdom of this.

In Figure 6.3 we have taken $n = 4$, and show the effect of interpolating the function $T_3(x)^3$ at the zeros of $T_5(x)$. (The expansion of $f_n(x)^3$ includes other products of three Chebyshev polynomials, of course, but this term will suffice.) Clearly the interpolation is poor, the reason being that

$$T_3(x)^3 = \tfrac{1}{4}\left(T_9(x) + 3T_3(x)\right),$$

which aliases to

$$\tfrac{1}{4}\left(-T_1(x) + 3T_3(x)\right).$$

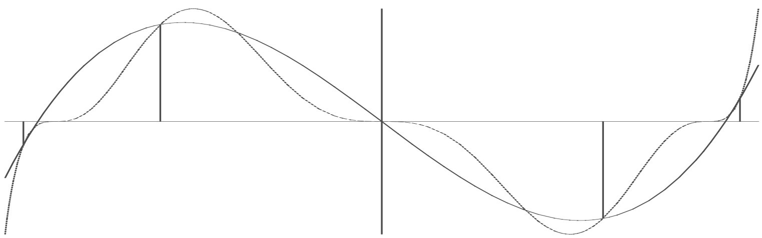


Figure 6.3: $T_3(x)^3$ interpolated at zeros of $T_5(x)$

In contrast, if we could have taken $n = 9$, we could have interpolated $T_3(x)^3$ exactly as shown in Figure 6.4. However, we should then have had

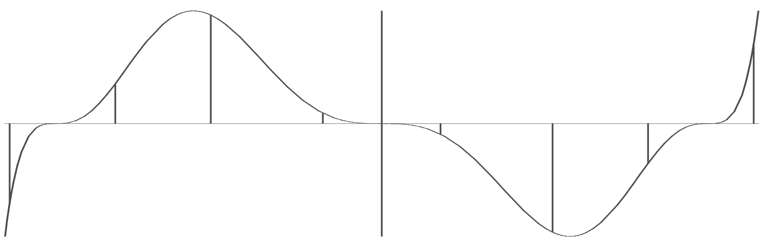


Figure 6.4: $T_3(x)^3$ interpolated (identically) at zeros of $T_{10}(x)$

to consider the effect of aliasing on further products of polynomials of higher order, such as those illustrated in Figures 6.5 and 6.6. There are ways of circumventing such difficulties, which we shall not discuss here.

Much use has been made of the concept of aliasing in estimating quadrature errors (see Section 8.4, where interpolation points and basis functions other than those discussed above are also considered).

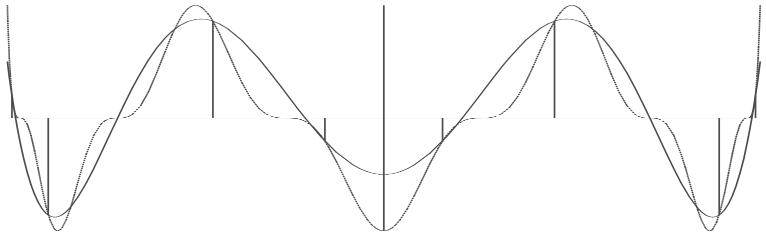


Figure 6.5: $T_6(x)^3$ interpolated at zeros of $T_{10}(x)$

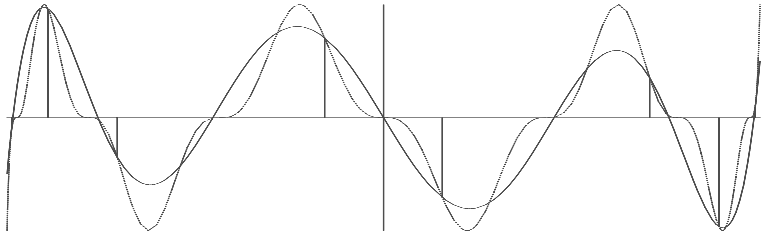


Figure 6.6: $T_7(x)^3$ interpolated at zeros of $T_{10}(x)$

6.3.2 Second-kind interpolation

Consider in this case interpolation by a weighted polynomial $\sqrt{1-x^2}\,p_n(x)$ on the zeros of $U_{n+1}(x)$, namely

$$y_k = \cos\frac{k\pi}{n+2} \quad (k = 1, \ldots, n+1).$$

Theorem 6.8 *The weighted interpolation polynomial to $f(x)$ is given by*

$$\sqrt{1-x^2}\,p_n(x) = \sqrt{1-x^2}\sum_{i=0}^{n} c_i U_i(x) \tag{6.23}$$

with coefficients given by

$$c_i = \frac{2}{n+1}\sum_{k=1}^{n+1}\sqrt{1-y_k^2} f(y_k) U_i(y_k). \tag{6.24}$$

Proof: From (4.50), with $n-1$ replaced by $n+1$,

$$\sum_{k=1}^{n+1}(1-y_k^2)U_i(y_k)U_j(y_k) = \begin{cases} 0, & i \neq j\ (\leq n); \\ \frac{1}{2}(n+1), & i = j \leq n. \end{cases} \tag{6.25}$$

If we set $\sqrt{1-y_k^2}p_n(y_k)$ equal to $f(y_k)$, we obtain

$$f(y_k) = \sqrt{1-y_k^2}\sum_{i=0}^{n} c_i U_i(y_k),$$

and hence, multiplying by $\frac{2}{n+1}\sqrt{1-y_k^2}\,U_j(y_k)$ and summing over k,

$$\begin{aligned}\frac{2}{n+1}\sum_{k=1}^{n+1}\sqrt{1-y_k^2}f(y_k)U_j(y_k) &= \sum_{i=0}^{n} c_i\left\{\frac{2}{n+1}\sum_{k=1}^{n+1}(1-y_k^2)U_i(y_k)U_j(y_k)\right\}\\ &= c_i\end{aligned}$$

by (6.25). ●●

Alternatively, we may want to interpolate at the zeros of $U_{n-1}(x)$ together with the points $x = \pm 1$, namely

$$y_k = \cos\frac{k\pi}{n} \quad (k = 0, \ldots, n).$$

In this case, however, we must express the interpolating polynomial as a sum of first-kind polynomials, when we can use the discrete orthogonality formula (4.45)

$$\sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k) = \begin{cases} 0, & i \neq j \ (\leq n); \\ \frac{1}{2}n, & 0 < i = j < n; \\ n, & i = j = 0;\ i = j = n. \end{cases} \tag{6.26}$$

(Note the double prime indicating that the first and last terms of the sum are to be halved.)

The interpolating polynomial is then

$$p_n(x) = \sum_{i=0}^{n}{}'' c_i T_i(x) \tag{6.27}$$

with coefficients given by

$$c_i = \frac{2}{n}\sum_{k=0}^{n}{}'' f(y_k)T_i(y_k). \tag{6.28}$$

Apart from a factor of $\sqrt{2/n}$, these coefficients make up the discrete Chebyshev transform of Section 4.7.

6.3.3 Third- and fourth-kind interpolation

Taking as interpolation points the zeros of $V_{n+1}(x)$, namely

$$x_k = \cos\frac{(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \quad (k = 1, \ldots, n+1),$$

we have the orthogonality formula, for $i, j \leq n$,

$$\sum_{k=1}^{n+1}(1+x_k)V_i(x_k)V_j(x_k) = \begin{cases} 0 & i \neq j \\ n+\frac{3}{2} & i = j \end{cases} \tag{6.29}$$

(See Problem 14 of Chapter 4.)

Theorem 6.9 *The weighted interpolation polynomial to $\sqrt{1+x}\, f(x)$ is given by*

$$\sqrt{1+x}\, p_n(x) = \sqrt{1+x} \sum_{i=0}^{n} c_i V_i(x) \tag{6.30}$$

where

$$c_i = \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} \sqrt{1+x_k} f(x_k) V_i(x_k). \tag{6.31}$$

Proof: If we set $\sqrt{1+x_k}\, p_n(x_k)$ equal to $\sqrt{1+x_k} f(x_k)$, we obtain

$$\sqrt{1+x_k}\, f(x_k) = \sqrt{1+x_k} \sum_{i=0}^{n} c_i V_i(x_k),$$

and hence, multiplying by $\dfrac{1}{n+\frac{3}{2}} \sqrt{1+x_k}\, V_j(x_k)$ and summing over k,

$$\begin{aligned} \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} (1+x_k) f(x_k) V_j(x_k) &= \sum_{i=0}^{n} c_i \left\{ \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} (1+x_k) V_i(x_k) V_j(x_k) \right\} \\ &= c_i \end{aligned}$$

by (6.29). ●●

The same goes for interpolation at the zeros of $W_{n+1}(x)$, namely

$$x_k = \cos \frac{(n-k+2)\pi}{n+\frac{3}{2}} \quad (k = 1, \ldots, n+1),$$

if we replace 'V' by 'W' and '$1+x$' by '$1-x$' throughout.

Alternatively, we may interpolate at the zeros of $V_n(x)$ together with one end point $x = -1$; i.e., at the points

$$x_k = \cos \frac{(k-\frac{1}{2})\pi}{n+\frac{1}{2}} \quad (k = 1, \ldots, n+1),$$

where we have the discrete orthogonality formulae (the notation $\sum^*$ indicating that the *last* term of the summation is to be halved)

$$\sum_{k=1}^{n+1}{}^{*}\, T_i(x_k) T_j(x_k) = \begin{cases} 0 & i \neq j \ (\leq n) \\ n+\frac{1}{2} & i = j = 0 \\ \frac{1}{2}(n+\frac{1}{2}) & 0 < i = j \leq n. \end{cases} \tag{6.32}$$

The interpolating polynomial is then

$$p_n(x) = \sum_{i=0}^{n}{}' \, c_i T_i(x) \tag{6.33}$$

with coefficients given by

$$c_i = \frac{2}{n+\frac{1}{2}} \sum_{k=1}^{n+1}{}^{*} f(x_k) T_i(x_k). \qquad (6.34)$$

6.3.4 Conditioning

In practice, one of the main reasons for the use of a Chebyshev polynomial basis is the good conditioning that frequently results. A number of comparisons have been made of the conditioning of calculations involving various polynomial bases, including $\{x^k\}$ and $\{T_k(x)\}$. A paper by Gautschi (1984) gives a particularly effective approach to this topic.

If a Chebyshev basis is adopted, there are usually three gains:

1. The coefficients generally decrease rapidly with the degree n of polynomial;
2. The coefficients converge individually with n;
3. The basis is well conditioned, so that methods such as collocation are well behaved numerically.

6.4 Best $\mathcal{L}_1$ approximation by Chebyshev interpolation

Up to now, we have concentrated on the $\mathcal{L}_\infty$ or minimax norm. However, the $\mathcal{L}_\infty$ norm is not the only norm for which Chebyshev polynomials can be shown to be minimal. Indeed, a minimality property holds, with a suitable weight function of the form $(1-x)^\gamma(1+x)^\delta$, in the $\mathcal{L}_1$ and $\mathcal{L}_2$ norms, and more generally in the $\mathcal{L}_p$ norm, where p is equal to 1 or an even integer, and this is true for all four kinds of Chebyshev polynomials. Here we look at minimality in the $\mathcal{L}_1$ norm.

The $\mathcal{L}_1$ norm (weighted by $w(x)$) of a function $f(x)$ on $[-1,1]$ is

$$\|f\|_1 := \int_{-1}^{1} w(x)\,|f(x)|\,\mathrm{d}x \qquad (6.35)$$

and the Chebyshev polynomials have the following minimality properties in $\mathcal{L}_1$.

Theorem 6.10 $2^{1-n}T_n(x)$ *($n>0$), $2^{-n}U_n(x)$, $2^{-n}V_n(x)$, $2^{-n}W_n(x)$ are the monic polynomials of minimal $\mathcal{L}_1$ norm with respect to the respective weight functions*

$$w(x) = \frac{1}{\sqrt{1-x^2}},\ 1,\ \frac{1}{\sqrt{1-x}},\ \frac{1}{\sqrt{1+x}}. \qquad (6.36)$$

Theorem 6.11 *The polynomial $p_{n-1}(x)$ of degree $n-1$ is a best $\mathcal{L}_1$ approximation to a given continuous function $f(x)$ with one of the four weights $w(x)$ given by (6.36) if $f(x) - p_{n-1}(x)$ vanishes at the n zeros of $T_n(x)$, $U_n(x)$, $V_n(x)$, $W_n(x)$, respectively, and at no other interior points of $[-1,1]$.*

(Note that the condition is sufficient but not necessary.)

Clearly Theorem 6.10 is a special case of Theorem 6.11 (with $f(x) = x^n$), and so it suffices to prove the latter. We first state a classical lemma on the characterisation of best $\mathcal{L}_1$ approximations (Rice 1964, Section 4–4).

Lemma 6.12 *If $f(x) - p_{n-1}(x)$ does not vanish on a set of positive measure (e.g., over the whole of a finite subinterval), where p_{n-1} is a polynomial of degree $n-1$ in x, then p_{n-1} is a best weighted $\mathcal{L}_1$ approximation to f on $[-1,1]$ if and only if*

$$I_n^{(r)} := \int_{-1}^{1} w(x)\ \mathrm{sgn}[f(x) - p_{n-1}(x)]\ \phi_r(x)\,\mathrm{d}x = 0 \tag{6.37}$$

for $r = 0, 1, \ldots, n-1$, where each $\phi_r(x)$ is any given polynomial of exact degree r.

Using this lemma, we can now establish the theorems.

Proof: (of Theorem 6.11 and hence of Theorem 6.10)

Clearly $\mathrm{sgn}(f(x) - p_{n-1}(x)) = \mathrm{sgn}\,P_n(x)$, where $P_r \equiv T_r,\ U_r,\ V_r,\ W_r$, respectively ($r = 0, 1, \ldots, n$).

Then, taking $\phi_r(x) = P_r(x)$ in (6.37) and making the usual change of variable,

$$I_n^{(r)} = \begin{cases} \displaystyle\int_0^\pi \mathrm{sgn}(\cos n\theta)\ \cos r\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^\pi \mathrm{sgn}(\sin(n+1)\theta)\ \sin(r+1)\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^\pi \mathrm{sgn}(\cos(n+\tfrac{1}{2})\theta)\ \cos(r+\tfrac{1}{2})\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^\pi \mathrm{sgn}(\sin(n+\tfrac{1}{2})\theta)\ \sin(r+\tfrac{1}{2})\theta\,\mathrm{d}\theta, \end{cases}$$

respectively. The proof that $I_n^{(r)} = 0$ is somewhat similar in each of these four cases.

Consider the first case. Here, since the zeros of $\cos n\theta$ occur at $(k - \frac{1}{2})\pi/n$ for $k = 1, \ldots, n$, we have

$$\begin{aligned} I_n^{(r)} &= \int_0^{\pi/2n} \cos r\theta\,\mathrm{d}\theta + \sum_{k=1}^{n-1} (-1)^k \int_{(k-\frac{1}{2})\pi/n}^{(k+\frac{1}{2})\pi/n} \cos r\theta\,\mathrm{d}\theta + (-1)^n \int_{(n-\frac{1}{2})\pi/n}^{\pi} \cos r\theta\,\mathrm{d}\theta \\ &= \frac{1}{r}\sin\frac{r\pi}{2n} + \sum_{k=1}^{n-1} (-1)^k \frac{1}{r}\left[\sin\frac{(k+\frac{1}{2})r\pi}{n} - \sin\frac{(k-\frac{1}{2})r\pi}{n}\right] + \end{aligned}$$

$$
\begin{aligned}
&\quad + (-1)^{n-1}\frac{1}{r}\sin\frac{(n-\frac{1}{2})r\pi}{n} \\
&= \frac{2}{r}\left[\sin\frac{r\pi}{2n} - \sin\frac{3r\pi}{2n} + \cdots + (-1)^{n-1}\sin\frac{(2n-1)r\pi}{2n}\right] \\
&= \frac{1}{r}\left[\sin\frac{r\pi}{n} - \left\{\sin\frac{r\pi}{n} + \sin\frac{2r\pi}{n}\right\} + \cdots + (-1)^{n-1}\sin\frac{(n-1)r\pi}{n}\right] \Big/ \cos\frac{r\pi}{2n} \\
&= 0.
\end{aligned}
$$

We can likewise show that $I_n^{(r)} = 0$ in each of the three remaining cases. Theorems 6.11 and 6.10 then follow very easily from Lemma 6.12 ●●

It follows (replacing n by $n+1$) that the nth degree polynomial $p_n(x)$ interpolating a function $f(x)$ at the zeros of one of the Chebyshev polynomials $T_{n+1}(x)$, $U_{n+1}(x)$, $V_{n+1}(x)$ or $W_{n+1}(x)$, which we showed how to construct in Section 6.3, will in many cases give a best weighted $\mathcal{L}_1$ approximation — subject only to the condition (which we cannot usually verify until after carrying out the interpolation) that $f(x) - p_n(x)$ vanishes nowhere else in the interval.

6.5 Near-minimax approximation by Chebyshev interpolation

Consider a continuous function $f(x)$ and denote the (first-kind) Chebyshev interpolation mapping by J_n. Then

$$
(J_n f)(x) = \sum_{k=1}^{n+1} f(x_k)\ell_k(x), \tag{6.38}
$$

by the Lagrange formula, and clearly J_n must be a projection, since (6.38) is linear in f and exact when f is a polynomial of degree n. From Lemma 5.13, J_n is near-minimax within a relative distance $\|J_n\|_\infty$.

Now

$$
|(J_n f)(x)| \le \sum_{k=1}^{n+1} \|f\|_\infty |\ell_k(x)|.
$$

Hence

$$
\begin{aligned}
\|J_n\|_\infty &= \sup_f \frac{\|J_n f\|_\infty}{\|f\|_\infty} \\
&= \sup_f \sup_{x\in[-1,1]} \frac{|(J_n f)(x)|}{\|f\|_\infty} \\
&\le \sup_f \sup_{x\in[-1,1]} \sum_{k=0}^{n} |\ell_k(x)| \\
&= \mu_n
\end{aligned} \tag{6.39}
$$

where

$$\mu_n = \sup_{x\in[-1,1]} \sum_{k=1}^{n+1} |\ell_k(x)|\,. \tag{6.40}$$

Now if $\sum_{k=0}^{n} |\ell_k(x)|$ attains its extremum at $x = \xi$, we can define a continuous function $\phi(x)$ such that

$$\begin{aligned} \|\phi\|_\infty &\le 1, \\ \phi(x_k) &= \operatorname{sgn}(\ell_k(\xi)). \end{aligned}$$

Then, from (6.38),

$$(J_n\phi)(\xi) = \sum_{k=1}^{n+1} |\ell_k(\xi)| = \mu_n,$$

whence

$$\|J_n\|_\infty \ge \|J_n\phi\|_\infty \ge \mu_n. \tag{6.41}$$

Inequalities (6.39) and (6.41) together give us

$$\|J_n\|_\infty = \mu_n.$$

What we have written so far applies to any Lagrange interpolation operator. If we specialise to first-kind Chebyshev interpolation, where $\ell_k(x)$ is as given by Corollary 6.4A, then we have the following asymptotic bound on $\|J_n\|_\infty$.

Theorem 6.13 *If $\{x_k\}$ are the zeros of $T_{n+1}(x)$, then*

1.

$$\mu_n = \frac{1}{\pi}\sum_{k=1}^{n+1}\left|\cot\frac{(k-\frac{1}{2})\pi}{2(n+1)}\right|,$$

2.

$$\mu_n = \frac{2}{\pi}\log n + 0.9625 + O(1/n) \text{ as } n \to \infty.$$

Proof: For the details of the proof, the reader is referred to Powell (1967) or Rivlin (1974). See also Brutman (1978). ●●

The following classical lemma then enables us to deduce convergence properties.

Lemma 6.14 (Jackson's theorem) *If $\omega(\delta)$ is the modulus of continuity of $f(x)$, then the minimax polynomial approximation $B_n f$ of degree n to f satisfies*

$$\|f - B_n f\|_\infty \le \omega(1/n).$$

Corollary 6.14A *If* $(J_n f)(x)$ *interpolates* $f(x)$ *in the zeros of* $T_{n+1}(x)$*, and if* $f(x)$ *is Dini–Lipschitz continuous, then* $(J_n f)(x)$ *converges uniformly to* $f(x)$ *as* $n \to \infty$*.*

Proof: By the definition of Dini–Lipschitz continuity, $\omega(\delta)\log\delta \to 0$ as $\delta \to 0$. By Theorem 5.12

$$\begin{aligned}
\|f - J_n f\|_\infty &\le (1 + \|J_n\|)_\infty \|f - B_n f\|_\infty \\
&\le (1 + \mu_n)\omega(1/n) \\
&\sim \frac{2}{\pi}\omega(1/n)\log n \\
&= -\frac{2}{\pi}\omega(\delta)\log\delta \quad (\delta = 1/n) \\
&\to 0 \text{ as } \delta \to 0;\ \text{i.e., as } n \to \infty. \ \bullet\bullet
\end{aligned}$$

In closing this chapter, we remind the reader that further interpolation results have been given earlier in Chapter 4 in the context of orthogonality. See in particular Sections 4.3.2 and 6.2.

6.6 Problems for Chapter 6

1. Prove Lemmas 6.3 and 6.4, and deduce Corollary 6.4A.

2. Prove Corollary 6.6A.

3. Find expressions for the coefficients (6.14) of the nth degree interpolating polynomial when $f(x) = \operatorname{sgn} x$ and $f(x) = |x|$, and compare these with the coefficients in the Chebyshev expansions (5.11) and (5.12).

4. List the possibilities of aliasing in the following interpolation situations:

   (a) Polynomials U_j of the second kind on the zeros of $T_{n+1}(x)$,

   (b) Polynomials V_j of the third kind on the zeros of $T_{n+1}(x)$,

   (c) Polynomials U_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (d) Polynomials T_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (e) Polynomials V_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (f) Polynomials U_j on the zeros of $U_{n+1}(x)$,

   (g) Polynomials T_j on the zeros of $U_{n+1}(x)$.

5. Give a proof of Theorem 6.11 for the case of the function U_r.

6. Using Theorem 6.11, consider the lacunary series partial sum

$$f_n(x) = \sum_{k=1}^{n} c_k U_{2^k-1}(x).$$

Assuming that the series is convergent to $f = \lim_{n\to\infty} f_n$, show that $f - f_n$, for instance, vanishes at the zeros of U_{2^n-1}. Give sufficient conditions for f_n to be a best $\mathcal{L}_1$ approximation to f for every n.

7. Show that the $n+1$ zeros of $T_{n+1}(z) - T_{n+1}(z^*)$ are distinct and lie on E_r, for a suitable fixed point z^* on E_r $(r > 1)$. Fixing r, find the zeros for the following choices of z^*:

   (a) $z^* = \frac{1}{2}(r + r^{-1})$,
   (b) $z^* = -\frac{1}{2}(r + r^{-1})$,
   (c) $z^* = \frac{1}{2}\mathrm{i}(r - r^{-1})$,
   (d) $z^* = -\frac{1}{2}\mathrm{i}(r - r^{-1})$.

8. If $f_n(z)$ is a polynomial of degree n interpolating $f(z)$, continuous on the ellipse E_r and analytic in its interior, find a set of interpolation points z_k $(k = 1, \ldots, n+1)$ on E_r such that

   (a) f_n is near-minimax within a computable relative distance σ_n on E_r, giving a formula for σ_n;
   (b) this result is valid as $r \to 1$; i.e., as the ellipse collapses to the line segment $[-1, 1]$.

   To effect (b), show that it is necessary to choose the interpolation points asymmetrically across the x-axis, so that points do not coalesce.

Chapter 7

Near-Best $\mathcal{L}_\infty$, $\mathcal{L}_1$ and $\mathcal{L}_p$ Approximations

7.1 Near-best $\mathcal{L}_\infty$ (near-minimax) approximations

We have already established in Section 5.5 that partial sums of first kind expansions

$$(S_n^T f)(x) = \sum_{k=0}^{n}{}' c_k U_k(x), \quad c_k = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x \tag{7.1}$$

yield near-minimax approximations within a relative distance of $O(\log n)$ in $\mathcal{C}[-1,1]$. Is this also the case for other kinds of Chebyshev polynomial expansions? The answer is in the affirmative, if we go about the expansion in the right way.

7.1.1 Second-kind expansions in $\mathcal{L}_\infty$

Consider the class $\mathcal{C}_{\pm 1}[-1,1]$ of functions continuous on $[-1,1]$ but constrained to vanish at ± 1. Let $S_n^{(2)} f$ denote the partial sum of the expansion of $f(x)/\sqrt{1-x^2}$ in Chebyshev polynomials of the second kind, $\{U_k(x) : k = 0,1,2,\ldots,n\}$, multiplied by $\sqrt{1-x^2}$. Then

$$(S_n^{(2)} f)(x) = \sqrt{1-x^2}\sum_{k=0}^{n} b_k U_k(x), \quad b_k = \frac{2}{\pi}\int_{-1}^{1} f(x)U_k(x)\,\mathrm{d}x. \tag{7.2}$$

If now we define

$$g(\theta) = \begin{cases} f(\cos\theta) & 0 \le \theta \le \pi \\ -f(\cos\theta) & -\pi \le \theta \le 0 \end{cases}$$

($g(\theta)$ being an odd, continuous and 2π-periodic function since $f(1) = f(-1) = 0$), then we obtain the equivalent Fourier sine series partial sum

$$(S_{n+1}^{FS} g)(\theta) = \sum_{k=0}^{n} b_k \sin(k+1)\theta, \quad b_k = \frac{2}{\pi}\int_0^{\pi} g(\theta)\sin(k+1)\theta\,\mathrm{d}\theta. \tag{7.3}$$

The operator S_{n+1}^{FS} can be identified as the restriction of the Fourier projection S_{n+1}^F to the space $\mathcal{C}_{2\pi,\mathrm{o}}^0$ of continuous functions that are both periodic of period 2π and odd; in fact we have $S_{n+1}^{FS} g = S_{n+1}^F g$ for odd functions g, where

$$(S_{n+1}^F g)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)\frac{\sin(n+\frac{3}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t. \tag{7.4}$$

If λ_n is the Lebesgue constant defined in (5.71)

$$\lambda_n = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\right|\,\mathrm{d}t$$

and partly tabulated in Table 5.1 on page 126, then, similarly to (5.75), we may show that

$$\left\|S_n^{(2)}\right\|_\infty = \left\|S_{n+1}^{FS}\right\|_\infty \le \lambda_{n+1} \qquad \text{(on the space } \mathcal{C}_{\pm 1}[-1,1]). \tag{7.5}$$

Therefore $(S_n^{(2)}f)(x)$ is near-minimax within a relative distance λ_{n+1}.

This constant λ_{n+1} is not, however, the best possible, as has been shown by Mason & Elliott (1995) — the argument of Section 5.5.1 falls down because the function

$$\mathrm{sgn}\left(\frac{\sin(n+\frac{3}{2})\theta}{\sin\frac{1}{2}\theta}\right)$$

is even, and cannot therefore be closely approximated by any function in $\mathcal{C}^0_{2\pi,\mathrm{o}}$.

However, g being odd, we may rewrite (7.4) as

$$\begin{aligned}(S_{n+1}^{FS}g)(\theta) &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\{g(t+\theta)-g(-t-\theta)\}\frac{\sin(n+\frac{3}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} g(t)\left\{\frac{\sin(n+\frac{3}{2})(t-\theta)}{\sin\frac{1}{2}(t-\theta)} - \frac{\sin(n+\frac{3}{2})(t+\theta)}{\sin\frac{1}{2}(t+\theta)}\right\}\mathrm{d}t \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} g(t)K_{n+1}^{FS}(\theta,t)\,\mathrm{d}t. \end{aligned} \tag{7.6}$$

This kernel $K_{n+1}^{FS}(\theta,t)$ is an odd function of θ and t, and an argument similar to that in Section 5.5.1 can now be used to show that

$$\left\|S_n^{(2)}\right\|_\infty = \left\|S_{n+1}^{FS}\right\|_\infty = \frac{1}{4\pi}\sup_\theta\int_{-\pi}^{\pi}\left|K_{n+1}^{FS}(\theta,t)\right|\,\mathrm{d}t = \lambda_n^{(2)}, \text{ say.} \tag{7.7}$$

Table 7.1: Lower bounds on $\lambda_n^{(2)}$

n	bound	n	bound	n	bound
1	1.327	10	1.953	100	2.836
2	1.467	20	2.207	200	3.114
3	1.571	30	2.362	300	3.278
4	1.653	40	2.474	400	3.394
5	1.721	50	2.561	500	3.484

Mason & Elliott (1995) have actually computed values of $\lambda_n^{(2)}$, which is no straightforward task since the points where the integrand $K_{n+1}^{FS}(\theta,t)$ changes sign are not in general easily determined. For a lower bound to the supremum for each n, however, we may evaluate the integral when $\theta = \pi/(2n+3)$, when the sign changes occur at the precisely-known points $t = 0$, $\pm 3\pi/(2n+3)$, $\pm 5\pi/(2n+3)$, …, $\pm\pi$. This gives the values shown in Table 7.1.

7.1.2 Third-kind expansions in $\mathcal{L}_\infty$

Following Mason & Elliott (1995) again, consider functions f in $\mathcal{C}_{-1}[-1,1]$, continuous on $[-1,1]$ but constrained to vanish at $x = -1$. Then the nth degree projection operator $S_n^{(3)}$, such that $S_n^{(3)} f$ is the partial sum of the expansion of $f(x)\sqrt{2/(1+x)}$ in Chebyshev polynomials of the third kind, $\{V_k(x) : k = 0,1,2,\ldots,n\}$, multiplied by $\sqrt{(1+x)/2}$, is defined by

$$\begin{aligned}(S_n^{(3)} f)(x) &= \sqrt{\frac{1+x}{2}} \sum_{k=0}^{n} c_k V_k(x) \\ &= \sum_{k=0}^{n} c_k \cos(k+\tfrac{1}{2})\theta \qquad (7.8)\end{aligned}$$

where $x = \cos\theta$ and

$$\begin{aligned}c_k &= \frac{1}{\pi}\int_{-1}^{1} \sqrt{\frac{2}{1-x}} f(x) V_k(x)\,\mathrm{d}x \\ &= \frac{2}{\pi}\int_{0}^{\pi} g(\theta)\cos(k+\tfrac{1}{2})\theta\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-2\pi}^{2\pi} g(\theta)\cos(k+\tfrac{1}{2})\theta\,\mathrm{d}\theta \qquad (7.9)\end{aligned}$$

with g defined as follows:

$$g(\theta) = \begin{cases} f(\cos\theta) & 0 \le \theta \le \pi \\ -g(2\pi-\theta) & \pi \le \theta \le 2\pi \\ g(-\theta) & -2\pi \le \theta \le 0. \end{cases}$$

The function $g(\theta)$ has been defined to be continuous (since $g(\pi) = f(-1) = 0$) and 4π-periodic, and is even about $\theta = 0$ and odd about $\theta = \pi$. Its Fourier expansion (in trigonometric functions of $\frac{1}{2}\theta$) therefore involves only terms in $\cos(2k+1)\frac{\theta}{2} = \cos(k+\frac{1}{2})\theta$ and is of the form (7.8) when truncated. From (7.8) and (7.9),

$$(S_n^{(3)} f)(x) = \frac{1}{2\pi}\int_{-2\pi}^{2\pi} g(t) \sum_{k=0}^{n} \cos(k+\tfrac{1}{2})t \cos(k+\tfrac{1}{2})\theta\,\mathrm{d}t.$$

Table 7.2: Values of $\lambda_n^{(3)}$

n	$\lambda_n^{(3)}$	n	$\lambda_n^{(3)}$	n	$\lambda_n^{(3)}$
1	1.552	10	2.242	100	3.140
2	1.716	20	2.504	200	3.420
3	1.832	30	2.662	300	3.583
4	1.923	40	2.775	400	3.700
5	1.997	50	2.864	500	3.790

We leave it as an exercise to the reader (Problem 1) to deduce that

$$(S_n^{(3)} f)(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t+\theta) \frac{\sin(n+1)t}{\sin \frac{1}{2}t} \, \mathrm{d}t.$$

Thus

$$\left| (S_n^{(3)} f)(x) \right| \leq \|g\|_\infty \lambda_n^{(3)} \tag{7.10}$$

where

$$\lambda_n^{(3)} = \frac{1}{\pi} \int_0^{\pi} \left| \frac{\sin(n+1)t}{\sin \frac{1}{2}t} \right| \mathrm{d}t. \tag{7.11}$$

Hence $\left\| S_n^{(3)} \right\|_\infty \leq \lambda_n^{(3)}$.

Arguing as in Section 5.5.1 as before, we again show that we have an equality

$$\left\| S_n^{(3)} \right\|_\infty = \lambda_n^{(3)}.$$

Numerical values of $\lambda_n^{(3)}$ are shown in Table 7.2, and clearly appear to approach those of λ_n (Table 5.1) asymptotically.

A fuller discussion is given by Mason & Elliott (1995), where it is conjectured that (as for λ_n in (5.77))

$$\lambda_n^{(3)} = \frac{4}{\pi^2} \log n + A_1 + O(1/n)$$

where $A_1 \simeq 1.2703$. (This follows earlier work by Luttman & Rivlin (1965) and by Cheney & Price (1970) on the asymptotic behaviour of λ_n.) Once more, then, we have obtained a near-minimax approximation within a relative distance asymptotic to $4\pi^{-2} \log n$.

For further detailed discussion of Lebesgue functions and constants for interpolation, see Brutman (1997).

7.2 Near-best $\mathcal{L}_1$ approximations

From Section 6.4 we would expect Chebyshev series partial sums to yield near-best $\mathcal{L}_1$ approximations with respect to the weights given in (6.36), namely $w(x) = 1/\sqrt{1-x^2}$, 1, $1/\sqrt{1-x}$, $1/\sqrt{1+x}$, since they already provide best $\mathcal{L}_1$ approximations for a function that is a polynomial of one degree higher. In fact, this can be shown to hold simply by pre-multiplying and post-dividing the functions expanded in Section 7.1 by the additional factor $\sqrt{1-x^2}$. The simplest case to consider here is that of the second-kind polynomials U_n, since the function expanded is then just the original function.

The partial sum of degree n of the second kind, for a continuous function $f(x)$, is defined by the projection

$$P_n^{(2)} : (P_n^{(2)} f)(x) = \sum_{k=0}^{n} b_k U_k(x), \quad b_k = \frac{2}{\pi} \int_{-1}^{1} \sqrt{1-x^2} f(x) U_k(x) \,\mathrm{d}x. \tag{7.12}$$

Defining the function g by

$$g(\theta) = \sin\theta \, f(\cos\theta), \tag{7.13}$$

which is naturally an odd periodic continuous function, we see that

$$b_k = \frac{2}{\pi} \int_0^{\pi} g(\theta) \sin(k+1)\theta \,\mathrm{d}\theta, \tag{7.14}$$

as in (7.3), and $(P_n^{(2)} f)(\cos\theta) = (S_{n+1}^{FS} g)(\theta)$.

Now, treating $f(x)$ as defined on $[-1, 1]$ and $g(\theta)$ as defined on $[-\pi, \pi]$ so that

$$\|g\|_1 = \int_{-\pi}^{\pi} |g(\theta)| \,\mathrm{d}\theta = \int_{-\pi}^{\pi} |\sin\theta \, f(\cos\theta)| \,\mathrm{d}\theta = 2 \int_{-1}^{1} f(x) \,\mathrm{d}x = 2 \|f\|_1 ,$$

we have

$$\begin{aligned}
\left\| P_n^{(2)} f \right\|_1 &= \tfrac{1}{2} \left\| S_{n+1}^{FS} g \right\|_1 \\
&= \tfrac{1}{2} \int_{-\pi}^{\pi} \left| \frac{1}{4\pi} \int_{-\pi}^{\pi} g(t) K_{n+1}^{FS}(\theta, t) \,\mathrm{d}t \right| \mathrm{d}\theta \\
&\le \tfrac{1}{2} \int_{-\pi}^{\pi} |g(t)| \,\mathrm{d}t \, \frac{1}{4\pi} \sup_t \int_{-\pi}^{\pi} \left| K_{n+1}^{FS}(\theta, t) \right| \mathrm{d}\theta \\
&= \tfrac{1}{2} \|g\|_1 \, \lambda_n^{(2)} \\
&= \|f\|_1 \, \lambda_n^{(2)},
\end{aligned}$$

where $\lambda_n^{(2)}$ is the constant defined in (7.7) above.

Hence

$$\left\| P_n^{(2)} \right\|_1 \leq \lambda_n^{(2)}. \tag{7.15}$$

Thus $\lambda_n^{(2)}$ is a bound on $\left\| P_n^{(2)} \right\|_1$, just as it was a bound on $\left\| S_n^{(2)} \right\|_\infty$ in Section 7.1.1, and so $(P_n^{(2)} f)(x)$, given by (7.12), is a near-best $\mathcal{L}_1$ approximation within the relative distance $\lambda_n^{(2)}$ defined in (7.7).

The discussion above is, we believe, novel. Freilich & Mason (1971) established that $\left\| P_n^{(2)} \right\|_1$ is bounded by λ_n, but the new bound (7.15) is smaller by about 0.27.

If we define $P_n^{(1)}$ and $P_n^{(3)}$ to be the corresponding partial sum projections of the first and third kinds,

$$\begin{aligned} (P_n^{(1)} f)(x) &= \frac{1}{\sqrt{1-x^2}} \sum_{k=0}^{n}{}' c_k T_k(x), \\ c_k &= \frac{2}{\pi} \int_{-1}^{1} f(x) T_k(x) \, \mathrm{d}x, \end{aligned} \tag{7.16}$$

$$\begin{aligned} (P_n^{(3)} f)(x) &= \frac{1}{2\sqrt{1-x}} \sum_{k=0}^{n} c_k V_k(x), \\ c_k &= \frac{1}{\pi} \int_{-1}^{1} \sqrt{2(1+x)} f(x) V_k(x) \, \mathrm{d}x, \end{aligned} \tag{7.17}$$

then it is straightforward to show in a similar way (see Problem 2) that

$$\left\| P_n^{(1)} \right\|_1 \leq \lambda_n \quad \text{(classical Lebesgue constant)}$$

and

$$\left\| P_n^{(3)} \right\|_1 \leq \lambda_n^{(3)} \quad \text{(given by (7.11))}.$$

7.3 Best and near-best $\mathcal{L}_p$ approximations

The minimal $\mathcal{L}_\infty$ and $\mathcal{L}_1$ properties of the weighted Chebyshev polynomials, discussed in Sections 3.3 and 6.4, are in fact special examples of general $\mathcal{L}_p$ minimality properties, which are discussed by Mason & Elliott (1995).

Theorem 7.1 *The monic polynomials* $2^{1-n} T_n(z)$, $2^{-n} U_n(z)$, $2^{-n} V_n(z)$, $2^{-n} W_n(z)$ *minimise the* $\mathcal{L}_p$ *norm*

$$\left[\int_{-1}^{1} w(x) \left| P_n(x) \right|^p \, \mathrm{d}x \right]^{\frac{1}{p}} \quad (1 < p < \infty) \tag{7.18}$$

over all monic polynomials $P_n(x)$ with

$$w(x) = (1-x)^{\frac{1}{2}(\alpha-1)}(1+x)^{\frac{1}{2}(\beta-1)}$$

for the respective values

$$(\alpha, \beta) = (0,0),\ (p,p),\ (0,p),\ (p,0).$$

The proof of this result depends on the characterisation of the best $\mathcal{L}_p$ approximation according to the following result, which we state without proof.

Lemma 7.2 *The $\mathcal{L}_p$ norm (7.18) is minimised if and only if*

$$\int_{-1}^{1} w(x)\, |P_n(x)|^{p-2}\, P_n(x) P_k(x)\, \mathrm{d}x = 0, \quad \forall k < n. \tag{7.19}$$

Proof: (of Theorem 7.1)

We shall concentrate on the first case, that of the first kind polynomials $T_n(x)$, and leave the remaining cases as exercises for the reader (Problem 3).

Define

$$P_n(x) = T_n(x), \quad w(x) = 1/\sqrt{(}1-x^2).$$

Then

$$\int_{-1}^{1} w(x)\, |P_n(x)|^{p-2}\, P_n(x) P_k(x)\, \mathrm{d}x = \int_0^{\pi} |\cos n\theta|^{p-2} \cos n\theta \cos k\theta \, \mathrm{d}\theta.$$

Now, for $0 \le y \le 1$, define

$$C_n(\theta, y) = \begin{cases} 1 & (|\cos n\theta| \le y), \\ 0 & (|\cos n\theta| > y). \end{cases}$$

Then if $y = \cos\eta$ we have $C_n(\theta, y) = 1$ over each range

$$\frac{(r-1)\pi + \eta}{n} \le \theta \le \frac{r\pi - \eta}{n}, \quad r = 1, 2, \ldots, n.$$

Thus, for any integer j with $0 < j < 2n$,

$$\begin{aligned}
\int_0^{\pi} C_n(\theta, y) \cos j\theta \, \mathrm{d}\theta &= \sum_{r=1}^{n} \int_{((r-1)\pi+\eta)/n}^{(r\pi-\eta)/n} \cos j\theta \, \mathrm{d}\theta \\
&= \sum_{r=1}^{n} \frac{1}{j} \left[\sin \frac{j\{(r-1)\pi + \eta\}}{n} - \sin \frac{j\{r\pi - \eta\}}{n} \right] \\
&= \sum_{r=1}^{2n} \frac{1}{j} \sin \frac{j\{(r-1)\pi + \eta\}}{n} \\
&= 0.
\end{aligned}$$

But now, for $0 \le k < n$,

$$\begin{aligned}
&\int_{-1}^{1} w(x)\,|P_n(x)|^{p-2}\,P_n(x)P_k(x)\,\mathrm{d}x \\
&= \int_0^{\pi} |\cos n\theta|^{p-2} \cos n\theta\ \cos k\theta\,\mathrm{d}\theta \\
&= \int_0^{\pi} |\cos n\theta|^{p-2}\,\tfrac{1}{2}[\cos(n+k)\theta + \cos(n-k)\theta]\,\mathrm{d}\theta \\
&= \int_0^{\pi} \left\{ \frac{1}{p-1} \int_0^1 y^{p-1}(1 - C_n(\theta, y))\,\mathrm{d}y \right\} \tfrac{1}{2}[\cos(n+k)\theta + \cos(n-k)\theta]\,\mathrm{d}\theta \\
&= \frac{1}{p-1} \int_0^1 y^{p-1} \left\{ \int_0^{\pi} (1 - C_n(\theta, y))\,\tfrac{1}{2}[\cos(n+k)\theta + \cos(n-k)\theta]\,\mathrm{d}\theta \right\} \mathrm{d}y \\
&= 0.
\end{aligned}$$

The result then follows from Lemma 7.2. ●●

(An alternative method of proof is to translate into polynomial terms the result on trigonometric polynomials, due to S. N. Bernstein, given in Achieser's book (Achieser 1956, Section 10).)

7.3.1 Complex variable results for elliptic-type regions

It is possible to obtain bounds for norms of projections, and hence measures of near-best $\mathcal{L}_p$ approximation, by using ideas of convexity over a family of $\mathcal{L}_p$ measure spaces for $1 \le p \le \infty$ (Mason 1983*b*, Mason 1983*a*). However, the settings for which there are results have been restricted to ones involving generalised complex Chebyshev series — based on results for Laurent series. Mason & Chalmers (1984) give $\mathcal{L}_p$ results for Fourier, Taylor and Laurent series; moreover Chalmers & Mason (1984) show these to be minimal projections on appropriate analytic function spaces. The settings, involving projection from space X to space Y, where $\mathcal{A}(D)$ denotes the space of functions analytic in D and continuous on $\bar{D}$, are:

1. Chebyshev, first kind: $X = \mathcal{A}(D_\rho)$, where D_ρ is the elliptical domain $\{z : \left|z + \sqrt{z^2-1}\right| < \rho\}$; $Y = Y_1 = \Pi_n$ (polynomials of degree n in z);

   $$P = G_n$$

   where G_n is the Chebyshev first-kind series projection of $\mathcal{A}(D_\rho)$ into Π_n.

2. Chebyshev, second kind: $X = \{f(z) = \sqrt{z^2-1}F(z),\ F \in \mathcal{A}(D_\rho)\}$, $Y = Y_2 = \{f(z) = \sqrt{z^2-1}F(z),\ F \in \Pi_n\}$;

   $$P = H^*_{n-1} : H^*_{n-1}f = \sqrt{z^2-1}H_{n-1}F,$$

where H_n is the Chebyshev second kind series projection of $\mathcal{A}(D_\rho)$ into Π_n.

3. Generalised Chebyshev: $X = \mathcal{A}(\{z : \rho_1 < \left|z + \sqrt{z^2 - 1}\right| < \rho_2\})$ (annulus between two ellipses); $Y = Y_1 \oplus Y_2$;

$$P = J_n = G_n + H_{n-1}^*.$$

Then it is proved by Mason (1983*b*), using convexity arguments, that for each of the three projections above

$$\|P\|_p \leq (\sigma_{2n})^{|2p^{-1}-1|} \quad (1 \leq p \leq \infty) \tag{7.20}$$

where

$$\sigma_n = \frac{1}{n}\int_0^\pi \left|\frac{\sin(n+1)\theta}{\sin\theta}\right| \mathrm{d}\theta.$$

Note that $\sigma_{2n} = \lambda_n$. So the generalised expansion is proved to be as close to minimax as the (separated) first kind one.

For $p = 1$, $p = \infty$, we obtain bounds increasing as $4\pi^{-2}\log n$, while $\|P\|_p \to 1$ as $p \to 2$.

It follows also (Chalmers & Mason 1984) that J_n is a minimal projection; indeed, this appears to be the only such result for Chebyshev series. The component projections G_n and H_{n-1}^* are essentially odd and even respectively, and correspond to the cosine and sine parts of a full Fourier series. In contrast, the projection G_n is not minimal.

The earliest near-best results for $\mathcal{L}_\infty$ and $\mathcal{L}_1$ approximation on elliptic domains appear to be those of Geddes (1978) and Mason (1978). See also Mason & Elliott (1993) for detailed results for all individual cases.

We should also note that it has long been known that

$$\|P\|_p \leq C_p \tag{7.21}$$

where C_p is some constant independent of n. Although this is superficially stronger than (7.20) from a theoretical point of view, the bounds (7.20) are certainly small for values of n up to around 500. Moreover, it is known that $C_p \to \infty$ as $p \to \infty$. See Zygmund (1959) for an early derivation of this result, and Mhaskar & Pai (2000) for a recent discussion.

7.4 Problems for Chapter 7

1. Show that

$$\cos(k+\tfrac{1}{2})t\cos(k+\tfrac{1}{2})\theta = \tfrac{1}{2}[\cos(k+\tfrac{1}{2})(t+\theta) + \cos(k+\tfrac{1}{2})(t-\theta)]$$

and that

$$\sum_{k=0}^{n} \cos(k+\tfrac{1}{2})u = \frac{\sin(n+1)u}{2\sin\frac{1}{2}u}.$$

Hence prove that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sum_{k=0}^{n}\cos(k+\tfrac{1}{2})t\cos(k+\tfrac{1}{2})\theta\,\mathrm{d}\theta = \frac{1}{\pi}\int_{0}^{\pi} g(t+\theta)\frac{\sin(n+1)t}{\sin\frac{1}{2}t}\,\mathrm{d}t$$

by showing that the pair of integrals involved are equal.

(This completes the proof of Section 7.1.2, showing that the weighted third-kind expansion has a partial sum which is near-minimax.)

2. Show that $\left\|P_n^{(1)}\right\|_1 \le \lambda_n$ and $\left\|P_n^{(3)}\right\|_1 \le \lambda_n^{(3)}$, where λ_n is the classical Lebesgue constant and $\lambda_n^{(3)}$ is given by (7.11).

3. Prove Theorem 7.1 in the case of polynomials of the second and third kinds.

4. If S_n is a partial sum of a Fourier series

$$(S_n f)(\theta) = \tfrac{1}{2}a_0 + \sum_{k=0}^{n}(a_k\cos k\theta + b_k\sin k\theta),$$

show how this may be written, for suitably defined functions, as a combined first-kind and (weighted) second-kind Chebyshev expansion.

[Hint: $f(\theta) = F(\cos\theta) + \sin\theta\, G(\cos\theta)$ = even part of f + odd part of f.]

5. Consider the Fejér operator $\tilde{F}_n$, which takes the mean of the first n partial sums of the Fourier series.

   (a) Show that $\tilde{F}_n$ is not a projection.

   (b) Show that

$$(\tilde{F}_n f)(\theta) = \frac{1}{n\pi}\int_{0}^{2\pi} f(t)\tilde{\sigma}_n(t-\theta)\,\mathrm{d}t$$

   where

$$\tilde{\sigma}_n(\theta) = \frac{\sin\frac{1}{2}(n+1)\theta\ \sin\frac{1}{2}n\theta}{2\sin\frac{1}{2}\theta}.$$

   (c) Show that $(\tilde{F}_n f)(\theta)$, under the transformation $x = \cos\theta$, becomes a combined third-kind and fourth-kind Chebyshev-Fejér sum, each part being appropriately weighted.

6. Derive the basic result for $p = \infty$, namely $\|P\|_\infty \le \sigma_{2n} = \lambda_n$, for the three projections listed in Section 7.3.1.

7. Derive the corresponding basic results for $p = 1$.

   Would it be possible to obtain a better set of results in this case by using an odd kernel, like that used in (7.6)?

8. Note that $\|P\|_2 = 1$ in Section 7.3.1 and that it is known that $\|P\|_p$ is bounded for any fixed p in the range $1 < p < \infty$. Discuss whether there is a 'better' result than the one quoted.

   (You might like to consider both the practical case $n \leq 500$ and the theoretical case of arbitrarily large n.)

9. Investigate the validity of letting $p \to 1$ in the results of Section 7.3.1, when the interior of the ellipse collapses to the interval $[-1, 1]$.

10. Compute by hand the bounds for $\left\| S_n^{(2)} \right\|_\infty$ in the case $n = 0$.

11. Compute some numerical values of $\lambda_n^{(2)}$ and compare them with the lower bounds given in Table 7.1.

Chapter 8

Integration Using Chebyshev Polynomials

In this chapter we show how Chebyshev polynomials and some of their fundamental properties can be made to play an important part in two key techniques of numerical integration.

- Gaussian quadrature estimates an integral by combining values of the integrand at zeros of orthogonal polynomials. We consider the special case of *Gauss–Chebyshev quadrature*, where particularly simple procedures follow for suitably weighted integrands.
- One can approximately integrate a function by expanding it in a series and then integrating a partial sum of the series. We show that, for Chebyshev expansions, this process — essentially the *Clenshaw–Curtis method* — is readily analysed and again provides a natural procedure for appropriately weighted integrands.

Although this could be viewed as an 'applications' chapter, which in an introductory sense it certainly is, our aim here is primarily to derive further basic properties of Chebyshev polynomials.

8.1 Indefinite integration with Chebyshev series

If we wish to approximate the indefinite integral

$$h(X) = \int_{-1}^{X} w(x) f(x)\,\mathrm{d}x,$$

where $-1 < X \leq 1$, it may be possible to do so by approximating $f(x)$ on $[-1, 1]$ by an nth degree polynomial $f_n(x)$ and integrating $w(x)f_n(x)$ between -1 and X, giving the approximation

$$h(X) \simeq h_n(X) = \int_{-1}^{X} w(x) f_n(x)\,\mathrm{d}x. \tag{8.1}$$

Suppose, in particular, that the weight $w(x)$ is one of the four functions

$$w(x) = \frac{1}{\sqrt{1-x^2}},\ 1,\ \frac{1}{\sqrt{1-x}},\ \frac{1}{\sqrt{1+x}}, \tag{8.2}$$

and that we take $f_n(x)$ as the partial sum of the expansion of $f(x)$ in Chebyshev polynomials of the corresponding one of the four kinds

$$P_k(x) = T_k(x),\ U_k(x),\ V_k(x),\ W_k(x). \tag{8.3}$$

Then we can use the fact that (excluding the case where $P_k(x) = T_k(x)$ with $k = 0$)

$$\int_{-1}^{X} w(x)P_k(x)\,\mathrm{d}x = C_k(X)Q_k(X) - C_k(-1)Q_k(-1)$$

where

$$Q_k(X) = U_{k-1}(X),\ T_{k+1}(X),\ W_k(X),\ V_k(X) \tag{8.4a}$$

and

$$C_k(X) = -\frac{\sqrt{1-X^2}}{k},\ \frac{1}{k+1},\ 2\frac{\sqrt{1-X}}{k+\frac{1}{2}},\ -2\frac{\sqrt{1+X}}{k+\frac{1}{2}}, \tag{8.4b}$$

respectively. (Note that $C_k(-1) = 0$ in the first and fourth cases.) This follows immediately from the fact that if $x = \cos\theta$ then we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\sin k\theta &= -\frac{k\cos k\theta}{\sin\theta}, \\
\frac{\mathrm{d}}{\mathrm{d}x}\cos(k+1)\theta &= \frac{(k+1)\sin(k+1)\theta}{\sin\theta}, \\
\frac{\mathrm{d}}{\mathrm{d}x}\sin(k+\tfrac{1}{2})\theta &= -\frac{(k+\frac{1}{2})\cos(k+\frac{1}{2})\theta}{\sin\theta}, \\
\frac{\mathrm{d}}{\mathrm{d}x}\cos(k+\tfrac{1}{2})\theta &= \frac{(k+\frac{1}{2})\sin(k+\frac{1}{2})\theta}{\sin\theta}.
\end{aligned}$$

In the excluded case, we use

$$\frac{\mathrm{d}}{\mathrm{d}x}\theta = -\frac{1}{\sin\theta}$$

to give

$$\int_{-1}^{X} \frac{1}{\sqrt{1-x^2}}T_0(x)\,\mathrm{d}x = \arccos(-1) - \arccos X = \pi - \arccos X.$$

Thus, for each of the weight functions (8.2) we are able to integrate the weighted polynomial and obtain the approximation $h_n(X)$ explicitly. Suppose that

$$f_n(x) = {\sum_{k=0}^{n}}' a_k T_k(x)\ [P_k = T_k] \ \text{ or } \sum_{k=0}^{n} a_k P_k(x)\ [P_k = U_k,\ V_k,\ W_k]. \tag{8.5}$$

Then in the first case

$$\begin{aligned}
h_n(X) &= {\sum_{k=0}^{n}}' a_k \int_{-1}^{X} w(x)T_k(x)\,\mathrm{d}x = \\
&= \tfrac{1}{2}a_0(\pi - \arccos X) - \sum_{k=1}^{n} a_k \frac{\sqrt{1-X^2}}{k}U_{k-1}(X),
\end{aligned} \tag{8.6}$$

while in the second, third and fourth cases

$$h_n(X) = \sum_{k=0}^{n} a_k \int_{-1}^{X} w(x) P_k(x)\,\mathrm{d}x = \sum_{k=0}^{n} a_k \left[C_k(x) Q_k(x)\right]_{-1}^{X}. \tag{8.7}$$

The above procedure is a very reliable one, as the following theorem demonstrates.

Theorem 8.1 *If $f(x)$ is $\mathcal{L}_2$-integrable with respect to one of the weights $w(x)$, as defined by (8.2), and $h_n(X)$ is defined by (8.6) or (8.7) as appropriate, if $Q_k(X)$ and $C_k(X)$ are defined by (8.4), and if a_k are the* exact *coefficients of the expansion of $f(x)$ in Chebyshev polynomials of the corresponding kind, then $h_n(X)$ converges uniformly to $h(X)$ on $[-1, 1]$.*

Proof: The idea of the proof is the same in all four cases. We give details of the second case here, and leave the others as exercises (Problems 1 and 2).

For $P_k = U_k$, $w = 1$,

$$\begin{aligned} h_n(X) &= \int_{-1}^{X} f_n(x)\,\mathrm{d}x \\ &= \int_{-1}^{X} \sum_{k=0}^{n} a_k \sin(k+1)\theta\,\mathrm{d}\theta. \end{aligned}$$

Thus the integrand is the partial Fourier sine series expansion of $\sin\theta\, f(\cos\theta)$, which converges in $\mathcal{L}_2$ and hence in $\mathcal{L}_1$ (Theorems 5.2 and 5.5).

Now

$$\begin{aligned} \|h - h_n\|_\infty &= \max_X \left| \int_{-1}^{X} \{f(x) - f_n(x)\}\,\mathrm{d}x \right| \\ &\le \max_X \int_{-1}^{X} |f(x) - f_n(x)|\,\mathrm{d}x \\ &= \int_{-1}^{1} |f(x) - f_n(x)|\,\mathrm{d}x \\ &= \int_{0}^{\pi} \left| \sin\theta\, f(\cos\theta) - \sum_{k=0}^{n} a_k \sin(k+1)\theta \right| \mathrm{d}\theta \\ &\to 0, \quad n \to \infty. \end{aligned}$$

Hence h_n converges uniformly to h. ●●

The coefficients a_k in (8.5) have been assumed to be exactly equal to the relevant Chebyshev series coefficients. In practice, we most often approximate these by the corresponding coefficients in a Chebyshev interpolation polynomial (see Chapter 6) — effectively evaluating the integral that defines a_k by the trapezoidal rule (see Section 6.2). In some circumstances, we may need to calculate the Chebyshev coefficients more accurately than this.

The method followed above is equivalent to methods well known in the literature. For the first choice ($P_k = T_k$) the method is that of Clenshaw & Curtis (1960) and for the second choice ($P_k = U_k$) that of Filippi (1964).

The analysis of Section 8.1 is taken mainly from Mason & Elliott (1995, and related papers).

8.2 Gauss–Chebyshev quadrature

Suppose that we now wish to calculate a definite integral of $f(x)$ with weight $w(x)$, namely

$$I = \int_a^b f(x)w(x)\,\mathrm{d}x. \tag{8.8}$$

Suppose also that I is to be approximated in the form

$$I \simeq \sum_{k=1}^{n} A_k f(x_k) \tag{8.9}$$

where A_k are certain coefficients and $\{x_k\}$ are certain abscissae in $[a, b]$ (all to be determined). The idea of Gauss quadrature is to find that formula (8.9) that gives an exact result for all polynomials of as high a degree as possible.

If $J_{n-1}f(x)$ is the polynomial of degree $n-1$ which interpolates $f(x)$ in any n distinct points $x_1, \ldots, x_n$, then

$$J_{n-1}f(x) = \sum_{k=1}^{n} f(x_k)\ell_k(x) \tag{8.10}$$

where ℓ_k is the Lagrange polynomial (as in (6.5))

$$\ell_k(x) = \prod_{\substack{r=1 \\ r\neq k}}^{n} \left(\frac{x - x_r}{x_k - x_r}\right) \tag{8.11}$$

The polynomial $J_{n-1}f(x)$ has the integral

$$\begin{aligned} I_n &= \int_a^b J_{n-1}f(x)w(x)\,\mathrm{d}x \\ &= \sum_{k=1}^{n} f(x_k) \int_a^b w(x)\ell_k(x)\,\mathrm{d}x \\ &= \sum_{k=1}^{n} A_k f(x_k) \end{aligned}$$

provided that the coefficients A_k are chosen to be

$$A_k = \int_a^b w(x)\ell_k(x)\,\mathrm{d}x. \tag{8.12}$$

With any n distinct abscissae, therefore, and with this choice (8.12) of coefficients, the formula (8.9) certainly gives an exact result whenever $f(x)$ is a polynomial of degree $n-1$ or less. We can improve on this degree, however, by a suitable choice of abscissae.

Notice too that, for general abscissae, there is no control over the signs and magnitudes of the coefficients A_k, so that evaluation of the formula (8.9) may involve heavy cancellation between large terms of opposite signs, and consequent large rounding error. When we choose the abscissae to maximise the degree of exactness, however, it can be shown that this problem ceases to arise.

Theorem 8.2 *If x_k $(k = 1,\ldots,n)$ are the n zeros of $\phi_n(x)$, and $\{\phi_k : k = 0,1,2,\ldots\}$ is the system of polynomials, ϕ_k having the exact degree k, orthogonal with respect to $w(x)$ on $[a,b]$, then (8.9) with coefficients (8.12) gives an exact result whenever $f(x)$ is a polynomial of degree $2n-1$ or less. Moreover, all the coefficients A_k are positive in this case.*

Proof: Since $\phi_n(x)$ is a polynomial exactly of degree n, any polynomial $f(x)$ of degree $2n-1$ can be written (using long division by ϕ_n) in the form

$$f(x) = \phi_n(x)Q(x) + J_{n-1}f(x)$$

where $Q(x)$ and $J_{n-1}f(x)$ are polynomials each of degree at most $n-1$. Then

$$\int_a^b f(x)w(x)\,\mathrm{d}x = \int_a^b \phi_n(x)Q(x)w(x)\,\mathrm{d}x + \int_a^b J_{n-1}f(x)w(x)\,\mathrm{d}x. \tag{8.13}$$

Now $\phi_n(x)$ is orthogonal to all polynomials of degree less than n, so that the first integral on the right-hand side of (8.13) vanishes. Thus

$$\begin{aligned}\int_a^b f(x)w(x)\,\mathrm{d}x &= \int_a^b J_{n-1}f(x)w(x)\,\mathrm{d}x \\ &= \sum_{k=1}^n A_k J_{n-1}f(x_k)\end{aligned}$$

since the coefficients have been chosen to give an exact result for polynomials of degree less than n. But now

$$f(x_k) = \phi_n(x_k)Q(x_k) + J_{n-1}f(x_k) = J_{n-1}f(x_k),$$

since x_k is a zero of $\phi_n(x)$. Hence

$$\int_a^b f(x)w(x)\,\mathrm{d}x = \sum_{k=1}^n A_k f(x_k),$$

and so (8.9) gives an exact result for $f(x)$, as required.

To show that the coefficients A_k are positive, we need only notice that $\ell_k(x)^2$ is a polynomial of degree $2n-2$, and is therefore integrated exactly, so that

$$A_k \equiv \sum_{j=1}^{n} A_j \ell_k(x_j)^2 = \int_a^b \ell_k(x)^2 w(x)\,\mathrm{d}x > 0$$

for each k. ●●

Thus we can expect to obtain very accurate integrals with the formula (8.9), and the formula should be numerically stable.

When the interval $[a,b]$ is $[-1,1]$ and the orthogonal polynomials $\phi_n(x)$ are one of the four kinds of Chebyshev polynomials, then the weight function $w(x)$ is $(1-x^2)^{-\frac{1}{2}}$, $(1-x^2)^{\frac{1}{2}}$, $(1+x)^{\frac{1}{2}}(1-x)^{-\frac{1}{2}}$ or $(1-x)^{\frac{1}{2}}(1+x)^{-\frac{1}{2}}$ and the zeros x_k are known explicitly. It remains to determine the coefficients A_k, which we may do by making use of the following lemma.

Lemma 8.3

$$\begin{aligned} ⨍_0^\pi \frac{\cos n\theta}{\cos\theta - \cos\phi}\,\mathrm{d}\theta &= \pi\frac{\sin n\phi}{\sin\phi}, \\ ⨍_0^\pi \frac{\sin n\theta\ \sin\theta}{\cos\theta - \cos\phi}\,\mathrm{d}\theta &= -\pi\cos n\phi, \end{aligned}$$

for any ϕ in $[0,\pi]$, $n = 1, 2, 3, \ldots$.

(We have stated this lemma in terms of the 'Cauchy principal value' integral $⨍ \cdots \mathrm{d}\theta$ since, if we allow ϕ to take an arbitrary value, the integrands have a non-integrable singularity at $\theta = \phi$. However, when we come to apply the lemma in this chapter, $\theta = \phi$ will always turn out to be a zero of the numerator, so that the singularity will in fact be removable and the principal value integrals will be equivalent to integrals in the ordinary sense.)

Proof: The lemma can be proved by induction on n, provided that we first establish the $n = 0$ case of the first result

$$⨍_0^\pi \frac{1}{\cos\theta - \cos\phi}\,\mathrm{d}\theta = 0.$$

We may do this as follows. Since $\cos\theta$ is an even function, we have

$$\begin{aligned} &⨍_0^\pi \frac{1}{\cos\theta - \cos\phi}\,\mathrm{d}\theta = \\ &= \tfrac{1}{2} ⨍_{-\pi}^\pi \frac{1}{\cos\theta - \cos\phi}\,\mathrm{d}\theta \\ &= ⨍_{-\pi}^\pi \frac{\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta}{(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{\mathrm{i}\phi})(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\phi})} \end{aligned}$$

$$
\begin{aligned}
&= \oint_{|z|=1} \frac{-\mathrm{i}\,\mathrm{d}z}{(z-\mathrm{e}^{\mathrm{i}\phi})(z-\mathrm{e}^{-\mathrm{i}\phi})} \\
&= \frac{-\mathrm{i}}{\mathrm{e}^{\mathrm{i}\phi}-\mathrm{e}^{-\mathrm{i}\phi}} \left[\oint_{|z|=1} \frac{\mathrm{d}z}{z-\mathrm{e}^{\mathrm{i}\phi}} - \oint_{|z|=1} \frac{\mathrm{d}z}{z-\mathrm{e}^{-\mathrm{i}\phi}} \right] \\
&= \frac{-1}{2\sin\phi}[\mathrm{i}\pi - \mathrm{i}\pi] = 0.
\end{aligned}
$$

We leave the subsequent induction as an exercise (Problem 3). ●●

The evaluation of A_k can now be carried out.

Theorem 8.4 *In the Gauss–Chebyshev formula*

$$
\int_{-1}^{1} f(x)w(x)\,\mathrm{d}x \simeq \sum_{k=1}^{n} A_k f(x_k), \tag{8.14}
$$

where $\{x_k\}$ are the n zeros of $\phi_n(x)$, the coefficients A_k are as follows:

1. *For $w(x) = (1-x^2)^{-\frac{1}{2}}$, $\phi_n(x) = T_n(x)$:*

$$
A_k = \frac{\pi}{n}.
$$

2. *For $w(x) = (1-x^2)^{\frac{1}{2}}$, $\phi_n(x) = U_n(x)$:*

$$
A_k = \frac{\pi}{n+1}(1-x_k^2).
$$

3. *For $w(x) = (1-x)^{-\frac{1}{2}}(1+x)^{\frac{1}{2}}$, $\phi_n(x) = V_n(x)$:*

$$
A_k = \frac{\pi}{n+\frac{1}{2}}(1+x_k).
$$

4. *For $w(x) = (1-x)^{\frac{1}{2}}(1+x)^{-\frac{1}{2}}$, $\phi_n(x) = W_n(x)$:*

$$
A_k = \frac{\pi}{n+\frac{1}{2}}(1-x_k).
$$

Proof: We prove case 1 and leave case 2 as an exercise (Problem 4). We shall prove cases 3 and 4 a little later.

In case 1, writing

$$
x_k = \cos\theta_k = \cos\frac{(k-\frac{1}{2})\pi}{n}
$$

for the zeros of $T_n(x)$,

$$\begin{aligned} A_k &= \int_{-1}^{1} \frac{T_n(x)}{(x - x_k)\, n\, U_{n-1}(x_k)} \frac{\mathrm{d}x}{\sqrt{1-x^2}} \\ &= \int_0^{\pi} \frac{\cos n\theta\ \sin\theta_k}{(\cos\theta - \cos\theta_k)\, n\ \sin n\theta_k}\, \mathrm{d}\theta \\ &= \frac{\pi}{n}, \end{aligned}$$

using Corollary 6.4A and Lemma 8.3. ●●

Case 1 above is particularly convenient to use, since all weights are equal and the formula (8.9) can thus be evaluated with just $n-1$ additions and one multiplication.

EXAMPLE 8.1: To illustrate the exactness of (8.9) for polynomials of degree $\leq 2n-1$, consider $n = 4$ and $f(x) = x^2$. Then

$$T_4(x) = 8x^4 - 8x^2 + 1$$

has zeros $x_1, \ldots, x_4$ with

$$x_1^2 = x_4^2 = \frac{2+\sqrt{2}}{4}, \quad x_2^2 = x_3^2 = \frac{2-\sqrt{2}}{4}.$$

Hence

$$\int_{-1}^{1} \frac{x^2}{\sqrt{1-x^2}}\, \mathrm{d}x \simeq \frac{\pi}{4} \sum_k x_k^2 = \frac{\pi}{4}\, 2 \left(\frac{2+\sqrt{2}}{4} + \frac{2-\sqrt{2}}{4} \right) = \frac{\pi}{2}$$

which is the exact value of the integral, as we expect. (See Problem 6 for a more challenging example.)

Cases 3 and 4 of Theorem 8.4, namely the Chebyshev polynomials of the third and fourth kinds, require a little more care. We first establish a lemma corresponding to Lemma 8.3.

Lemma 8.5

1.

$$\int_0^{\pi} \frac{\cos(n+\frac{1}{2})\theta}{\cos\theta - \cos\phi} \cos\tfrac{1}{2}\theta\, \mathrm{d}\theta = \frac{\pi}{2} \frac{\sin(n+\frac{1}{2})\phi}{\sin\frac{1}{2}\phi}.$$

2.

$$\int_0^{\pi} \frac{\sin(n+\frac{1}{2})\theta}{\cos\theta - \cos\phi} \sin\tfrac{1}{2}\theta\, \mathrm{d}\theta = -\frac{\pi}{2} \frac{\cos(n+\frac{1}{2})\phi}{\sin\frac{1}{2}\phi}.$$

Proof: (of the Lemma) From the first equation of Lemma 8.3, if we replace $\cos\theta$ by x and $\cos\phi$ by y,

$$\int_{-1}^{1} \frac{T_n(x)}{x-y} \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \pi U_{n-1}(y). \tag{8.15}$$

Writing $x = 2u^2 - 1$, $y = 2v^2 - 1$, where $u = \cos\frac{1}{2}\theta$, $v = \cos\frac{1}{2}\phi$,

$$\begin{aligned}
\int_{-1}^{1} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \frac{V_n(x)}{x-y}\,\mathrm{d}x &= \int_{0}^{1} \frac{2u}{\sqrt{1-u^2}} \frac{T_{2n+1}(u)}{u^2-v^2}\,\mathrm{d}u \\
&= \tfrac{1}{2}\int_{-1}^{1} T_{2n+1}(u)\left(\frac{1}{u+v}+\frac{1}{u-v}\right)\frac{\mathrm{d}u}{\sqrt{1-u^2}} \\
&= \int_{-1}^{1} \frac{T_{2n+1}(u)}{u-v}\frac{\mathrm{d}u}{\sqrt{1-u^2}} \\
&= \pi U_{2n}(v), \qquad \text{by (8.15)}.
\end{aligned}$$

Rewriting this in terms of θ and ϕ, we get

$$\int_{0}^{\pi} \frac{1}{\sin\frac{1}{2}\theta} \frac{\cos(n+\frac{1}{2})\theta}{\cos\theta - \cos\phi} \sin\theta\,\mathrm{d}\theta = \pi\frac{\sin(2n+1)\frac{1}{2}\phi}{\sin\frac{1}{2}\phi}, \tag{8.16}$$

and this proves part 1 of the Lemma.

Part 2 may be proved similarly, starting from the second equation of Lemma 8.3, which gives

$$\int_{-1}^{1} (1-x^2)^{\frac{1}{2}} \frac{U_{n-1}(x)}{x-y}\,\mathrm{d}x = \pi T_n(y),$$

and making similar substitutions. ●●

Proof: (of Theorem 8.4, case 3) Here

$$\begin{aligned}
A_k &= \int_{-1}^{1} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \prod_{r\neq k}\left(\frac{x-x_r}{x_k-x_r}\right)\mathrm{d}x \\
&= \int_{-1}^{1} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \frac{V_n(x)}{(x-x_k)V_n'(x_k)} \\
&= \int_{0}^{\pi} \frac{1}{\sin\frac{1}{2}\theta} \frac{\cos(n+\frac{1}{2})\theta\ \cos\frac{1}{2}\theta_k\ \sin\theta_k\ \sin\theta}{(\cos\theta-\cos\theta_k)\ (n+\frac{1}{2})\ \sin(n+\frac{1}{2})\theta_k}\,\mathrm{d}\theta \\
&= \frac{2\pi}{n+\frac{1}{2}}\cos^2\tfrac{1}{2}\theta_k, \qquad \text{by (8.16)} \\
&= \frac{\pi}{n+\frac{1}{2}}(1+x_k).
\end{aligned}$$

Thus case 3 is proved. Case 4 follows, on replacing x by $-x$. ●●

EXAMPLE 8.2: To illustrate this case, consider, for example, $f(x) = x^2$ and $n = 2$ for case 3, so that

$$I = \int_{-1}^{1} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} x^2 \, \mathrm{d}x.$$

Now $V_2(x) = 4x^2 - 2x - 1$ has zeros $x_1, x_2 = \frac{1}{4}(1 \pm \sqrt{5})$, with $x_1^2, x_2^2 = \frac{1}{8}(3 \pm \sqrt{5})$. Hence

$$\begin{aligned} I &\simeq \frac{2\pi}{5}[(1+x_1)x_1^2 + (1+x_2)x_2^2] \\ &= \frac{2\pi}{5}[\tfrac{1}{4}(5+\sqrt{5})\tfrac{1}{8}(3+\sqrt{5}) + \tfrac{1}{4}(5-\sqrt{5})\tfrac{1}{8}(3-\sqrt{5})] \\ &= \tfrac{1}{2}\pi. \end{aligned}$$

This is exact, as we can verify:

$$I = \int_0^{\pi} \frac{\cos\frac{1}{2}\theta}{\sin\frac{1}{2}\theta} (\cos\theta)^2 \, \sin\theta \, \mathrm{d}\theta = \int_0^{\pi} \tfrac{1}{2}(1+\cos\theta)(1+\cos 2\theta) \, \mathrm{d}\theta = \tfrac{1}{2}\pi.$$

The Gauss–Chebyshev quadrature formulae are the only Gauss formulae whose nodes x_k and weights A_k (given by Theorem 8.4) can be written down explicitly.

8.3 Quadrature methods of Clenshaw–Curtis type

8.3.1 Introduction

The Gauss–Chebyshev quadrature method of Section 8.2 is based on the continuous orthogonality properties of the Chebyshev polynomials. However, as we showed in Section 4.6, the four kinds of polynomials also have discrete orthogonality properties, and it is this kind of property that was exploited in the original quadrature method of Clenshaw & Curtis (1960). Their method has been developed in a considerable literature of papers by many authors (Piessens & Branders 1983, Adam 1987, Adam & Nobile 1991); a particularly nice presentation is given by Sloan & Smith (1978), who provide a version based on a general weight function together with a calculation of error estimates. Our treatment here is based on Sloan and Smith's formulation and techniques, which we can extend to all four kinds of Chebyshev polynomials.

The basic idea is to replace the integrand by an interpolating polynomial, and then to integrate this between the required limits. Suppose that we wish to determine the integral

$$I(f) := \int_{-1}^{1} w(x)f(x) \, \mathrm{d}x; \qquad (8.17)$$

then we replace $f(x)$ by the polynomial $J_n f(x)$ of degree n which interpolates f in abscissae $\{x_k : k = 1, \ldots, n+1\}$, and hence we obtain the approximation

$$I_n(f) := \int_{-1}^{1} w(x) J_n f(x) \, \mathrm{d}x \tag{8.18}$$

to evaluate, either exactly or approximately. So far, this only repeats what we have said earlier. However, if Chebyshev polynomial abscissae are adopted as interpolation points then, as we saw in Section 6.3, discrete orthogonality properties lead to very economical interpolation formulae, expressing the polynomial $J_n f(x)$ in forms which can readily be integrated — in many cases exactly.

There are a few important cases in which Gauss–Chebyshev and Clenshaw–Curtis quadrature lead to the same formulae, although they differ in general.

8.3.2 First-kind formulae

Suppose that

$$J_n f(x) = \sum_{j=0}^{n} b_j T_j(x) \tag{8.19}$$

interpolates $f(x)$ in the zeros $\{x_k\}$ of $T_{n+1}(x)$. Then, using the discrete orthogonality results (4.40) and (4.42), we have

$$d_{ij} := \sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) = 0, \quad i \neq j, \quad i, j \leq n \tag{8.20a}$$

and

$$d_{ii} = \begin{cases} (n+1), & i = 0, \\ \frac{1}{2}(n+1), & i \neq 0. \end{cases} \tag{8.20b}$$

Hence

$$\sum_{k=1}^{n+1} f(x_k) T_i(x_k) = \sum_{k=1}^{n+1} J_n f(x_k) T_i(x_k) = \sum_{j=0}^{n} b_j \sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) = b_i d_{ii}$$

and so

$$b_i = \frac{1}{d_{ii}} \sum_{k=1}^{n+1} f(x_k) T_i(x_k). \tag{8.21}$$

From (8.18)

$$I_n(f) = \sum_{j=0}^{n} b_j a_j, \tag{8.22}$$

where

$$a_j = \int_{-1}^{1} w(x)T_j(x)\,\mathrm{d}x = \int_0^\pi w(\cos\theta)\ \cos j\theta\ \sin\theta\,\mathrm{d}\theta. \tag{8.23}$$

Formulae (8.21)–(8.23) give the quadrature rule

$$I_n(f) = \sum_{k=1}^{n+1} w_k f(x_k), \tag{8.24a}$$

$$w_k = \sum_{j=0}^{n} \frac{a_j}{d_{jj}} T_j(x_k) = \sum_{j=0}^{n}{}' \frac{2a_j}{n+1} T_j(x_k). \tag{8.24b}$$

Hence I_n is readily determined, provided that the integrals (8.23) defining a_j are straightforward to calculate.

- For the specific weighting

  $$w(x) = (1-x^2)^{-\frac{1}{2}} \tag{8.25}$$

  we have

  $$a_j = \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_j(x)\,\mathrm{d}x = \int_0^\pi \cos j\theta\,\mathrm{d}\theta = \begin{cases} \pi, & j=0, \\ 0, & j>0, \end{cases} \tag{8.26}$$

  giving

  $$w_k = \frac{a_0}{d_{00}} T_0(x_k) = \frac{\pi}{n+1}.$$

  Hence

  $$\int_{-1}^{1} f(x) \frac{\mathrm{d}x}{\sqrt{1-x^2}} \simeq I_n(f) = \frac{\pi}{n+1} \sum_{k=1}^{n+1} f(x_k). \tag{8.27}$$

  Thus we get the first-kind Gauss–Chebyshev formula of Theorem 8.4.

An alternative Clenshaw–Curtis formula may be obtained by defining $J_n f(x)$ to be the polynomial interpolating the values of $f(x)$ at the abscissae

$$y_k = \cos\frac{k\pi}{n}, \quad k = 0, \ldots, n,$$

which are the zeros of $(1-x^2)U_{n-1}(x)$. In this case we use the discrete orthogonality results (4.45) and (4.46) to give us

$$d_{ij} := \sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k) = 0, \quad i \neq j \tag{8.28a}$$

and

$$d_{ii} = \begin{cases} n, & i=0, i=n, \\ \frac{1}{2}n, & 0<i<n. \end{cases} \tag{8.28b}$$

We readily deduce, in place of (8.19), that

$$J_n f(x) = \sum_{j=0}^{n} b_j T_j(x) \tag{8.29}$$

where in this case

$$b_i = \frac{1}{d_{ii}} \sum_{k=0}^{n}{}'' f(y_k) T_i(y_k), \tag{8.30}$$

and that

$$I_n(f) = \sum_{j=0}^{n} b_j a_j$$

where a_j are given by the same formula (8.23) as before. This gives us the rule

$$I_n(f) = \sum_{k=0}^{n} w_k f(y_k) \tag{8.31a}$$

$$w_k = \sum_{j=0}^{n} \frac{a_j}{d_{jj}} T_j(y_k) = \sum_{j=0}^{n}{}'' \frac{2a_j}{n} T_j(y_k). \tag{8.31b}$$

- For $w(x) = (1 - x^2)^{-\frac{1}{2}}$, this reduces to the formula

$$\int_{-1}^{1} f(x) \frac{\mathrm{d}x}{\sqrt{1-x^2}} \simeq I_n(f) = \pi b_0 = \frac{\pi}{n} \sum_{j=0}^{n}{}'' f(y_k). \tag{8.32}$$

  This is nearly equivalent to the second-kind Gauss–Chebyshev formula of Theorem 8.4, applied to the function $\dfrac{f(x)}{1-x^2}$, except that account is taken of the values of $f(x)$ at the end points $x = \pm 1$. This may better reflect the inverse-square-root singularities of the integrand at these points.

8.3.3 Second-kind formulae

It is clear that the key to the development of a Clenshaw–Curtis integration method is the finding of a discrete orthogonality formula. In fact, there exist at least sixteen such formulae, listed in Problem 14 of Chapter 4, some of which are covered in Section 4.6.

An example of a second-kind discrete orthogonality formula, given by (4.50) and (4.51), is

$$d_{ij} = \sum_{k=1}^{n+1} (1 - y_k^2) U_i(y_k) U_j(y_k) = \begin{cases} \frac{1}{2}(n+2), & i = j \leq n, \\ 0, & i \neq j, \end{cases} \tag{8.33}$$

where $\{y_k\}$ are the zeros of $U_{n+1}(x)$:

$$y_k = \cos\frac{k\pi}{n+2}, \quad k = 1, \ldots, n+1.$$

To make use of this, we again approximate the required integral $I(f)$ of (8.17) by the integral $I_n(f)$ of the form (8.18), but now interpolating $f(x)$ by a function of the form

$$J_n f(x) = (1-x^2)^{\frac{1}{2}} \sum_{j=0}^{n} b_j U_j(x); \tag{8.34}$$

that is, a polynomial weighted by $(1-x^2)^{\frac{1}{2}}$. There is thus an implicit assumption that $f(x)$ vanishes at $x = \pm 1$, and that it possibly has a square-root singularity at these points (though this is not essential).

Now

$$b_i = \frac{2}{n+2} \sum_{k=1}^{n+1} (1-y_k^2)^{\frac{1}{2}} f(y_k) U_i(y_k) \tag{8.35}$$

from (8.33). Integrating (8.18) gives us

$$I_n(f) = \sum_{j=0}^{n} b_j a_j \tag{8.36}$$

where

$$a_j = \int_{-1}^{1} w(x)(1-x^2)^{\frac{1}{2}} U_j(x)\,\mathrm{d}x = \int_0^{\pi} w(\cos\theta)\ \sin(j+1)\theta\ \sin\theta\,\mathrm{d}\theta. \tag{8.37}$$

This gives the rule

$$I_n(f) = \sum_{k=1}^{n+1} w_k f(y_k), \tag{8.38a}$$

$$w_k = (1-y_k^2)^{\frac{1}{2}} \sum_{j=0}^{n} \frac{2a_j}{n+2} U_j(y_k). \tag{8.38b}$$

- In the special case where $w(x) = 1$,

$$a_j = \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} U_j(x)\,\mathrm{d}x = \int_0^{\pi} \sin(j+1)\theta\ \sin\theta\,\mathrm{d}\theta = \begin{cases} \frac{1}{2}\pi, & j = 0, \\ 0, & j > 0. \end{cases}$$

  Hence, from (8.36), (8.37),

$$\int_{-1}^{1} f(x)\,\mathrm{d}x = I_n(f) = \frac{\pi}{n+2} \sum_{k=1}^{n+1} (1-y_k^2)^{\frac{1}{2}} f(y_k). \tag{8.39}$$

This is equivalent to the second-kind Gauss–Chebyshev formula of Theorem 8.4, applied to the function

$$\frac{f(x)}{\sqrt{1-x^2}}.$$

8.3.4 Third-kind formulae

A third-kind formula is obtained from the orthogonality formula

$$d_{ij} = \sum_{k=1}^{n+1}(1+x_k)V_i(x_k)V_j(x_k) = \begin{cases} n+\frac{3}{2}, & i=j, \\ 0, & i \neq j, \end{cases} \tag{8.40}$$

where $\{x_k\}$ are the zeros of $V_{n+1}(x)$. (See Problem 14 of Chapter 4.)

In this case, we choose

$$J_n f(x) = (1+x)^{\frac{1}{2}} \sum_{j=0}^{n} b_j V_j(x), \tag{8.41}$$

a polynomial weighted by $(1+x)^{\frac{1}{2}}$ (implicitly supposing that $f(-1) = 0$). Now, from (8.39), we can show that

$$b_i = \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} (1+x_k)^{\frac{1}{2}} f(x_k) V_i(x_k). \tag{8.42}$$

Integrating (8.18) gives us again

$$I_n(f) = \sum_{j=0}^{n} b_j a_j$$

where now

$$a_j = \int_{-1}^{1} w(x)(1+x)^{\frac{1}{2}} V_j(x)\,\mathrm{d}x. \tag{8.43}$$

So we have the rule

$$I_n(f) = \sum_{k=1}^{n+1} w_k f(x_k) \tag{8.44a}$$

$$w_k = (1+x_k)^{\frac{1}{2}} \sum_{j=0}^{n} \frac{2a_j}{2n+3} V_j(x_k). \tag{8.44b}$$

- For the special case in which

$$w(x) = (1-x)^{-\frac{1}{2}}, \tag{8.45}$$

then

$$a_j = \int_{-1}^{1} (1+x) V_j(x) \frac{\mathrm{d}x}{(1-x^2)^{\frac{1}{2}}} =$$
$$= \int_0^{\pi} 2\cos(j+\tfrac{1}{2})\theta \, \cos\tfrac{1}{2}\theta \, \mathrm{d}\theta = \begin{cases} \pi, & j = 0, \\ 0, & j > 0. \end{cases}$$

Hence

$$\int_{-1}^{1} f(x) \frac{\mathrm{d}x}{\sqrt{1-x}} = I_n(f) = \frac{2\pi}{2n+3} \sum_{k=1}^{n+1} (1+x_k)^{\frac{1}{2}} f(x_k). \tag{8.46}$$

This is equivalent to the third-kind Gauss–Chebyshev formula of Theorem 8.4, applied to the function

$$\frac{f(x)}{\sqrt{1+x}}.$$

8.3.5 General remark on methods of Clenshaw–Curtis type

There are effectively two types of quadrature formula considered above.

- For special choices of weight function $w(x)$, such that all but one of the Chebyshev transforms b_i vanish, the formula involves only a single summation — such as (8.27) — and is identical or very similar to a Gauss–Chebyshev formula.

- For a more general weight function, provided that the integral (8.23), (8.37) or (8.43) defining a_j can be exactly evaluated by some means, we obtain a formula involving a double summation — such as (8.24) — one set of summations to compute the weights w_k and a final summation to evaluate the integral.

8.4 Error estimation for Clenshaw–Curtis methods

There are a number of papers on error estimation in Clenshaw–Curtis methods (Fraser & Wilson 1966, O'Hara & Smith 1968, Smith 1982, Favati et al. 1993, for instance). However, we emphasise here the approach of Sloan & Smith (1980), which seems to be particularly robust, depends on interesting properties of Chebyshev polynomials, and is readily extendible to cover all four kinds of Chebyshev polynomial and the plethora of abscissae that were discussed in Section 8.3.

8.4.1 First-kind polynomials

Suppose that the function $f(x)$ being approximated is continuous and of bounded variation, and therefore has a uniformly convergent first-kind Chebyshev expansion

$$f(x) \approx \sum_{j=0}^{\infty}{}' \beta_j T_j(x). \tag{8.47}$$

Then the error in the integration method (8.31) (based on $\{y_k\}$) is

$$\begin{aligned} E_n(f) &:= I(f) - I_n(f) \\ &= (I - I_n)\left(\sum_{j=n+1}^{\infty} \beta_j T_j(x)\right) \\ &= \sum_{j=n+1}^{\infty} \beta_j \{I(T_j) - I_n(T_j)\}. \end{aligned} \tag{8.48}$$

Now

$$I(T_j) = \int_{-1}^{1} w(x) T_j(x)\,\mathrm{d}x = a_j \tag{8.49}$$

and (J_n again denoting the operator interpolating in the points $\{y_k\}$)

$$I_n(T_j) = \int_{-1}^{1} w(x) J_n T_j(x)\,\mathrm{d}x. \tag{8.50}$$

But

$$J_n T_j(y_k) = T_j(y_k) = T_{j'}(y_k) \tag{8.51}$$

where (as shown in Table 8.1) $j' = j'(n, j)$ is an integer in the range $0 \le j' \le n$ defined by

$$\left.\begin{array}{rcll} j'(n,j) &=& j, & 0 \le j \le n \\ j'(n,j) &=& 2n - j, & n \le j \le 2n \\ j'(n, 2n+j) &=& j'(n,j) & \end{array}\right\}. \tag{8.52}$$

This follows immediately from the observation that, j, k and n being integers,

$$T_j(y_k) = \cos\frac{jk\pi}{n} = \cos\frac{(2n \pm j)k\pi}{n} = T_{2n\pm j}(y_k).$$

Thus the interpolation operator J_n has the so-called *aliasing*[1] effect of identifying any Chebyshev polynomial T_j with a polynomial $T_{j'}$ of degree at most n, and it follows from (8.51) that, identically,

$$J_n T_j(x) = T_{j'}(x), \tag{8.53}$$

[1] See Section 6.3.1.

Table 8.1: $T_{j'}(x)$ interpolates $T_j(x)$ in the zeros of $(1-x^2)U_{n-1}(x)$

$j =$	0	1	2	$\rightarrow$	$n-1$	n
	$2n$	$2n-1$	$2n-2$	$\leftarrow$	$n+1$	n
	$2n$	$2n+1$	$2n+2$	$\rightarrow$	$3n-1$	$3n$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$j' =$	0	1	2	$\cdots$	$n-1$	n

and

$$I_n(T_j) = I_n(T_{j'}) = I(T_{j'}) = a_{j'}. \tag{8.54}$$

Therefore

$$E_n(f) = \sum_{j=n+1}^{\infty} \beta_j (a_j - a_{j'}). \tag{8.55}$$

Sloan & Smith (1980) assume that the weight function $w(x)$ is smooth enough for a_j (8.23) to be neglected for $j > 2n$ and that the integrand $f(x)$ itself is smooth enough for β_j (8.47) to be neglected beyond $j = 3n$. Then (8.55) yields, referring to Table 8.1,

$$\begin{aligned} |E_n(f)| \le\ & |a_{n+1} - a_{n-1}|\,|\beta_{n+1}| + |a_{n+2} - a_{n-2}|\,|\beta_{n+2}| + \cdots \\ & \cdots + |a_{2n} - a_0|\,|\beta_{2n}| + |a_1|\,|\beta_{2n+1}| + \cdots + |a_n|\,|\beta_{3n}|\,. \end{aligned}$$

If we then assume a geometric decay in the β_js, say

$$|\beta_{n+j}| \le c_n r_n^j$$

for some c_n, r_n with $r_n < 1$, then

$$|E_n(f)| \le c_n\{|a_{n+1} - a_{n-1}|\,r_n + \cdots + |a_{2n} - a_0|\,r_n^n + |a_1|\,r_n^{n+1} + \cdots + |a_n|\,r_n^{2n}\}. \tag{8.56}$$

If we change the notation slightly, replacing b_j by b_{nj}, the additional subscript being introduced to show the dependence on n,

$$b_{nj} = \frac{2}{\pi} \sum_{k=0}^{n}{}'' f(y_k) T_j(y_k),$$

it is clear that b_{nj} is an approximation to

$$\beta_j = \frac{2}{\pi} \int_{-1}^{1} f(x) T_j(x) \frac{\mathrm{d}x}{\sqrt{1-x^2}},$$

which becomes increasingly accurate with increasing n. Hence, a succession of values of b_{nj} (for various values of n) may be used to estimate β_j. (For the case $j = n$, β_j would be approximated by $\frac{1}{2} b_{nj}$.)

Sloan and Smith's 'second method' is based on obtaining estimates of r_n and c_n, and then using them in (8.56). Essentially, r_n is estimated from ratios of coefficients and c_n from the coefficients themselves. One algorithm, which takes account of the observed fact that odd and even coefficients tend to have somewhat different behaviours, and which uses three or four coefficients to construct each estimate, is as follows:

- Compute

$$\begin{aligned} z_1 &= \max\{\tfrac{1}{2}\left|b_{nn}\right|, \left|b_{n,n-2}\right|, \left|b_{n,n-4}\right|, \left|b_{n,n-6}\right|\}, \\ z_2 &= \max\{\left|b_{n,n-1}\right|, \left|b_{n,n-3}\right|, \left|b_{n,n-5}\right|. \end{aligned}$$

- If $z_1 > z_2$ then if $\left|b_{n,n-6}\right| > \cdots > \frac{1}{2}\left|b_{nn}\right|$ then

$$r_n^2 = \max\left\{\frac{\frac{1}{2}\left|b_{nn}\right|}{\left|b_{n,n-2}\right|}, \frac{\left|b_{n,n-2}\right|}{\left|b_{n,n-4}\right|}, \frac{\left|b_{n,n-4}\right|}{\left|b_{n,n-6}\right|}\right\}, \tag{8.57}$$

  otherwise $r_n = 1$.

- If $z_1 < z_2$ then if $\left|b_{n,n-5}\right| > \cdots > \left|b_{n,n-1}\right|$ then

$$r_n^2 = \max\left\{\frac{\left|b_{n,n-1}\right|}{\left|b_{n,n-3}\right|}, \frac{\left|b_{n,n-3}\right|}{\left|b_{n,n-5}\right|}\right\}, \tag{8.58}$$

  otherwise $r_n = 1$.

- Set

$$c_n = \max\{\tfrac{1}{2}\left|b_{nn}\right|, \left|b_{n,n-1}\right| r_n, \ldots, \left|b_{n,n-6}\right| r_n^6\}. \tag{8.59}$$

8.4.2 Fitting an exponential curve

A similar but somewhat neater procedure for estimating c_n and r_n is to fit the coefficients

$$b_{nn}, b_{n,n-1}, b_{n,n-2}, \ldots, b_{n,n-k}$$

(or the even or odd subsequences of them) by the sequence

$$c_n r_n^n, c_n r_n^{n-1}, c_n r_n^{n-2}, \ldots, c_n r_n^{n-k}.$$

This is in effect a discrete approximation of a function $g(x) = b_{nx}$ by

$$c_n (r_n)^x \equiv \mathrm{e}^{A+Bx}$$

at $x = n, n-1, n-2, \ldots, n-k$, where $A = \ln c_n$ and $B = \ln r_n$.

Then

$$g(x) = \mathrm{e}^{A+Bx} + e(x)$$

where $e(x)$ is the error. Hence

$$\ln g(x) + \ln(1 - e(x)/g(x)) = A + Bx$$

so that, to the first order of approximation,

$$\ln g(x) - e(x)/g(x) \approx A + Bx$$

and

$$g(x) \ln g(x) - e(x) \approx g(x)(A + Bx).$$

Hence a discrete least-squares fit of $\ln g(x)$ by $A+Bx$, weighted throughout by $g(x)$, can be expected to give a good model of the least-squares fitting of $g(x)$ by e^{A+Bx}.

This is an example of an algorithm for approximation by a 'function of a linear form' — more general discussion of such algorithms is given in Mason & Upton (1989).

8.4.3 Other abscissae and polynomials

Analogous procedures to those of Section 8.4.1 can be found for all four kinds of Chebyshev polynomials, and for all sets of abscissae that provide discrete orthogonality.

For example:

- For first-kind polynomials on the zeros $\{x_k\}$ of $T_{n+1}(x)$ (8.24), equations (8.47)–(8.50) still hold, but now

  $$J_n T_j(x_k) = T_j(x_k) = \pm T_{j'}(x_k)$$

  where (as in Table 8.2)

  $$\left.\begin{array}{rcll} j'(n,j) & = & j, & 0 \le j \le n \text{ (with + sign)} \\ j'(n,n+1) & = & n+1 & \text{(with zero coefficient)} \\ j'(n,j) & = & 2n+2-j, & n+2 \le j \le 2n+2 \text{ ($-$ sign)} \\ j'(n,j+2n+2) & = & j'(n,j) & \text{(with changed sign)} \end{array}\right\}. \tag{8.60}$$

  This follows immediately from

  $$T_j(x_k) = \begin{cases} \cos\frac{j'(k-\frac{1}{2})\pi}{n+1} & (0 \le j \le n) \\ 0 & (j = n+1) \\ \cos\frac{(2n+2-j')(k-\frac{1}{2})\pi}{n+1} = -\cos\frac{j'(k-\frac{1}{2})\pi}{n+1} & (n+2 \le j \le 2n+2) \\ \cos\frac{(2n+2+j')(k-\frac{1}{2})\pi}{n+1} = -\cos\frac{j'(k-\frac{1}{2})\pi}{n+1} & (2n+3 \le j \le 3n+2) \end{cases}$$

Table 8.2: $\pm T_{j'}(x)$ interpolating $T_j(x)$ in the zeros of $T_{n+1}(x)$

$j =$	0	1	$\rightarrow$	n	$n+1$	$n+2$	$\rightarrow$	$2n+1$	$2n+2$
	$4n+4$	$4n+3$	$\leftarrow$	$3n+4$	$3n+3$	$3n+2$	$\leftarrow$	$2n+3$	$2n+2$
	$4n+4$	$4n+5$	$\rightarrow$	$5n+4$	$5n+5$	$5n+6$	$\rightarrow$	$6n+5$	$6n+6$
	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$j' =$	0	1	$\cdots$	n	$n+1$	n	$\cdots$	1	0
sign	$+$	$+$	$\cdots$	$+$	0	$-$	$\cdots$	$-$	$-$

We now deduce that

$$\begin{aligned} |E_n(f)| \le\ & |a_{n+1}|\,|\beta_{n+1}| + |a_{n+2} + a_n|\,|\beta_{n+2}| + \cdots \\ & \cdots + |a_{2n+2} + a_0|\,|\beta_{2n+2}| + |a_1|\,|\beta_{2n+3}| + \cdots \\ & \cdots + |a_{n+1}|\,|\beta_{3n+3}|\,. \end{aligned} \tag{8.61}$$

- For second-kind polynomials on the zeros of U_{n+1} (8.38), we require an expansion

$$f(x) = \sum_{j=0}^{\infty} \beta_j U_j(x)$$

so that β_j is approximated by b_j from (8.35).

Then

$$E_n(f) = \sum_{j=n+1}^{\infty} \beta_j [I(U_j) - I_n(U_j)]$$

where now

$$I(U_j) = \int_{-1}^{1} w(x)(1-x^2)^{1/2} U_j(x)\,\mathrm{d}x = a_j \tag{8.62}$$

and

$$I_n(U_j) = \int_{-1}^{1} w(x)(1-x^2)^{1/2} J_n U_j(x)\,\mathrm{d}x. \tag{8.63}$$

If $\{y_k\}$ are the zeros of $U_{n+1}(x)$, then

$$J_n U_j(y_k) = U_j(y_k) = \pm U_{j'}(y_k)$$

where (taking $U_{-1} \equiv 0$)

$$\left.\begin{array}{rcll} j'(n,j) & = & j, & 0 \le j \le n \text{ (with + sign)} \\ j'(n,n+1) & = & n+1 & \text{(with zero coefficient)} \\ j'(n,j) & = & 2n+2-j, & n+2 \le j \le 2n+2 \text{ ($-$ sign)} \\ j'(n,2n+3) & = & -1 & \text{(with zero coefficient)} \\ j'(n,j+2n+4) & = & j'(n,j) & \text{(with unchanged sign)} \end{array}\right\}. \tag{8.64}$$

This is shown in Table 8.3, and follows from

$$y_k = \cos\theta_k = \cos\frac{k\pi}{n+2}, \quad k = 1, \ldots, n+1.$$

For

$$\begin{aligned} U_j(y_k)\sin\theta_k &= \sin(j+1)\theta_k \quad (j = 0, \ldots, n) \\ &= \sin(2n+2-j'+1)\theta_k = -\sin(j'+1)\theta_k \\ &= -U_{j'}(y_k)\sin\theta_k \quad (j' = n+1, \ldots) \end{aligned}$$

and

$$\sin(j+2n+4+1)\theta_k = \sin(j+1)\theta_k.$$

Table 8.3: $\pm U_{j'}(x)$ interpolating $U_j(x)$ in the zeros of $U_{n+1}(x)$

$j =$	0	$\to$	n	$n+1$	$n+2$	$\to$	$2n+2$	$2n+3$
	$2n+4$	$\to$	$3n+4$	$3n+5$	$3n+6$	$\to$	$4n+6$	$4n+7$
	$4n+8$	$\to$	$5n+8$	$5n+9$	$5n+10$	$\to$	$6n+10$	$6n+11$
	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$j' =$	0	$\cdots$	n	$n+1$	n	$\cdots$	0	-1
sign	$+$	$\cdots$	$+$	0	$-$	$\cdots$	$-$	0

From (8.62) and (8.63):

$$E_n(f) = \sum_{j=n+1}^{\infty} \beta_j(a_j - a_{j'})$$

and

$$\begin{aligned} |E_n(f)| \le\ & |a_{n+1}|\,|\beta_{n+1}| + |a_{n+2} + a_n|\,|\beta_{n+2}| + \cdots \\ & \cdots + |a_{2n+2} + a_0|\,|\beta_{2n+2}| + |a_0|\,|\beta_{2n+4}| + \cdots \\ & \cdots + |a_{n+1}|\,|\beta_{3n+5}|\,. \end{aligned} \tag{8.65}$$

- For third-kind polynomials on the zeros of V_{n+1} (8.44), we use an expansion

$$f(x) = (1+x)^{1/2} \sum_{j=0}^{\infty} \beta_j V_j(x). \tag{8.66}$$

Then

$$E_n(f) = \sum_{j=n+1}^{\infty} \beta_j [I(V_j) - I_n(V_j)]$$

where

$$I(V_j) = \int_{-1}^{1} w(x)(1+x)^{1/2} V_j(x)\,\mathrm{d}x = a_j \tag{8.67}$$

and

$$I_n(V_j) = \int_{-1}^{1} w(x)(1+x)^{1/2} J_n V_j(x)\,\mathrm{d}x. \tag{8.68}$$

Choose $\{x_k\}$ as the zeros of $V_{n+1}(x)$. Then

$$J_n V_j(x_k) = V_j(x_k) = \pm V_{j'}(x_k)$$

where

$$\left.\begin{array}{rcll} j'(n,j) & = & j, & 0 \le j \le n \text{ (with + sign)} \\ j'(n,n+1) & = & n+1 & \text{ (with zero coefficient)} \\ j'(n,j) & = & 2n+2-j, & n+2 \le j \le 2n+2 \text{ (} - \text{ sign)} \\ j'(n,j+2n+3) & = & j'(n,j) & \text{(with changed sign)} \end{array}\right\}. \tag{8.69}$$

This is shown in Table 8.4, and follows from

$$x_k = \cos\theta_k = \cos\frac{(k-\frac{1}{2})\pi}{n+\frac{3}{2}},$$

giving

$$\begin{aligned} \cos\tfrac{1}{2}\theta_k V_j(x_k) &= \cos\frac{(j+\frac{1}{2})(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \\ &= \cos\frac{(2n+2-j'+\frac{1}{2})(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \\ &= \cos\frac{\{2(n+\frac{3}{2})-(j'+\frac{1}{2})\}(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \\ &= -\cos\frac{(j'+\frac{1}{2})(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \end{aligned}$$

and

$$\cos\frac{(j+2n+3+\frac{1}{2})(k-\frac{1}{2})\pi}{n+\frac{3}{2}} = -\cos\frac{(j+\frac{1}{2})(k-\frac{1}{2})\pi}{n+\frac{3}{2}}.$$

Table 8.4: $\pm V_{j'}(x)$ interpolating $V_j(x)$ in the zeros of $V_{n+1}(x)$

$j =$	0	$\rightarrow$	n	$n+1$	$n+2$	$\rightarrow$	$2n+2$
	$4n+5$	$\leftarrow$	$3n+5$	$3n+4$	$3n+3$	$\leftarrow$	$2n+3$
	$4n+6$	$\rightarrow$	$5n+6$	$5n+7$	$5n+8$	$\rightarrow$	$6n+8$
	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
$j' =$	0	$\cdots$	n	$n+1$	n	$\cdots$	0
sign	$+$	$\cdots$	$+$	0	$-$	$\cdots$	$-$

From (8.67) and (8.68):

$$E_n(f) = \sum_{j=n+1}^{\infty} \beta_j (a_j - a_{j'})$$

and

$$\begin{aligned} |E_n(f)| \le\ & |a_{n+1}|\,|\beta_{n+1}| + |a_{n+2} + a_n|\,|\beta_{n+2}| + \cdots \\ & \cdots + |a_{2n+2} + a_0|\,|\beta_{2n}| + |a_0|\,|\beta_{2n+3}| + \cdots \\ & \cdots + |a_n|\,|\beta_{3n+3}|\,. \end{aligned} \tag{8.70}$$

We note that there are only very slight differences between Tables 8.2, 8.3 and 8.4 and between the corresponding error bounds (8.61), (8.65) and (8.70).

8.5 Some other work on Clenshaw–Curtis methods

There is now a significant amount of literature on Clenshaw–Curtis methods, built up over about forty years, from which we shall draw attention to a selection of items.

Of particular interest are applications to Bessel function integrals (Piessens & Branders 1983), oscillatory integrals (Adam 1987), Fourier transforms of singular functions (Piessens & Branders 1992), Cauchy principal-value integrals (Hasegawa & Torii 1991) and Volterra integral equations (Evans et al. 1981).

Among contributions specific to error bounds and error estimates are the early work of Chawla (1968), Locher (1969) and O'Hara & Smith (1968), together with more recent work of Smith (1982) and Favati et al. (1993)—the last being concerned with analytic functions.

Product integration (including error estimation) has been well studied, in particular by Sloan & Smith (1978, 1980, 1982) and Smith & Paget (1992).

There has been an important extension of the Clenshaw–Curtis method to integration over a d-dimensional hypercube, by Novak & Ritter (1996).

8.6 Problems for Chapter 8

1. If $w = (1 - x^2)^{-\frac{1}{2}}$ and $P_k(x) = T_k(x)$ in Section 8.1, show that

$$\begin{aligned}\|h - h_n\|_\infty &= \int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} \left| f(x) - \sum_{k=0}^{n}{}' a_k T_k(x) \right| \mathrm{d}x \\ &= \int_{0}^{\pi} \left| f(\cos\theta) - \sum_{k=0}^{n}{}' a_k \cos k\theta \right| \mathrm{d}\theta.\end{aligned}$$

   By considering the Fourier cosine series expansion of $f(\cos\theta)$, deduce Theorem 8.1 for the first case.

2. If $w = [\frac{1}{2}(1 - x)]^{-\frac{1}{2}}$ and $P_k(x) = V_k(x)$ in Section 8.1, show that

$$\begin{aligned}\|h - h_n\|_\infty &= \int_{-1}^{1} [\tfrac{1}{2}(1 - x)]^{-\frac{1}{2}} \left| f(x) - \sum_{k=0}^{n} a_k V_k(x) \right| \mathrm{d}x \\ &= 2\int_{0}^{\pi} \left| \cos\tfrac{1}{2}\theta\, f(\cos\theta) - \sum_{k=0}^{n} a_k \cos(k + \tfrac{1}{2})\theta \right| \mathrm{d}\theta \\ &= 4\int_{0}^{\pi/2} \left| \cos\phi\, f(\cos 2\phi) - \sum_{k=0}^{n} a_k \cos(2k + 1)\phi \right| \mathrm{d}\phi.\end{aligned}$$

   By considering the Fourier cosine series expansion of $\cos\phi\, f(\cos 2\phi)$ (which is odd about $\phi = \frac{1}{2}\pi$), deduce Theorem 8.1 for the third case.

3. Complete the proof of Lemma 8.3, by performing an induction on n for the pair of formulae together.

4. Use Lemma 8.3 to prove the second part of Theorem 8.4. Verify that this quadrature formula is exact for $n = 3$ in the case of the integral

$$\int_{-1}^{1} \sqrt{1 - x^2}\, x^2 \,\mathrm{d}x.$$

5. Prove in detail the second part of Lemma 8.5.

6. Verify the exactness of Gauss–Chebyshev quadrature using first-kind polynomials, by testing it for $n = 4$ and $f(x) = x^6$, $f(x) = x^7$.

7. Verify the Gauss–Chebyshev rule for fourth-kind polynomials, by testing it for $n = 1$ and $f(x) = 1$, $f(x) = x$.

8. Verify that there is a Gauss–Chebyshev quadrature rule based on the zeros of $(1 - x^2)U_{n-1}(x)$ and the polynomials $T_n(x)$, and derive a formula. (This type of formula, which uses both end points, is called a *Lobatto* rule.) When would this rule be useful?

9. Show that there is a Gauss–Chebyshev quadrature rule based on the zeros of $(1+x)V_n(x)$ and the polynomials $T_n(x)$, and derive a formula. (This type of formula, which uses one end point, is called a *Radau* rule.) When would this rule be useful?

CHAPTER 9

Solution of Integral Equations

9.1 Introduction

In this chapter we shall discuss the application of Chebyshev polynomial techniques to the solution of Fredholm (linear) integral equations, which are classified into three kinds taking the following generic forms:

First kind: Given functions $K(x, y)$ and $g(x)$, find a function $f(y)$ on $[a, b]$ such that for all $x \in [a, b]$

$$\int_a^b K(x, y) f(y) \, \mathrm{d}y = g(x); \tag{9.1}$$

Second kind: Given functions $K(x, y)$ and $g(x)$, and a constant λ not providing a solution of (9.3) below, find a function $f(y)$ on $[a, b]$ such that for all $x \in [a, b]$

$$f(x) - \lambda \int_a^b K(x, y) f(y) \, \mathrm{d}y = g(x); \tag{9.2}$$

Third kind: Given a function $K(x, y)$, find values *(eigenvalues)* of the constant λ for which there exists a function *(eigenfunction)* $f(y)$, not vanishing identically on $[a, b]$, such that for all $x \in [a, b]$

$$f(x) - \lambda \int_a^b K(x, y) f(y) \, \mathrm{d}y = 0. \tag{9.3}$$

Equations of these three kinds may be written in more abstract terms as the functional equations

$$\mathcal{K} f = g, \tag{9.4}$$

$$f - \lambda \mathcal{K} f = g, \tag{9.5}$$

$$f - \lambda \mathcal{K} f = 0, \tag{9.6}$$

where $\mathcal{K}$ represents a linear mapping (here an *integral transformation*) from some function space F into itself or possibly (for an equation of the first kind) into another function space G, g represents a given element of F or G as appropriate and f is an element of F to be found.

A detailed account of the theory of integral equations is beyond the scope of this book — we refer the reader to Tricomi (1957), for instance. However, it is broadly true (for most of the kernel functions $K(x, y)$ that one is likely to meet) that equations of the second and third kinds have well-defined and well-behaved solutions. Equations of the first kind are quite another matter—here the problem will very often be ill-posed mathematically in the sense of either having no solution, having infinitely many solutions, or having a solution f that is infinitely sensitive to variations in the function g. It is essential that one reformulates such a problem as a well-posed one by some means, before attempting a numerical solution.

In passing, we should also mention integral equations of Volterra type, which are similar in form to Fredholm equations but with the additional property that $K(x, y) = 0$ for $y > x$, so that

$$\int_a^b K(x, y) f(y)\, \mathrm{d}y$$

is effectively

$$\int_a^x K(x, y) f(y)\, \mathrm{d}y.$$

A Volterra equation of the first kind may often be transformed into one of the second kind by differentiation. Thus

$$\int_a^x K(x, y) f(y)\, \mathrm{d}y = g(x)$$

becomes, on differentiating with respect to x,

$$K(x, x) f(x) + \int_a^x \frac{\partial}{\partial x} K(x, y) f(y)\, \mathrm{d}y = \frac{\mathrm{d}}{\mathrm{d}x} g(x).$$

It is therefore unlikely to suffer from the ill-posedness shown by general Fredholm equations of the first kind. We do not propose to discuss the solution of Volterra equations any further here.

9.2 Fredholm equations of the second kind

In a very early paper, Elliott (1961) studied the use of Chebyshev polynomials for solving non-singular equations of the second kind

$$f(x) - \lambda \int_a^b K(x, y) f(y)\, \mathrm{d}y = g(x), \quad a \le x \le b, \tag{9.7}$$

and this work was later updated by him (Elliott 1979). Here $K(x, y)$ is bounded in $a \le x, y \le b$, and we suppose that λ is not an eigenvalue of (9.3). (If λ were such an eigenvalue, corresponding to the eigenfunction $\phi(y)$, then any solution $f(y)$ of (9.7) would give rise to a multiplicity of solutions of the form $f(y) + \alpha\phi(y)$ where α is an arbitrary constant.)

For simplicity, suppose that $a = -1$ and $b = 1$. Assume that $f(x)$ may be approximated by a finite sum of the form

$$\sum_{j=0}^{N}{}'' a_j T_j(x). \tag{9.8}$$

Then we can substitute (9.8) into (9.7) so that the latter becomes the approximate equation

$$\sum_{j=0}^{N}{}'' a_j T_j(x) - \lambda \sum_{j=0}^{N}{}'' a_j \int_{-1}^{1} K(x,y) T_j(y)\, \mathrm{d}y \sim g(x), \quad -1 \le x \le 1. \tag{9.9}$$

We need to choose the coefficients a_j so that (9.9) is satisfied as well as possible over the interval $-1 \le x \le 1$. A reasonably good way of achieving this is by *collocation* — requiring equation (9.9) to be an exact equality at the $N + 1$ points (the extrema of $T_N(x)$ on the interval)

$$x = y_{i,N} = \cos \frac{i\pi}{N},$$

so that

$$\sum_{j=0}^{N}{}'' a_j (P_{ij} - \lambda Q_{ij}) = g(y_{i,N}), \quad i = 0, \ldots, N, \tag{9.10}$$

where

$$P_{ij} = T_j(y_{i,N}), \quad Q_{ij} = \int_{-1}^{1} K(y_{i,N}, y) T_j(y)\, \mathrm{d}y. \tag{9.11}$$

We thus have $N + 1$ linear equations to solve for the $N + 1$ unknowns a_j.

As an alternative to collocation, we may choose the coefficients $b_{i,k}$ so that $K_M(y_{i,N}, y)$ gives a least squares or minimax approximation to $K(y_{i,N}, y)$.

If we cannot evaluate the integrals in (9.11) exactly, we may do so approximately, for instance, by replacing each $K(y_{i,N}, y)$ with a polynomial

$$K_M(y_{i,N}, y) = \sum_{k=0}^{M}{}'' b_{i,k} T_k(y) \tag{9.12}$$

for some[1] $M > 0$, with Chebyshev coefficients given by

$$b_{i,k} = \frac{2}{M} \sum_{m=0}^{M}{}'' K(y_{i,N}, y_{m,M}) T_k(y_{m,M}), \quad k = 0, \ldots, M, \tag{9.13}$$

[1]There does not need to be any connection between the values of M and N.

where

$$y_{m,M} = \cos \frac{m\pi}{M}, \quad m = 0, \ldots, M.$$

As in the latter part of Section 6.3.2, we can then show that

$$K_M(y_{i,N}, y_{m,M}) = K(y_{i,N}, y_{m,M}), \quad m = 0, \ldots, M,$$

so that, for each i, $K_M(y_{i,N}, y)$ is the polynomial of degree M in y, interpolating $K(y_{i,N}, y)$ at the points $y_{m,M}$.

From (2.43) it is easily shown that

$$\int_{-1}^{1} T_n(x)\, \mathrm{d}x = \begin{cases} \dfrac{-2}{n^2-1}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases} \tag{9.14}$$

Hence

$$\begin{aligned} Q_{ij} &\approx \int_{-1}^{1} K_M(y_{i,N}, y) T_j(y)\, \mathrm{d}y \\ &= \sum_{k=0}^{M}{}'' b_{i,k} \int_{-1}^{1} T_k(y) T_j(y)\, \mathrm{d}y \\ &= \tfrac{1}{2} \sum_{k=0}^{M}{}'' b_{i,k} \int_{-1}^{1} \{T_{j+k}(y) + T_{|j-k|}(y)\}\, \mathrm{d}y \\ &= - \sum_{\substack{k=0 \\ j \pm k \text{ even}}}^{M}{}'' b_{i,k} \left\{ \frac{1}{(j+k)^2 - 1} + \frac{1}{(j-k)^2 - 1} \right\} \\ &= -2 \sum_{\substack{k=0 \\ j \pm k \text{ even}}}^{M}{}'' b_{i,k} \frac{j^2 + k^2 - 1}{(j^2 + k^2 - 1)^2 - 4j^2k^2}, \end{aligned} \tag{9.15}$$

giving us the approximate integrals we need.

Another interesting approach, based on 'alternating polynomials' (whose equal extrema occur among the given data points), is given by Brutman (1993). It leads to a solution in the form of a sum of Chebyshev polynomials, with error estimates.

9.3 Fredholm equations of the third kind

We can attack integral equations of the third kind in exactly the same way as equations of the second kind, with the difference that we have $g(x) = 0$.

Thus the linear equations (9.10) become

$$\sum_{j=0}^{N}{}'' a_j (P_{ij} - \lambda Q_{ij}) = 0, \quad i = 0, \ldots, N. \tag{9.16}$$

Multiplying each equation by $T_\ell(y_{i,N})$ and carrying out a $\sum''$ summation over i (halving the first and last terms), we obtain (after approximating K by K_M) the equations

$$\frac{MN}{4} a_\ell = \lambda \sum_{j=0}^{N}{}'' a_j \times$$
$$\times \left(\sum_{k=0}^{M}{}'' \sum_{i=0}^{N}{}'' \sum_{m=0}^{M}{}'' T_\ell(y_{i,N}) K(y_{i,N}, y_{m,M}) T_k(y_{m,M}) \int_{-1}^{1} T_k(y) T_j(y) \, \mathrm{d}y \right), \tag{9.17}$$

which is of the form

$$\frac{MN}{4} a_\ell = \lambda \sum_{j=0}^{N} a_j A_{\ell j} \tag{9.18}$$

or, written in terms of vectors and matrices,

$$\frac{MN}{4} \mathbf{a} = \lambda \mathbf{A} \mathbf{a}. \tag{9.19}$$

Once the elements of the $(N+1) \times (N+1)$ matrix $\mathbf{A}$ have been calculated, this is a straightforward (unsymmetric) matrix eigenvalue problem, which may be solved by standard techniques to give approximations to the dominant eigenvalues of the integral equation.

9.4 Fredholm equations of the first kind

Consider now a Fredholm integral equation of the first kind, of the form

$$g(x) = \int_a^b K(x, y) f(y) \, \mathrm{d}y, \quad c \le x \le d. \tag{9.20}$$

We can describe the function $g(x)$ as an integral transform of the function $f(y)$, and we are effectively trying to solve the 'inverse problem' of determining $f(y)$ given $g(x)$.

For certain special kernels $K(x, y)$, a great deal is known. In particular, the choices

$$K(x, y) = \cos xy, \quad K(x, y) = \sin xy \text{ and } K(x, y) = \mathrm{e}^{-xy},$$

with $[a, b] = [0, \infty)$, correspond respectively to the well-known Fourier cosine transform, Fourier sine transform and Laplace transform. We shall not pursue these topics specifically here, but refer the reader to the relevant literature (Erdélyi et al. 1954, for example).

Smooth kernels in general will often lead to inverse problems that are ill-posed in one way or another.

For example, if K is continuous and f is integrable, then it can be shown that $g = \mathcal{K}f$ must be continuous — consequently, if we are given a g that is not continuous then no (integrable) solution f of (9.20) exists.

Uniqueness is another important question to be considered. For example Groetsch (1984) notes that the equation

$$\int_0^\pi x \sin y\, f(y)\ \mathrm{d}y = x$$

has a solution

$$f(y) = \tfrac{1}{2}.$$

However, it has an infinity of further solutions, including

$$f(y) = \tfrac{1}{2} + \sin ny \quad (n = 2, 3, \ldots).$$

An example of a third kind of ill-posedness, given by Bennell (1996), is based on the fact that, if K is absolutely integrable in y for each x, then by the Riemann–Lebesgue theorem,

$$\phi_n(x) \equiv \int_a^b K(x, y) \cos ny\, \mathrm{d}y \to 0 \text{ as } n \to \infty.$$

Hence

$$\int_a^b K(x, y)(f(y) + \alpha \cos ny)\ \mathrm{d}y = g(x) + \alpha\phi_n(x) \to g(x) \text{ as } n \to \infty,$$

where α is an arbitrary positive constant. Thus a small perturbation

$$\delta g(x) = \alpha\phi_n(x)$$

in $g(x)$, converging to a zero limit as $n \to \infty$, can lead to a perturbation

$$\delta f(y) = \alpha \cos ny$$

in $f(y)$ which remains of finite magnitude α for all n. This means that the solution $f(y)$ does not depend continuously on the data $g(x)$, and so the problem is ill-posed.

We thus see that it is not in fact necessarily advantageous for the function K to be smooth. Nevertheless, there are ways of obtaining acceptable numerical solutions to problems such as (9.20). They are based on the technique of *regularisation*, which effectively forces an approximate solution to be appropriately smooth. We return to this topic in Section 9.6 below.

9.5 Singular kernels

A particularly important class of kernels, especially in the context of the study of Chebyshev polynomials in integral equations, comprises the *Hilbert kernel*

$$K(x,y) = \frac{1}{x-y} \tag{9.21}$$

and other related 'Hilbert-type' kernels that behave locally like (9.21) in the neighbourhood of $x = y$.

9.5.1 Hilbert-type kernels and related kernels

If $[a,b] = [-1,1]$ and

$$K(x,y) = \frac{w(y)}{y-x},$$

where $w(y)$ is one of the weight functions $(1+y)^\alpha(1-y)^\beta$ with $\alpha, \beta = \pm\frac{1}{2}$, then there are direct links of the form (9.20) between Chebyshev polynomials of the four kinds (Fromme & Golberg 1981, Mason 1993).

Theorem 9.1

$$\pi U_{n-1}(x) = \int_{-1}^{1} K_1(x,y) T_n(y)\,\mathrm{d}y, \tag{9.22a}$$

$$-\pi T_n(x) = \int_{-1}^{1} K_2(x,y) U_{n-1}(y)\,\mathrm{d}y, \tag{9.22b}$$

$$\pi W_n(x) = \int_{-1}^{1} K_3(x,y) V_n(y)\,\mathrm{d}y, \tag{9.22c}$$

$$-\pi V_n(x) = \int_{-1}^{1} K_4(x,y) W_n(y)\,\mathrm{d}y \tag{9.22d}$$

where

$$K_1(x,y) = \frac{1}{\sqrt{1-y^2}\,(y-x)},$$

$$K_2(x,y) = \frac{\sqrt{1-y^2}}{(y-x)},$$

$$K_3(x,y) = \frac{\sqrt{1+y}}{\sqrt{1-y}\,(y-x)},$$

$$K_4(x,y) = \frac{\sqrt{1-y}}{\sqrt{1+y}\,(y-x)},$$

and each integral is to be interpreted as a Cauchy principal value integral.

Proof: In fact, formulae (9.22a) and (9.22b) correspond under the transformation $x = \cos\theta$ to the trigonometric formulae

$$\int_0^\pi \frac{\cos n\phi}{\cos\phi - \cos\theta}\,\mathrm{d}\phi = \pi\frac{\sin n\theta}{\sin\theta}, \tag{9.23}$$

$$\int_0^\pi \frac{\sin n\phi \sin\phi}{\cos\phi - \cos\theta}\,\mathrm{d}\phi = -\pi\cos n\theta, \tag{9.24}$$

which have already been proved in another chapter (Lemma 8.3). Formulae (9.22c) and (9.22d) follow similarly from Lemma 8.5. ●●

From Theorem 9.1 we may immediately deduce integral relationships between Chebyshev series expansions of functions as follows.

Corollary 9.1A

1. *If* $f(y) \sim \sum_{n=1}^{\infty} a_n T_n(y)$ *and* $g(x) \sim \pi\sum_{n=1}^{\infty} a_n U_{n-1}(x)$ *then*

$$g(x) = \int_{-1}^{1} \frac{f(y)}{\sqrt{1-y^2}\,(y-x)}\,\mathrm{d}y. \tag{9.25a}$$

2. *If* $f(y) \sim \sum_{n=1}^{\infty} b_n U_{n-1}(y)$ *and* $g(x) \sim \pi\sum_{n=1}^{\infty} b_n T_n(x)$ *then*

$$g(x) = -\int_{-1}^{1} \frac{\sqrt{1-y^2}\,f(y)}{(y-x)}\,\mathrm{d}y. \tag{9.25b}$$

3. *If* $f(y) \sim \sum_{n=1}^{\infty} c_n V_n(y)$ *and* $g(x) \sim \pi\sum_{n=1}^{\infty} c_n W_n(x)$ *then*

$$g(x) = \int_{-1}^{1} \frac{\sqrt{1+y}\,f(y)}{\sqrt{1-y}\,(y-x)}\,\mathrm{d}y. \tag{9.25c}$$

4. *If* $f(y) \sim \sum_{n=1}^{\infty} d_n W_n(y)$ *and* $g(x) \sim \pi\sum_{n=1}^{\infty} d_n V_n(x)$ *then*

$$g(x) = -\int_{-1}^{1} \frac{\sqrt{1-y}\,f(y)}{\sqrt{1+y}\,(y-x)}\,\mathrm{d}y. \tag{9.25d}$$

Note that these expressions do not necessarily provide general solutions to the integral equations (9.25a)–(9.25d), but they simply show that the relevant formal expansions are integral transforms of each other.

These relationships are useful in attacking certain engineering problems. Gladwell & England (1977) use (9.25a) and (9.25b) in elasticity analysis and Fromme & Golberg (1979) use (9.25c), (9.25d) and related properties of V_n and W_n in analysis of the flow of air near the tip of an airfoil.

To proceed to other kernels, we note that by integrating equations (9.22a)–(9.22d) with respect to x, after premultiplying by the appropriate weights, we can deduce the following eigenfunction properties of Chebyshev polynomials for logarithmic kernels. The details are left to the reader (Problem 4).

Theorem 9.2 *The integral equation*

$$\lambda\phi(x) = \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}}\phi(y)K(x,y)\;\mathrm{d}y \tag{9.26}$$

has the following eigensolutions and eigenvalues λ for the following kernels K.

1. $K(x,y) = K_5(x,y) = \log|y-x|;$
   $\phi(x) = \phi_n(x) = T_n(x), \quad \lambda = \lambda_n = \pi/n.$

2. $K(x,y) = K_6(x,y) = \log|y-x| - \log\left|1 - xy - \sqrt{(1-x^2)(1-y^2)}\right|;$
   $\phi(x) = \phi_n(x) = \sqrt{1-x^2}\,U_{n-1}(x), \quad \lambda = \lambda_n = \pi/n.$

3. $K(x,y) = K_7(x,y) = \log|y-x| - \log\left|2 + x + y - 2\sqrt{1+x}\sqrt{1+y}\right|;$
   $\phi = \phi_n(x) = \sqrt{1+x}\,V_n(x), \quad \lambda = \lambda_n = \pi/(n+\frac{1}{2}).$

4. $K(x,y) = K_8(x,y) = \log|y-x| - \log\left|2 - x - y - 2\sqrt{1-x}\sqrt{1-y}\right|;$
   $\phi = \phi_n(x) = \sqrt{1-x}\,W_n(x), \quad \lambda = \lambda_n = \pi/(n+\frac{1}{2}).$

Note that each of these four kernels has a (weak) logarithmic singularity at $x = y$. In addition, K_6 has logarithmic singularities at $x = y = \pm 1$, K_7 at $x = y = -1$ and K_8 at $x = y = +1$.

From Theorem 9.2 we may immediately deduce relationships between formal Chebyshev series of the four kinds as follows.

Corollary 9.2A *With the notations of Theorem 9.2, in each of the four cases considered, if*

$$f(y) \sim \sum_{k=1}^{\infty} a_k\phi_k(y) \quad \textit{and} \quad g(x) \sim \sum_{k=1}^{\infty} \lambda_k a_k \phi_k(x)$$

then

$$g(x) = \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}} K(x,y)f(y)\,\mathrm{d}y.$$

Thus again four kinds of Chebyshev series may in principle be used to solve (9.26) for $K = K_5,\ K_6,\ K_7,\ K_8$, respectively.

The most useful results in Theorem 9.2 and its corollary are those relating to polynomials of the first kind, where we find from Theorem 9.2 that

$$-\frac{\pi}{n}T_n(x) = \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}} T_n(y)\log|y-x|\;\mathrm{d}y \tag{9.27}$$

and from Corollary 9.2A that, if

$$f(y) \sim \sum_{k=1}^{\infty} a_k T_k(y) \quad \text{and} \quad g(x) \sim \sum_{k=1}^{\infty} -\frac{\pi}{k} a_k T_k(x), \tag{9.28}$$

then

$$g(x) = \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}} \log|y-x|\, f(y)\, \mathrm{d}y. \tag{9.29}$$

Equation (9.29) is usually referred to as *Symm's integral equation*, and clearly a Chebyshev series method is potentially very useful for such problems. We shall discuss this specific problem further in Section 9.5.2.

By differentiating rather than integrating in (9.22a), (9.22b), (9.22c) and (9.22d), we may obtain the further results quoted in Problem 5. The second of these yields the simple equation

$$-n\pi U_{n-1}(x) = \fint_{-1}^{1} \frac{\sqrt{1-y^2}}{(y-x)^2} U_{n-1}(y)\, \mathrm{d}y. \tag{9.30}$$

This integral equation, which has a stronger singularity than (9.22a)–(9.22d), is commonly referred to as a *hypersingular* equation, in which the integral has to be evaluated as a Hadamard finite-part integral (Martin 1991, for example). A rather more general hypersingular integral equation is solved by a Chebyshev method, based on (9.30), in Section 9.7.1 below.

The ability of a Chebyshev series of the first or second kind to handle both Cauchy principal value and hypersingular integral transforms leads us to consider an integral equation that involves both. This can be successfully attacked, and Mason & Venturino (2002) give full details of a Galerkin method, together with both $\mathcal{L}_2$ and $\mathcal{L}_\infty$ error bounds, and convergence proofs.

9.5.2 Symm's integral equation

Consider the integral equation (Symm 1966)

$$G(x) = \mathcal{V}F(x) = \frac{1}{\pi} \int_{a}^{b} \log|y-x| F(y)\, \mathrm{d}y, \quad x \in [a,b], \tag{9.31}$$

which is of importance in potential theory.

This equation has a unique solution $F(y)$ (Jorgens 1970) with endpoint singularities of the form $(y-a)^{-\frac{1}{2}}(b-y)^{-\frac{1}{2}}$. In the case $a = -1$, $b = +1$, the required singularity is $(1-y^2)^{-\frac{1}{2}}$, and so we may write

$$F(y) = (1-y^2)^{-\frac{1}{2}} f(y), \quad G(x) = -\pi^{-1} g(x),$$

whereupon (9.31) becomes

$$g(x) = \mathcal{V}^* f(x) = \int_{-1}^{1} \frac{\log|y-x|}{\sqrt{1-y^2}} f(y)\,\mathrm{d}y \tag{9.32}$$

which is exactly the form (9.29) obtained from Corollary 9.2A.

We noted then (9.29) that if

$$f(y) \sim \sum_{k=1}^{\infty} a_k T_k(y)$$

then

$$g(x) \sim \sum_{k=1}^{\infty} -\frac{\pi}{k} a_k T_k(x)$$

(and vice versa).

Sloan & Stephan (1992), adopt such an idea and furthermore note that

$$\mathcal{V}^* T_0(x) = -\pi \log 2,$$

so that

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

if

$$g(x) \sim -\tfrac{1}{2} a_0 \pi \log 2 - \sum_{k=1}^{\infty} \frac{\pi}{k} a_k T_k(x).$$

Their method of approximate solution is to write

$$f^*(y) = \sum_{k=0}^{n-1}{}' a_k^* T_k(y) \tag{9.33}$$

and to require that

$$\mathcal{V}^* f^*(x) = g(x)$$

holds at the zeros $x = x_i$ of $T_n(x)$. Then

$$g(x_i) = -\tfrac{1}{2} a_0^* \pi \log 2 - \sum_{k=1}^{n-1} \frac{\pi}{k} a_k^* T_k(x_i), \quad i = 0, \ldots, n-1. \tag{9.34}$$

Using the discrete orthogonality formulae (4.42), we deduce that

$$a_0^* = -\frac{2}{n\pi \log 2} \sum_{i=0}^{n-1} g(x_i), \tag{9.35}$$

$$a_k^* = -\frac{2k}{n\pi} \sum_{i=0}^{n-1} g(x_i) T_k(x_i) \quad (k > 0). \tag{9.36}$$

Thus values of the coefficients $\{a_k^*\}$ are determined explicitly.

The convergence properties of the approximation f^* to f have been established by Sloan & Stephan (1992).

9.6 Regularisation of integral equations

Consider again an integral equation of the first kind, of the form

$$g(x) = \int_a^b K(x, y) f(y)\, \mathrm{d}y, \quad c \le x \le d, \tag{9.37}$$

i.e. $g = \mathcal{K} f$, where $\mathcal{K} : F \to G$ and where the given $g(x)$ may be affected by noise (Bennell & Mason 1989). Such a problem is said to be *well posed* if:

- for each $g \in G$ there exists a solution $f \in F$;
- this solution f is always unique in F;
- f depends continuously on g (i.e., the inverse of $\mathcal{K}$ is continuous).

Unfortunately it is relatively common for an equation of the form (9.37) to be ill posed, so that a method of solution is needed which ensures not only that a computed f is close to being a solution but also that f is an appropriately smooth function. The standard approach is called a *regularisation method*; Tikhonov (1963*b*, 1963*a*) proposed an L_2 approximation which minimises

$$I[f^*] := \int_a^b [\mathcal{K} f^*(x) - g(x)]^2\, \mathrm{d}x + \lambda \int_a^b [p(x) f^*(x)^2 + q(x) {f^*}'(x)^2]\, \mathrm{d}x \tag{9.38}$$

where p and q are specified positive weight functions and λ a positive 'smoothing' parameter. The value of λ controls the trade-off between the smoothness of f^* and the fidelity to the data g.

9.6.1 Discrete data with second derivative regularisation

We shall first make two changes to (9.38) on the practical assumptions that we seek a visually smooth (i.e., twice continuously differentiable) solution, and that the data are discrete. We therefore assume that $g(x)$ is known only at n ordinates x_i and then only subject to white noise contamination $\epsilon(x_i)$;

$$g^*(x_i) = \int_a^b K(x_i, y) f(y)\, \mathrm{d}y + \epsilon(x_i) \tag{9.39}$$

where each $\epsilon(x_i) \sim N(0, \sigma^2)$ is drawn from a normal distribution with zero mean and (unknown) variance σ^2. We then approximate f by the $f_\lambda^* \in L_2[a, b]$

that minimises

$$I[f^*] \equiv \frac{1}{n}\sum_{i=1}^{n}[\mathcal{K}f^*(x_i) - g^*(x_i)]^2 + \lambda\int_a^b \left[f^{*\prime\prime}(y)\right]^2 \,\mathrm{d}y, \tag{9.40}$$

thus replacing the first integral in (9.38) by a discrete sum and the second by one involving the second derivative of f^*.

Ideally, a value λ_{opt} of λ should be chosen (in an outer cycle of iteration) to minimise the true mean-square error

$$R(\lambda) \equiv \frac{1}{n}\sum_{i=1}^{n}[\mathcal{K}f^*_\lambda(x_i) - g(x_i)]^2. \tag{9.41}$$

This is not directly possible, since the values $g(x_i)$ are unknown. However, Wahba (1977) has shown that a good approximation to λ_{opt} may be obtained by choosing the 'generalised cross-validation' (GCV) estimate λ^*_{opt} that minimises

$$V(\lambda) = \frac{\frac{1}{n}\left\|(\mathbf{I} - \mathbf{A}(\lambda)\mathbf{g}\right\|^2}{\left[\frac{1}{n}\,\mathrm{trace}(\mathbf{I} - \mathbf{A}(\lambda))\right]^2}, \tag{9.42}$$

where

$$\mathcal{K}\mathbf{f}^*_\lambda = \mathbf{A}(\lambda)\mathbf{g}, \tag{9.43}$$

i.e., $\mathbf{A}(\lambda)$ is the matrix which takes the vector of values $g(x_i)$ into $\mathcal{K}f^*_\lambda(x_i)$.

An approximate representation is required for $\mathbf{f}^*_\lambda$. Bennell & Mason (1989) adopt a basis of polynomials orthogonal on $[a, b]$, and more specifically the Chebyshev polynomial sum

$$f^*_\lambda(y) = \sum_{j=0}^{m} a_j T_j(y) \tag{9.44}$$

when $[a, b] = [-1, 1]$.

9.6.2 Details of a smoothing algorithm (second derivative regularisation)

Adopting the representation (9.44), the smoothing term in (9.40) is

$$\int_{-1}^{1}\left[f^{*\prime\prime}_\lambda(y)\right]^2 \mathrm{d}y = \hat{\mathbf{a}}^T\mathbf{B}\hat{\mathbf{a}} \tag{9.45}$$

where $\hat{\mathbf{a}} = (a_2, a_3, \ldots, a_m)^T$ and $\mathbf{B}$ is a matrix with elements

$$B_{ij} = \int_{-1}^{1} P_i''(y)P_j''(y)\,\mathrm{d}y \quad (i, j = 2, \ldots, m). \tag{9.46}$$

The matrix $\mathbf{B}$ is symmetric and positive definite, with a Cholesky decomposition $\mathbf{B} = \mathbf{L}\mathbf{L}^T$, giving

$$\int_{-1}^{1} \left[f_\lambda^{*\prime\prime}(y)\right]^2 \,\mathrm{d}y = \left\|\mathbf{L}^T\hat{\mathbf{a}}\right\|^2.$$

Then, from (9.40),

$$I[f_\lambda^*] = \frac{1}{n}\left\|\mathbf{M}\mathbf{a} - \mathbf{g}^*\right\|^2 + \lambda\left\|\mathbf{L}^T\hat{\mathbf{a}}\right\|^2 \tag{9.47}$$

where $M_{ij} = \int_{-1}^{1} K(x_i, y)T_j(y)\,\mathrm{d}y$ and $\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$.

Bennell & Mason (1989) show that a_0 and a_1 may be eliminated by considering the $\mathbf{QU}$ decomposition of $\mathbf{M}$,

$$\mathbf{M} = \mathbf{QU} = \mathbf{Q}\left[\frac{\mathbf{V}}{\mathbf{0}}\right], \tag{9.48}$$

where $\mathbf{Q}$ is orthogonal, $\mathbf{V}$ is upper triangular of order $m+1$ and then

$$\mathbf{U} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{0} & \mathbf{R}_3 \end{bmatrix} \tag{9.49}$$

where $\mathbf{R}_1$ is a 2×2 matrix.

Defining $\tilde{\mathbf{a}} = (a_0, a_1)^T$,

$$\begin{aligned} \left\|\mathbf{M}\mathbf{a} - \mathbf{g}^*\right\|^2 &= \left\|\mathbf{QU}\mathbf{a} - \mathbf{g}^*\right\|^2 \\ &= \left\|\mathbf{Q}^T(\mathbf{QU}\mathbf{a} - \mathbf{g}^*)\right\|^2 \\ &= \left\|\mathbf{U}\mathbf{a} - \mathbf{e}\right\|^2, \text{ where } \mathbf{e} = \mathbf{Q}^T\mathbf{g}^*, \\ &= \left\|\mathbf{R}_1\tilde{\mathbf{a}} + \mathbf{R}_2\hat{\mathbf{a}} - \tilde{\mathbf{e}}\right\|^2 + \left\|\mathbf{R}_3\hat{\mathbf{a}} - \hat{\mathbf{e}}\right\|^2. \end{aligned} \tag{9.50}$$

Setting $\tilde{\mathbf{a}} = \mathbf{R}_1^{-1}(\tilde{\mathbf{e}} - \mathbf{R}_2\hat{\mathbf{a}})$,

$$I[f_\lambda^*] = \frac{1}{n}\left\|\mathbf{R}_3\hat{\mathbf{a}} - \hat{\mathbf{e}}\right\|^2 + \lambda\left\|\mathbf{L}^T\hat{\mathbf{a}}\right\|^2. \tag{9.51}$$

The problem of minimising I over $\hat{\mathbf{a}}$ now involves only the independent variables $a_2, \ldots, a_m$, and requires us to solve the equation

$$(\mathbf{H}^T\mathbf{H} + n\lambda\mathbf{I})\mathbf{b} = \mathbf{H}^T\hat{\mathbf{e}} \tag{9.52}$$

where $\mathbf{b} = \mathbf{L}^T\hat{\mathbf{a}}$ and $\mathbf{H} = \mathbf{R}_3(\mathbf{L}^T)^{-1}$.

Hence

$$\mathbf{b} = (\mathbf{H}^T\mathbf{H} + n\lambda\mathbf{I})^{-1}\mathbf{H}^T\hat{\mathbf{e}} \tag{9.53}$$

and it can readily be seen that the GCV matrix is

$$\mathbf{A}(\lambda) = \mathbf{H}(\mathbf{H}^T\mathbf{H} + n\lambda\mathbf{I})^{-1}\mathbf{H}^T. \tag{9.54}$$

The algorithm thus consists of solving the linear system (9.52) for a given λ while minimising $V(\lambda)$ given by (9.42).

Formula (9.42) may be greatly simplified by first determining the singular value decomposition (SVD) of $\mathbf{H}$

$$\mathbf{H} = \mathbf{W}\mathbf{\Lambda}\mathbf{X}^T \quad (\mathbf{W},\ \mathbf{X} \text{ orthogonal})$$

where

$$\mathbf{\Lambda} = \left[\begin{array}{c}\mathbf{\Delta} \\ \hline \mathbf{0}\end{array}\right] \quad (\mathbf{\Delta} = \operatorname{diag}(d_i)).$$

Then

$$V(\lambda) = \frac{\sum_{k=1}^{m-1}[n\lambda(d_k^2 + n\lambda)^{-1}]^2 z_k^2 + \sum_{k=1}^{m-2} z_k^2}{\sum_{k=1}^{m-1}\lambda(d_k^2 + n\lambda)^{-1} + (n-m-1)n^{-1}} \tag{9.55}$$

where $\mathbf{z} = \mathbf{W}^T\hat{\mathbf{e}}$.

The method has been successfully tested by Bennell & Mason (1989) on a number of problems of the form (9.37), using Chebyshev polynomials. It was noted that there was an optimal choice of the number m of basis functions, beyond which the approximation f_λ^* deteriorated on account of ill-conditioning. In Figures 9.1–9.3, we compare the true solution (dashed curve) with the computed Chebyshev polynomial solution (9.44) (continuous curve) for the function $f(y) = \mathrm{e}^{-y}$ and equation

$$\int_0^\infty \mathrm{e}^{-xy} f(y)\,\mathrm{d}y = \frac{1}{1+x}, \quad 0 < x < \infty,$$

with

- $\epsilon(x) \sim N(0, .005^2)$ and $m = 5$,
- $\epsilon(x) \sim N(0, .01^2)$ and $m = 5$,
- $\epsilon(x) \sim N(0, .01^2)$ and $m = 10$.

No significant improvement was obtained for any other value of m.

9.6.3 A smoothing algorithm with weighted function regularisation

Some simplifications occur in the above algebra if, as proposed by Mason & Venturino (1997), in place of (9.40) we minimise the functional

$$I[f^*] \equiv \frac{1}{n}\sum_{i=1}^{n}\left[\mathcal{K}f^*(x_i) - g(x_i)\right]^2 + \lambda\int_a^b w(y)[f^*(y)]^2\,\mathrm{d}y. \tag{9.56}$$

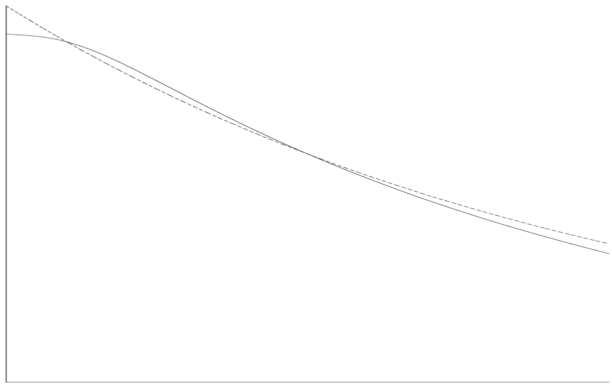


Figure 9.1: Data error $N(0, 0.005^2)$; 5 approximation coefficients

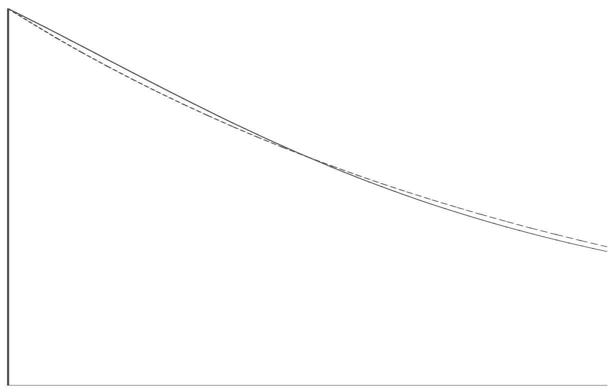


Figure 9.2: Data error $N(0, 0.01^2)$; 5 approximation coefficients

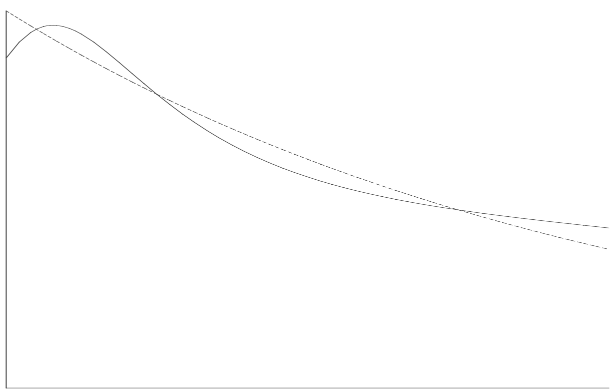


Figure 9.3: Data error $N(0, 0.01^2)$; 10 approximation coefficients

This is closer to the Tikhonov form (9.38) than is (9.40), and involves weaker assumptions about the smoothness of f.

Again we adopt an orthogonal polynomial sum to represent f^*. We choose $w(x)$ to be the weight function corresponding to the orthogonality. In particular, for the first-kind Chebyshev polynomial basis on $[-1, 1]$, and the approximation

$$f^*_\lambda(y) = \sum_{j=0}^{m} a_j T_j(y), \tag{9.57}$$

the weight function is of course $w(x) = 1/\sqrt{1-x^2}$.

The main changes to the procedure of Section 9.6.2 are that

- in the case of (9.56) we do *not* now need to separate a_0 and a_1 off from the other coefficients a_r, and
- the smoothing matrix corresponding to $\mathbf{B}$ in (9.46) becomes diagonal, so that no $\mathbf{LL}^T$ decomposition is required.

For simplicity, we take the orthonormal basis on $[-1, 1]$, replacing (9.57) with

$$f^*_\lambda(y) = \sum_{j=0}^{m} a_j \phi_j(y) = \sum_{j=0}^{m} a_j \left[T_j(y)/n_j\right], \tag{9.58}$$

where

$$n_j^2 = \begin{cases} 2/\pi, & j > 0; \\ 1/\pi, & j = 0. \end{cases} \tag{9.59}$$

Define two inner products (respectively discrete and continuous);

$$\langle u\,,\, v\rangle_{\mathrm{d}} = \sum_{k=1}^{n} u(x_k)v(x_k); \tag{9.60}$$

$$\langle u\,,\, v\rangle_{\mathrm{c}} = \int_{-1}^{1} w(x)u(x)v(x)\,\mathrm{d}x. \tag{9.61}$$

Then the minimisation of (9.56), for f^* given by (9.58), leads to the system of equations

$$\sum_{j=0}^{m} a_j \langle \mathcal{K}\phi_i\,,\, \mathcal{K}\phi_j\rangle_{\mathrm{d}} - \langle \mathcal{K}\phi_i\,,\, \mathbf{g}^*\rangle_{\mathrm{d}} + n\lambda \sum_{j=0}^{m} a_j \langle \phi_i\,,\, \phi_j\rangle_{\mathrm{c}} = 0, \quad i = 0, \ldots, m. \tag{9.62}$$

Hence

$$(\mathbf{Q}^T\mathbf{Q} + n\lambda\mathbf{I})\mathbf{a} = \mathbf{Q}^T\mathbf{g}^* \tag{9.63}$$

(as a consequence of the orthonormality), where

$$Q_{k,j} = (\mathcal{K}\phi_j)(x_k) \tag{9.64}$$

and $\mathbf{a}$, $\mathbf{g}^*$ are vectors with components a_j and $g^*(x_j)$, respectively.

To determine f^* we need to solve (9.63) for a_j, with λ minimising $V(\lambda)$ as defined in (9.42). The matrix $\mathbf{A}(\lambda)$ in (9.42) is to be such that

$$\mathcal{K}\mathbf{f}^* = \mathbf{A}(\lambda)\mathbf{g}^*. \tag{9.65}$$

Now $\mathcal{K}\mathbf{f}^* = \left\{\sum_j a_j \mathcal{K}\phi_j(x_k)\right\} = \mathbf{Qa}$ and hence, from (9.63)

$$\mathbf{A}(\lambda) = \mathbf{Q}(\mathbf{Q}^T\mathbf{Q} + n\lambda\mathbf{I})^{-1}\mathbf{Q}^T. \tag{9.66}$$

9.6.4 Evaluation of $V(\lambda)$

It remains to clarify the remaining details of the algorithm of Section 9.6.3, and in particular to give an explicit formula for $V(\lambda)$ based on (9.66).

Let

$$\mathbf{Q} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{X}^T \tag{9.67}$$

be the singular value decomposition of $\mathbf{Q}$, where $\mathbf{W}$ is $n \times n$ orthogonal, $\mathbf{X}$ is $(m+1) \times (m+1)$ orthogonal and

$$\boldsymbol{\Lambda} = \left[\begin{array}{c} \boldsymbol{\Delta}_m \\ \hline \mathbf{0} \end{array}\right]_{n\times(m+1)}. \tag{9.68}$$

Define

$$\mathbf{z} = [z_k] = \mathbf{W}^T\mathbf{g}. \tag{9.69}$$

From (9.68),

$$\boldsymbol{\Lambda}^T\boldsymbol{\Lambda} = \mathrm{diag}(d_0^2, \ldots, d_m^2). \tag{9.70}$$

It follows that

$$\mathbf{A}(\lambda) = \mathbf{W}\mathbf{B}(\lambda)\mathbf{W}^T \tag{9.71}$$

where

$$\mathbf{B}(\lambda) = \Lambda(\boldsymbol{\Lambda}^T\boldsymbol{\Lambda} + n\lambda\mathbf{I})^{-1}\boldsymbol{\Lambda}^T \tag{9.72}$$

so that $\mathbf{B}(\lambda)$ is the $n \times n$ diagonal matrix with elements

$$B_{kk} = \frac{d_k^2}{d_k^2 + n\lambda} \quad (0 \le k \le m); \qquad B_{kk} = 0 \quad (k > m). \tag{9.73}$$

From (9.71) and (9.72)

$$\begin{aligned}
\|(\mathbf{I}-\mathbf{A}(\lambda))\mathbf{g}\|^2 &= \|(\mathbf{I}-\mathbf{A}(\lambda))\mathbf{W}\mathbf{z}\|^2 \\
&= \left\|\mathbf{W}^T(\mathbf{I}-\mathbf{A}(\lambda))\mathbf{W}\mathbf{z}\right\|^2 \\
&= \left\|(\mathbf{I}-\mathbf{W}^T\mathbf{A}(\lambda)\mathbf{W})\mathbf{z}\right\|^2 \\
&= \|(\mathbf{I}-\mathbf{B}(\lambda))\mathbf{z}\|^2 \\
&= \sum_{i=0}^{m}\left(\frac{n\lambda}{d_i^2+n\lambda}\right)^2 z_i^2 + \sum_{i=m+1}^{n-1} z_i^2.
\end{aligned}$$

Thus

$$\|(\mathbf{I}-\mathbf{A}(\lambda))\mathbf{g}\|^2 = \sum_{i=0}^{m} n^2 e_i^2 z_i^2 + \sum_{i=m+1}^{n} z_i^2 \tag{9.74}$$

where

$$e_i(\lambda) = \frac{\lambda}{d_i^2+n\lambda}. \tag{9.75}$$

Also

$$\begin{aligned}
\operatorname{trace}(\mathbf{I}-\mathbf{A}(\lambda)) &= \operatorname{trace}\mathbf{W}^T(\mathbf{I}-\mathbf{A}(\lambda))\mathbf{W} \\
&= \operatorname{trace}(\mathbf{I}-\mathbf{B}(\lambda)) \\
&= \sum_{i=0}^{m}\frac{n\lambda}{d_i^2+n\lambda} + \sum_{i=m+1}^{n-1} 1.
\end{aligned}$$

Thus

$$\operatorname{trace}(\mathbf{I}-\mathbf{A}(\lambda)) = \sum_{i=0}^{m} n^2 e_i^2 + (n-m-1). \tag{9.76}$$

Finally, from (9.75) and (9.76), together with (9.42), it follows that

$$V(\lambda) = \frac{\displaystyle\sum_{0}^{m} n e_i^2 z_i^2 + \frac{1}{n}\sum_{m+1}^{n} z_i^2}{\left[\displaystyle\sum_{0}^{m} n e_i^2 + \frac{n-m-1}{n}\right]^2}. \tag{9.77}$$

9.6.5 Other basis functions

It should be pointed out that Chebyshev polynomials are certainly not the only basis functions that could be used in the solution of (9.37) by regularisation. Indeed there is a discussion by Bennell & Mason (1989, Section ii) of three alternative basis functions, each of which yields an efficient algorithmic procedure, namely:

1. a kernel function basis $\{K(x_i, y)\}$,
2. a B-spline basis, and
3. an eigenfunction basis.

Of these, an eigenfunction basis is the most convenient (provided that eigenfunctions are known), whereas a kernel function basis is rarely of practical value. A B-spline basis is of general applicability and possibly comparable to, or slightly more versatile than, the Chebyshev polynomial basis. See Rodriguez & Seatzu (1990) and also Bennell & Mason (1989) for discussion of B-spline algorithms.

9.7 Partial differential equations and boundary integral equation methods

Certain classes of partial differential equations, with suitable boundary conditions, can be transformed into integral equations on the boundary of the domain. This is particularly true for equations related to the Laplace operator. Methods based on the solution of such integral equations are referred to as *boundary integral equation* (BIE) methods (Jaswon & Symm 1977, for instance) or, when they are based on discrete element approximations, as *boundary element methods* (BEM) (Brebbia et al. 1984). Chebyshev polynomials have a part to play in the solution of BIEs, since they lead typically to kernels related to the Hilbert kernel discussed in Section 9.5.1.

We now illustrate the role of Chebyshev polynomials in BIE methods for a particular mixed boundary value problem for Laplace's equation, which leads to a hypersingular boundary integral equation.

9.7.1 A hypersingular integral equation derived from a mixed boundary value problem for Laplace's equation

Derivation

In this section we tackle a 'hard' problem, which relates closely to the hypersingular integral relationship (9.30) satisfied by Chebyshev polynomials of the second kind. The problem and method are taken from Mason & Venturino (1997).

Consider Laplace's equation for $u(x, y)$ in the positive quadrant

$$\Delta u = 0, \quad x, y \geq 0, \tag{9.78}$$

subject to (see Figure 9.4)

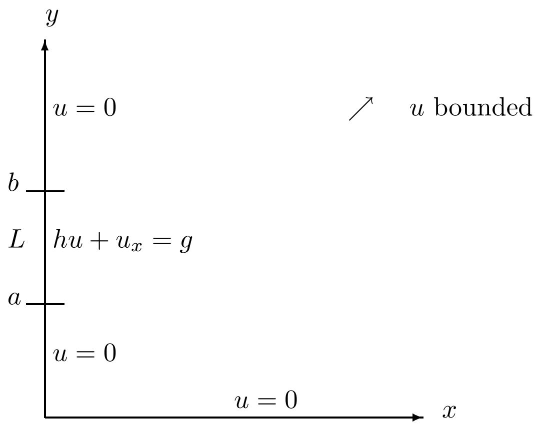


Figure 9.4: Location of the various boundary conditions (9.79)

$$u(x,0) = 0, \quad x \geq 0, \tag{9.79a}$$

$$hu(0,y) + u_x(0,y) = g(y), \quad 0 < a \leq y \leq b, \tag{9.79b}$$

$$u(0,y) = 0, \quad 0 \leq y < a;\ b < y, \tag{9.79c}$$

$$u(x,y) \text{ is bounded}, \quad x,y \to \infty. \tag{9.79d}$$

Thus the boundary conditions are homogeneous apart from a window $L \equiv [a,b]$ of radiation boundary conditions, and the steady-state temperature distribution in the positive quadrant is sought. Here the boundary conditions are 'mixed' in two senses: involving both u and u_x on L and splitting into two different operators on $x = 0$. Such problems are known to lead to Cauchy singular integral equations (Venturino 1986), but in this case a different approach leads to a hypersingular integral equation closely related to (9.30).

By separation of variables in (9.78), using (9.79a) and (9.79d), we find that

$$u(x,y) = \int_0^\infty A(\mu)\sin(\mu y)\exp(-\mu x)\,\mathrm{d}\mu. \tag{9.80}$$

The zero conditions (9.79c) on the complement L^c of L give

$$u(0,y) = \lim_{x\to 0+} \int_0^\infty A(\mu)\sin(\mu y)\exp(-\mu x)\,\mathrm{d}\mu = 0, \quad y \in L^c, \tag{9.81}$$

and differentiation of (9.80) with respect to x in L gives

$$u_x(0,y) = -\lim_{x\to 0+} \int_0^\infty \mu A(\mu)\sin(\mu y)\exp(-\mu x)\,\mathrm{d}\mu = 0, \quad y \in L. \tag{9.82}$$

Substitution of (9.80) and (9.82) into (9.79b) leads to

$$\lim_{x\to 0+} \int_0^\infty (h-\mu)A(\mu)\sin(\mu y)\exp(-\mu x)\,\mathrm{d}\mu = g(y), \quad y \in L. \tag{9.83}$$

Then (9.81) and (9.83) are a pair of dual integral equations for $A(\mu)$, and from which we can deduce u by using (9.80).

To solve (9.81) and (9.83), we define a function $B(y)$ as

$$B(y) := u(0,y) = \int_0^\infty A(\mu)\sin(\mu y)\,\mathrm{d}\mu, \quad y \geq 0. \tag{9.84}$$

Then, from (9.81)

$$B(y) = 0, \quad y \in L^c, \tag{9.85}$$

and, inverting the sine transform (9.84) and using (9.85),

$$\int_L B(t)\sin(st)\,\mathrm{d}t = \tfrac{1}{2}\pi A(s). \tag{9.86}$$

Substituting (9.86) in the integral equation (9.83) gives us

$$hB(y) - \frac{2}{\pi}\int_L I(t)\,\mathrm{d}t = g(t), \quad y \in L \tag{9.87}$$

where

$$\begin{aligned} I(t) &= \lim_{x\to 0+} \int_0^\infty \mu \sin(\mu t)\exp(-\mu x)\,\mathrm{d}\mu \\ &= \tfrac{1}{2}\lim_{x\to 0+} \int_0^\infty \mu\left[\cos\mu(t-y) - \cos\mu(t+y)\right]\exp(-\mu x)\,\mathrm{d}\mu. \end{aligned} \tag{9.88}$$

This simplifies (see Problem 7) to

$$\begin{aligned} I(t) &= \tfrac{1}{2}\lim_{x\to 0+}\left[\frac{x^2-(t-y)^2}{(x^2+(t-y)^2)^2} - \frac{x^2-(t+y)^2}{(x^2+(t+y)^2)^2}\right] \\ &= -\tfrac{1}{2}\left[\frac{1}{(t-y)^2} - \frac{1}{(t+y)^2}\right]. \end{aligned} \tag{9.89}$$

Substituting (9.89) into (9.87), we obtain the hypersingular integral equation, with strong singularity at $t = y$,

$$hB(y) + \frac{1}{\pi}\int_L B(t)\left[\frac{1}{(t-y)^2} - \frac{1}{(t+y)^2}\right]\mathrm{d}t = g(y), \quad y \in L, \tag{9.90}$$

from which $B(y)$ is to be determined, and hence $A(s)$ from (9.86) and $u(x,y)$ from (9.80).

Method of solution

Continuing to follow Mason & Venturino (1997), equation (9.90) can be rewritten in operator form as

$$\mathcal{A}\phi \equiv (h + \mathcal{H} + \mathcal{K})\phi = f \tag{9.91}$$

where $\mathcal{H}$ is a Hadamard finite-part integral and $\mathcal{K}$ is a compact perturbation, given by

$$(\mathcal{H}\phi)(x) = \int_{-1}^{1} \frac{\phi(s)}{(s-x)^2}\,\mathrm{d}s, \quad -1 < x < 1, \tag{9.92}$$

$$(\mathcal{K}\phi)(x) \equiv \int_{-1}^{1} K(x,s)\phi(s)\,\mathrm{d}s = \int_{-1}^{1} \frac{\phi(s)}{(s+x)^2}\,\mathrm{d}s, \quad -1 < x < 1, \tag{9.93}$$

and

$$f(x) = g\left(\tfrac{1}{2}(b-a)x + \tfrac{1}{2}(b+a)\right). \tag{9.94}$$

It is clear that $\phi(x)$ must vanish at the end points ± 1, since it represents boundary values, and moreover it should possess a square-root singularity (Martin 1991). Hence we write

$$\phi(x) = w(x)y(x), \text{ where } w(x) = \sqrt{1-x^2}. \tag{9.95}$$

We note also that the Hadamard finite-part operator maps second kind Chebyshev polynomials into themselves, as shown by Mason (1993) and Martin (1992) and indicated in (9.30) above; in fact

$$\mathcal{H}(wU_\ell)(x) = -\pi(\ell+1)U_\ell(x), \quad \ell \geq 0. \tag{9.96}$$

Solution of (9.86) in terms of second-kind polynomials is clearly suggested, namely

$$y(x) = \sum_{\ell=0}^{\infty} c_\ell U_\ell(x), \tag{9.97}$$

where the coefficients c_ℓ are to be determined, and we therefore define a weighted inner product

$$\langle u\,,\,v\rangle_w := \int_{-1}^{1} w(t)u(t)v(t)\,\mathrm{d}t$$

and observe that

$$\|U_\ell\|_w^2 = \tfrac{1}{2}\pi, \quad \ell \geq 0. \tag{9.98}$$

We also expand both $f(x)$, the right-hand side of (9.91), and $K(x,t)$ in second-kind polynomials

$$f(x) = \sum_{i=0}^{\infty} f_j U_j(x), \text{ where } f_j = \frac{2}{\pi}\langle f\,,\,U_j\rangle_w, \tag{9.99}$$

$$K(x,t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} K_{ij} U_i(x) U_j(t), \tag{9.100}$$

so that (9.93), (9.97) and (9.98) give

$$\begin{aligned} \mathcal{K}\phi &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} c_\ell K_{ij} \int_{-1}^{1} w(t) U_i(x) U_j(t) u_\ell(t) \, \mathrm{d}t \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} c_\ell K_{ij} U_i(x) \left\| U_j \right\|_w^2 \delta_{j\ell} \\ &= \tfrac{1}{2}\pi \sum_{\ell=0}^{\infty} c_\ell \sum_{i=0}^{\infty} K_{i\ell} U_i(x). \end{aligned} \tag{9.101}$$

Substituting (9.95), (9.97), (9.99), (9.101) and (9.99) into (9.91):

$$hw \sum_{\ell=0}^{\infty} c_\ell U_\ell(x) - \pi \sum_{\ell=0}^{\infty} (\ell+1) c_\ell U_\ell(x) + \tfrac{1}{2}\pi \sum_{\ell=0}^{\infty} c_\ell \sum_{i=0}^{\infty} K_{i\ell} U_i(x) = \sum_{j=0}^{\infty} f_j U_j(x). \tag{9.102}$$

Taking the weighted inner product with U_j:

$$\begin{aligned} h \sum_{\ell=0}^{\infty} c_\ell \left\langle wU_\ell \, , \, U_j \right\rangle_w - \pi \sum_{\ell=0}^{\infty} (\ell+1) c_\ell \left\langle U_\ell \, , \, U_j \right\rangle_w + \\ + \tfrac{1}{2}\pi \sum_{\ell=0}^{\infty} c_\ell \sum_{i=0}^{\infty} K_{i\ell} \left\langle U_i \, , \, U_j \right\rangle_w = \frac{1}{2}\pi f_j. \end{aligned} \tag{9.103}$$

Define

$$b_{jl} := \left\langle wU_\ell \, , \, U_j \right\rangle_w = \int_{-1}^{1} (1-x^2) U_\ell(x) U_j(x) \, \mathrm{d}x. \tag{9.104}$$

Then it can be shown (Problem 8) that

$$b_{j\ell} = \begin{cases} \dfrac{1}{(\ell+j+2)^2 - 1} - \dfrac{1}{(\ell-j)^2 - 1}, & j+\ell \text{ even}, \\ 0, & \text{otherwise}. \end{cases} \tag{9.105}$$

Hence, from (9.103),

$$h \sum_{\ell=0}^{\infty} b_{j\ell} c_\ell - \tfrac{1}{2}\pi^2 (j+1) c_j + (\tfrac{1}{2}\pi)^2 \sum_{\ell=0}^{\infty} K_{j\ell} c_\ell = \frac{1}{2}\pi f_j, \quad 0 \le j < \infty. \tag{9.106}$$

Reducing (9.106) to a finite system, to solve for approximate coefficients $\hat{c}_\ell$, we obtain

$$h \sum_{\ell=0}^{N-1} b_{j\ell} \hat{c}_\ell - \tfrac{1}{2}\pi^2 (j+1) \hat{c}_j + (\tfrac{1}{2}\pi)^2 \sum_{\ell=0}^{N-1} K_{j\ell} \hat{c}_\ell = \frac{1}{2}\pi f_j, \quad 0 \le j < N-1. \tag{9.107}$$

Table 9.1: Results for $\mathcal{K} = 0$, $\phi(x) = \sqrt{1 - x^2}\exp x$

N	condition number	$\|e_N\|_\infty$
1	1.93	
2	2.81	5×10^{-1}
4	4.17	5×10^{-3}
8	6.54	2×10^{-8}
16	10.69	1.5×10^{-13}

EXAMPLE 9.1: The method is tested by Mason & Venturino (1997) for a slightly different problem, where the well-behaved part $\mathcal{K}$ of the problem is set to zero and the function f is chosen so that $\phi(x) \equiv \sqrt{1 - x^2}\exp x$. The condition number of the matrix of the linear system (9.107) defining $\hat{c}_j$ is compared in Table 9.1 with the maximum error $\|e_N\|_\infty$ for various values of N, and it is clear that the conditioning is relatively good and the accuracy achieved is excellent.

Error analysis

A rigorous error analysis has been carried out by Mason & Venturino (1997), but the detail is much too extensive to quote here. However, the conclusion reached was that, if $f \in \mathcal{C}^{p+1}[-1, 1]$ and the integral operator $\mathcal{K}$ satisfies certain inequalities, then the method is convergent and

$$\|e_N\|_\infty \leq C.N^{-(p+1)} \tag{9.108}$$

where the constant C depends on the smoothness of K and f but not on N.

For further studies of singular integral equations involving a Cauchy kernel, see Elliott (1989) and Venturino (1992, 1993).

9.8 Problems for Chapter 9

1. Follow through all steps in detail of the proofs of Theorem 9.1 and Corollary 9.1A.

2. Using Corollary 9.1A, find a function $g(x)$ such that

$$g(x) = -\!\!\!\!\!\!\int_{-1}^{1} \frac{\sqrt{1 - y^2}f(y)}{(y - x)}\,\mathrm{d}y$$

   in the cases

(a) $f(y) = 1$;
(b) $f(y) = y^6$;
(c) $f(y) = \mathrm{e}^y$;
(d) $f(y) = \sqrt{1-y^2}$.

3. Using Corollary 9.1A, find a function $g(x)$ such that

$$g(x) = -\int_{-1}^{1} \frac{f(y)}{\sqrt{1-y^2}(y-x)}\,\mathrm{d}y$$

in the cases

(a) $g(x) = \mathrm{e}^x$;
(b) $g(x) = (1+x)^{\frac{1}{2}}(1-x)^{-\frac{1}{2}}$;
(c) $g(x) = x^5$;
(d) $g(x) = 1$.

4. Prove Theorem 9.2 in detail. For instance, the second kernel K_6 in the theorem is derived from

$$K_6(x,y) = \int_{-1}^{x} \frac{K_2(x,y)}{\sqrt{1-x^2}}\,\mathrm{d}x.$$

Setting $x = \cos 2\phi$, $y = \cos 2\psi$ and $\tan\phi = t$, show that K_6 simplifies to

$$\sin 2\psi \int_{\infty}^{t} \frac{\mathrm{d}t}{\sin^2\psi - t^2\cos^2\psi} = \log\left|\frac{\sin(\phi+\psi)}{\sin(\phi-\psi)}\right|.$$

Then, by setting $x = 2u^2 - 1$, $y = 2v^2 - 1$ and noting that $\sqrt{1-x^2} = 2u\sqrt{1-u^2}$, $\sqrt{1-y^2} = 2v\sqrt{1-v^2}$, show that $K_6(x,y)$ simplifies to

$$\log|x-y| - \log\left|1 - xy - \sqrt{(1-x^2)(1-y^2)}\right|.$$

5. By differentiating rather than integrating in (9.22a), (9.22b), (9.22c) and (9.22d), and using the properties

$$\begin{aligned}
[\sqrt{1-x^2}\,U_{n-1}(x)]' &= -nT_n(x)/\sqrt{1-x^2}, \\
[T_n(x)]' &= nU_{n-1}(x), \\
[\sqrt{1-x}\,W_n(x)]' &= (n+\tfrac{1}{2})V_n(x)/\sqrt{1-x}, \\
[\sqrt{1+x}\,V_n(x)]' &= (n+\tfrac{1}{2})W_n(x)/\sqrt{1+x},
\end{aligned}$$

deduce that the integral equation

$$\lambda\phi(x) = \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}}\phi(y)K(x,y)\,\mathrm{d}y$$

has the following eigensolutions ϕ and eigenvalues λ for the following kernels K:

(a) $K(x,y) = K_9(x,y) = \dfrac{\sqrt{1-x^2}}{(y-x)^2} - \dfrac{x}{\sqrt{1-x^2}(y-x)}$;
$\phi = \phi_n(x) = T_{n-1}(x)/\sqrt{1-x^2}$, $\lambda = \lambda_n = -n\pi$.

(b) $K(x,y) = K_{10}(x,y) = \dfrac{1-y^2}{(y-x)^2}$;
$\phi = \phi_n(x) = U_{n-1}(x)$, $\lambda = \lambda_n = -n\pi$.

(c) $K(x,y) = K_{11}(x,y) = \dfrac{\sqrt{(1-x)(1+y)}}{(y-x)^2} - \dfrac{\sqrt{1+y}}{2\sqrt{1-x}(y-x)}$;
$\phi = \phi_n(x) = V_n(x)/\sqrt{1-x}$, $\lambda = \lambda_n = -(n+\frac{1}{2})\pi$.

(d) $K(x,y) = K_{12}(x,y) = \dfrac{\sqrt{(1+x)(1-y)}}{(y-x)^2} + \dfrac{\sqrt{1-y}}{2\sqrt{1+x}(y-x)}$;
$\phi = \phi_n(x) = W_n(x)/\sqrt{1+x}$, $\lambda = \lambda_n = -(n+\frac{1}{2})\pi$.

6. (a) Describe and discuss possible amendments that you might make to the regularisation methods of Section 9.6 in case K has any one of the four singular forms listed in Theorem 9.1. Does the method simplify?

   (b) Discuss whether or not it might be better, for a general K, to use one of the Chebyshev polynomials other than $T_j(x)$ in the approximation (9.44).

7. Show that

$$\int_0^\infty \mu \sin \mu t \sin \mu y \exp(-\mu x)\, d\mu = \frac{x^2 - (t-y)^2}{(x^2 + (t-y)^2)^2} - \frac{x^2 - (t+y)^2}{(x^2 + (t+y)^2)^2}.$$

(This completes the determination of $I(t)$, given by (9.88), so as to derive the hypersingular equation (9.90).)

8. Show that

$$\int_{-1}^{1} (1-x^2) U_\ell(x) U_j(x)\, dx = \frac{1}{(\ell + j + 2)^2 - 1} - \frac{1}{(\ell - j)^2 - 1}$$

for $\ell + j$ even, and that the integral vanishes otherwise. (This is a step required in the derivation of the solution of the hypersingular equation (9.90).)

Chapter 10

Solution of Ordinary Differential Equations

10.1 Introduction

While, historically, finite-difference methods have been and remain the standard numerical technique for solving ordinary differential equations, newer alternative methods can be more effective in certain contexts. In particular we consider here methods founded on orthogonal expansions—the so-called *spectral* and *pseudospectral* methods—with special reference to methods based on expansions in Chebyshev polynomials.

In a typical finite-difference method, the unknown function $u(x)$ is represented by a table of numbers $\{y_0, y_1, \ldots, y_n\}$ approximating its values at a set of discrete points $\{x_0, x_1, \ldots, x_n\}$, so that $y_j \approx u(x_j)$. (The points are almost always equally spaced through the range of integration, so that $x_{j+1} - x_j = h$ for some small fixed h.)

In a spectral method, in contrast, the function $u(x)$ is represented by an infinite expansion $u(x) = \sum_k c_k \phi_k(x)$, where $\{\phi_k\}$ is a chosen sequence of prescribed basis functions. One then proceeds somehow to estimate as many as possible of the coefficients $\{c_k\}$, thus approximating $u(x)$ by a finite sum such as

$$u_n(x) = \sum_{k=0}^{n} c_k \phi_k(x). \tag{10.1}$$

One clear advantage that spectral methods have over finite-difference methods is that, once approximate spectral coefficients have been found, the approximate solution can immediately be evaluated at any point in the range of integration, whereas to evaluate a finite-difference solution at an intermediate point requires a further step of interpolation.

A pseudospectral method, at least according to some writers, is one in which $u(x)$ is still approximated by a function of the form $u_n(x)$ of (10.1), as in a spectral method, but this approximation is actually represented not by its coefficients but by its values $u_n(x_j)$ at a number ($n+1$ in this particular instance) of discrete points $\{x_j\}$. These points may be equally spaced, but equal spacing gives no advantages, and other spacings are frequently better.

The oldest and probably the most familiar spectral methods are based on the idea of Fourier series. Supposing for the moment, for convenience, that the independent variable x is confined to the interval $-\pi \le x \le \pi$, the technique is to assume that the unknown function has an expansion in the form

$$u(x) = \tfrac{1}{2}a_0 + \sum_{k=1}^{\infty} \{a_k \cos kx + b_k \sin kx\} \tag{10.2}$$

and to attempt to determine values for the coefficients $\{a_k, b_k\}$ such that the required differential equation and other conditions are satisfied.

Fourier methods may well be suitable when the problem is inherently periodic; for instance where the function $u(x)$ satisfies a second-order differential equation subject to the periodicity boundary conditions $u(\pi) = u(-\pi)$ and $u'(\pi) = u'(-\pi)$. If we have a second-order differential equation with the more usual boundary conditions $u(-\pi) = a$ and $u(\pi) = b$, however, with $a \neq b$, then obviously any finite partial sum

$$\tfrac{1}{2}a_0 + \sum_{k=1}^{n} \{a_k \cos kx + b_k \sin kx\}$$

of (10.2) is periodic and cannot satisfy both boundary conditions simultaneously; more importantly, very many terms are needed if this partial sum is to represent the function $u(x)$ at all closely near both ends of the range at the same time. It is better in such a case to take for $\{\phi_k\}$ a sequence of polynomials, so that the partial sum

$$\sum_{k=0}^{n} c_k \phi_k(x)$$

is a polynomial of degree n.

From now on, we shall suppose that the independent variable x is confined to the interval $-1 \leq x \leq 1$, so that a reasonable choice for $\phi_k(x)$ is the Chebyshev polynomial $T_k(x)$. The choice of basis is only the beginning of the story, however; we are still left with the task of determining the coefficients $\{c_k\}$.

10.2 A simple example

For the purposes of illustration, we shall consider the simple linear two-point boundary-value problem on the range $[-1, 1]$:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b, \tag{10.3}$$

where the function f and the boundary values a and b are given.

We can start out in several ways, of which the two following are the simplest:

- We may write $u(x)$ directly in the form

$$u(x) = {\sum_{k=0}^{\infty}}' c_k T_k(x). \qquad (10.4)$$

Using the result quoted in Problem 16 of Chapter 2, namely

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} T_k(x) = {\sum_{\substack{r=0 \\ (k-r) \text{ even}}}^{k-2}}' (k-r)k(k+r)T_r(x), \ (k \geq 2), \qquad (10.5)$$

we express (10.3) in the form

$$\sum_{k=2}^{\infty} {\sum_{\substack{r=0 \\ (k-r) \text{ even}}}^{k-2}}' (k-r)k(k+r)c_k T_r(x) = f(x), \qquad (10.6a)$$

$${\sum_{k=0}^{\infty}}' (-1)^k c_k = a, \qquad {\sum_{k=0}^{\infty}}' c_k = b. \qquad (10.6b)$$

- An alternative procedure, incorporating the boundary conditions in the representation itself, is to write $u(x)$ in the form

$$u(x) = (1-x^2) \sum_{k=0}^{\infty} \gamma_k U_k(x) + \frac{1-x}{2} a + \frac{1+x}{2} b. \qquad (10.7)$$

Then, using a result given from Problem 12 of Chapter 3,

$$(1-x^2)U_k(x) = \tfrac{1}{2}(T_k(x) - T_{k+2}(x)),$$

together with (10.5), we get (10.3) in a single equation of the form

$$\tfrac{1}{2} \sum_{k=2}^{\infty} {\sum_{\substack{r=0 \\ (k-r) \text{ even}}}^{k-2}}' (k-r)k(k+r)(\gamma_k - \gamma_{k-2})T_r(x) = f(x). \qquad (10.8)$$

We should be treading on dangerous ground, however, if we went ahead blindly with translating either of the above infinite systems of equations (10.6) or (10.8) into an algorithm, since non-trivial questions arise relating to convergence of the infinite series and the validity of differentiating term by term. In any event, since one can never calculate the whole of an infinite sequence of coefficients, the realistic approach is to accept the fact that we must perforce approximate $u(x)$ by a finite sum of terms, and to go on from there.

If we truncate the summation (10.4) or (10.7) after a finite number of terms, we obviously cannot in general satisfy (10.6a) or (10.8) throughout the range $-1 \leq x \leq 1$. We can, however, attempt to satisfy either equation approximately in some sense. We shall discuss two ways of doing this: *collocation* methods and *projection* or *tau* (τ) methods.

10.2.1 Collocation methods

Suppose that we approximate $u(x)$ by

$$u_n(x) := \sum_{k=0}^{n}{}' c_k T_k(x), \tag{10.9}$$

involving $n+1$ unknown coefficients $\{c_k\}$, then we may select $n-1$ points $\{x_1, \ldots, x_{n-1}\}$ in the range of integration and require $u_n(x)$ to satisfy the differential equation (10.3) at just these $n-1$ points, the so-called *collocation points*, in addition to obeying the boundary conditions. This requires us to solve just the system of $n+1$ linear equations

$$\sum_{k=2}^{n} \sum_{\substack{r=0 \\ (k-r)\text{ even}}}^{k-2}{}' (k-r)k(k+r)c_k T_r(x_j) = f(x_j), \qquad j = 1, \ldots, n-1, \tag{10.10a}$$

$$\sum_{k=0}^{n}{}' (-1)^k c_k = a, \tag{10.10b}$$

$$\sum_{k=0}^{n}{}' c_k = b, \tag{10.10c}$$

These equations may be reduced to a simpler form, especially if the $n-1$ points are carefully chosen so that we can exploit discrete orthogonality properties of the Chebyshev polynomials. Suppose that we choose for our collocation points the zeros of $T_{n-1}(x)$, namely

$$x_j = \cos \frac{(j-\frac{1}{2})\pi}{n-1}. \tag{10.11}$$

Multiply (10.10a) by $2T_\ell(x_j)$, where ℓ is an integer with $0 \le \ell \le n-2$, and sum from $j = 1$ to $j = n-1$. We can then use the discrete orthogonality relations (4.42) (for $0 \le r \le n-2$, $0 \le \ell \le n-2$)

$$\sum_{j=1}^{n-1} T_r(x_j)T_\ell(x_j) = \begin{cases} n-1, & r = \ell = 0, \\ \frac{1}{2}(n-1), & r = \ell \ne 0, \\ 0, & \text{otherwise} \end{cases} \tag{10.12}$$

to deduce that

$$\sum_{\substack{k=\ell+2 \\ (k-\ell)\text{ even}}}^{n} (k-\ell)k(k+\ell)c_k = \frac{2}{n-1} \sum_{j=1}^{n-1} T_\ell(x_j) f(x_j), \quad \ell = 0, \ldots, n-2. \tag{10.13}$$

The matrix of coefficients on the left-hand side of equation (10.13) is upper triangular, with elements

$$\begin{pmatrix}
8 & & 64 & & 216 & & 512 & & 1000 & \cdots \\
& 24 & & 120 & & 336 & & 720 & & \cdots \\
& & 48 & & 192 & & 480 & & 960 & \cdots \\
& & & 80 & & 280 & & 648 & & \cdots \\
& & & & 120 & & 384 & & 840 & \cdots \\
& & & & & 168 & & 504 & & \cdots \\
& & & & & & 224 & & 640 & \cdots \\
& & & & & & & 288 & & \cdots \\
& & & & & & & & 360 & \cdots \\
& & & & & & & & & \ddots
\end{pmatrix}.$$

We now have the following algorithm for generating an approximate solution of the problem (10.3) by collocation.

1. Find the collocation points $x_j = \cos\frac{(j-\frac{1}{2})\pi}{n-1}$, for $j = 1, \ldots, n-1$, and evaluate $f(x_j)$;
2. Use the recurrence (1.3) to evaluate $T_\ell(x_j)$, for $\ell = 0, \ldots, n-2$ and $j = 1, \ldots, n-1$;
3. Use equations (10.13), in reverse order, to determine the coefficients $c_n, \ldots, c_3, c_2$ one by one;
4. Use the boundary conditions (10.10b), (10.10c) to determine c_0 and c_1.

This algorithm can be made considerably more efficient in the case where $n-1$ is a large power of 2, as we can then use a technique derived from the fast Fourier transform algorithm to compute the right-hand sides of (10.13) in $O(n \log n)$ operations, without going through step 2 above which requires $O(n^2)$ operations (see Section 4.7).

Alternatively, we approximate $u(x)$ by

$$u_n(x) := (1-x^2)\sum_{k=0}^{n-2} \gamma_k U_k(x) + \frac{1-x}{2}a + \frac{1+x}{2}b, \tag{10.14}$$

involving $n-1$ unknown coefficients $\{\gamma_k\}$ and satisfying the boundary conditions automatically. With the same $n-1$ collocation points $\{x_1, \ldots, x_{n-1}\}$ we now solve the system of $n-1$ linear equations

$$\sum_{k=2}^{n} \sideset{}{'}\sum_{\substack{r=0 \\ (k-r) \text{ even}}}^{k-2} \tfrac{1}{2}(k-r)k(k+r)(\gamma_k - \gamma_{k-2})T_r(x_j) = f(x_j) \tag{10.15}$$

(with $\gamma_n = \gamma_{n-1} = 0$). If the collocation points are again taken as the zeros of $T_{n-1}(x)$, discrete orthogonality gives the equations

$$\sum_{\substack{k=\ell+2 \\ (k-\ell)\text{ even}}}^{n} \tfrac{1}{2}(k-\ell)k(k+\ell)(\gamma_k - \gamma_{k-2}) = \frac{2}{n-1}\sum_{j=1}^{n-1} T_\ell(x_j)f(x_j) \qquad (10.16)$$

or, equivalently,

$$\sum_{\substack{k=\ell \\ (k-\ell)\text{ even}}}^{n-2} (\ell^2 - 3k^2 - 6k - 4)\gamma_k = \frac{2}{n-1}\sum_{j=1}^{n-1} T_\ell(x_j)f(x_j), \quad \ell = 0, \ldots, n-2. \qquad (10.17)$$

The matrix of coefficients on the left-hand side of equation (10.17) is upper triangular, with elements

$$\begin{pmatrix}
-4 & & -28 & & -76 & & -148 & & -244 & \cdots \\
 & -12 & & -48 & & -108 & & -192 & & \cdots \\
 & & -24 & & -72 & & -144 & & -240 & \cdots \\
 & & & -40 & & -100 & & -184 & & \cdots \\
 & & & & -60 & & -132 & & -228 & \cdots \\
 & & & & & -84 & & -168 & & \cdots \\
 & & & & & & -112 & & -208 & \cdots \\
 & & & & & & & -144 & & \cdots \\
 & & & & & & & & -180 & \cdots \\
 & & & & & & & & & \ddots
\end{pmatrix}.$$

We then have the following algorithm.

1. Find the collocation points $x_j = \cos\frac{(j-\frac{1}{2})\pi}{n-1}$, for $j = 1, \ldots, n-1$, and evaluate $f(x_j)$;
2. Use the recurrence (1.3) to evaluate $T_\ell(x_j)$, for $\ell = 0, \ldots, n-2$ and $j = 1, \ldots, n-1$;
3. Solve equations (10.17), in reverse order, to determine the coefficients $\gamma_{n-2}, \ldots, \gamma_1, \gamma_0$ one by one.

EXAMPLE 10.1: Taking the differential equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) + 6\,|x| = 0, \quad u(\pm 1) = 0, \qquad (10.18)$$

whose known solution is $u(x) = 1 - x^2\,|x|$, and applying the above method with $n = 10$, we get the results in Table 10.1, where we show the values of the exact and approximate solutions at the collocation points. (Both solutions are necessarily even.)

Table 10.1: Solution of (10.18) by collocation, $n = 10$

x_j	$u(x_j)$	$u_n(x_j)$
0.0000	1.0000	0.9422
±0.3420	0.9600	0.9183
±0.6427	0.7344	0.7129
±0.8660	0.3505	0.3419
±0.9848	0.0449	0.0440

10.2.2 Error of the collocation method

We may analyse the error of the preceding collocation algorithms by the use of *backward error analysis.*

Here we shall look only at the first of the two alternative representations of $u(x)$. Let $u(x)$ denote the true solution to the problem (10.3), let

$$u_n(x) := \sum_{k=0}^{n}{}' c_k T_k(x)$$

be the approximate solution obtained by collocation and let $f_n(x)$ be the second derivative of $u_n(x)$. Then $u_n(x)$ is itself the exact solution to a similar problem

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u_n(x) = f_n(x), \quad u_n(-1) = a, \quad u_n(+1) = b. \tag{10.19}$$

Since equations (10.3) and (10.19) are both linear, and have the same boundary conditions, the error $e(x) := u(x) - u_n(x)$ must be the solution to the homogeneous boundary-value problem

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} e(x) = \delta f(x), \quad e(-1) = e(+1) = 0, \tag{10.20}$$

where

$$\delta f(x) := f(x) - f_n(x). \tag{10.21}$$

We can write down the solution of (10.20) in integral form

$$e(x) = \int_{-1}^{1} G(x,\xi)\delta f(\xi)\,\mathrm{d}\xi, \tag{10.22}$$

where $G(x, \xi)$ is the *Green's function* for this problem

$$G(x, \xi) = \begin{cases} -\frac{1}{2}(1-x)(1+\xi), & \xi \le x, \\ -\frac{1}{2}(1+x)(1-\xi), & \xi \ge x. \end{cases} \tag{10.23}$$

Equation (10.22) may be used to derive bounds on the error $e(x)$ from various possible norms of the difference $\delta f(x)$. In particular:

$$|e(x)| \le \tfrac{1}{2}(1-x^2)\,\|\delta f\|_\infty\,, \tag{10.24a}$$

$$|e(x)| \le \frac{1}{\sqrt{6}}(1-x^2)\,\|\delta f\|_2\,, \tag{10.24b}$$

$$|e(x)| \le \tfrac{1}{2}(1-x^2)\,\|\delta f\|_1\,. \tag{10.24c}$$

We know that $u_n(x)$ is a polynomial of degree n in x, so that its second derivative $f_n(x)$ must be a polynomial of degree $n-2$. The collocation equations (10.10a), however, tell us that

$$f_n(x_j) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} u_n(x_j) = f(x_j),$$

so that $f_n(x)$ coincides with $f(x)$ at the $n-1$ points $\{x_1, \ldots, x_{n-1}\}$. Therefore $f_n(x)$ must be the unique $(n-2)$nd degree polynomial interpolating $f(x)$ at the zeros of $T_{n-1}(x)$, and $\|\delta f\|$ must be the corresponding interpolation error.

Now we may apply one of the standard formulae for interpolation error, for instance (Davis 1961, Chapter 3):

- [Hermite] Assuming that $f(x)$ has an analytic continuation into the complex plane, we have

$$\delta f(x) = f(x) - f_n(x) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{T_{n-1}(x) f(z)\,\mathrm{d}z}{T_{n-1}(z)(z-x)} \tag{10.25}$$

  where C is a closed contour in the complex plane, encircling the interval $[-1, 1]$ but enclosing no singularities of $f(z)$.

- [Cauchy] Assuming alternatively that $f(x)$ has $(n-1)$ continuous derivatives, we have

$$\delta f(x) = f(x) - f_n(x) = \frac{T_{n-1}(x)}{2^n (n-1)!} f^{(n-1)}(\xi) \tag{10.26}$$

  for some real ξ in the interval $-1 \le \xi \le 1$.

To take (10.25) a little further, suppose that the analytic continuation $f(z)$ of $f(x)$ is regular on and within the ellipse E_r defined in Section 1.4.1, with foci at $z = \pm 1$. Then we know from (1.50) that

$$|T_{n-1}(z)| \ge \frac{1}{2}(r^{n-1} - r^{1-n})$$

for every z on E_r, while $|T_{n-1}(x)| \le 1$ for every x in $[-1, 1]$. Therefore

$$|\delta f(x)| \le \frac{1}{\pi(r^{n-1} - r^{1-n})} \oint_{E_r} \frac{|f(z)|}{|z-x|} \,|\,\mathrm{d}z| = O(r^{-n}) \text{ as } n \to \infty. \tag{10.27}$$

Applying (10.24), we deduce that the collocation solution $u_n(x)$ converges exponentially to the exact solution $u(x)$ in this case, as the number $n+1$ of terms and the number $n-1$ of collocation points increase.

10.2.3 Projection (tau) methods

Just as collocation methods are seen to be related to approximation by interpolation, so there are methods that are related to approximation by least squares or, more generally, by projection.

- Approximating $u(x)$ by

$$u_n(x) := {\sum_{k=0}^{n}}' c_k T_k(x),$$

as in (10.9), suppose now that we select $n-1$ independent test functions $\{\psi_1(x), \ldots, \psi_{n-1}(x)\}$ and a positive weight function $w(x)$, and solve for the $n+1$ coefficients $\{c_k\}$ the system of $n+1$ linear equations

$$\begin{aligned}
&\int_{-1}^{1} w(x) \left\{ \frac{\mathrm{d}^2}{\mathrm{d}x^2} u_n(x) - f(x) \right\} \psi_\ell(x) \,\mathrm{d}x \\
&= \int_{-1}^{1} w(x) \left\{ {\sum_{k=2}^{n}}' \sum_{\substack{r=0 \\ (k-r) \text{ even}}}^{k-2} (k-r)k(k+r) c_k T_r(x) - f(x) \right\} \psi_\ell(x) \,\mathrm{d}x \\
&\qquad\qquad = 0, \quad \ell = 1, \ldots, n-1, \\
&{\sum_{k=0}^{n}}' (-1)^k c_k = a, \\
&{\sum_{k=0}^{n}}' c_k = b,
\end{aligned} \tag{10.28}$$

so that u_n satisfies the boundary conditions and the residual

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u_n(x) - f(x)$$

is orthogonal to each of the $n-1$ test functions $\psi_1(x), \ldots \psi_{n-1}(x)$ with respect to the weight $w(x)$.

If we take $\psi_\ell(x) = T_{\ell-1}(x)$ and $w(x) = \frac{2}{\pi}(1-x^2)^{-1/2}$, this is equivalent to saying that the residual may be represented in the form

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u_n(x) - f(x) = \sum_{k=n}^{\infty} \tau_{k-1} T_{k-1}(x),$$

for some sequence of undetermined coefficients $\{\tau_k\}$. The method is for this reason often referred to as the *tau method*[1] (Ortiz 1969), although differing slightly from Lanczos's original tau method (see Section 10.3 below), in which the approximation $u_n(x)$ was represented simply as a sum of powers of x

$$u_n(x) = \sum_{k=0}^{n} a_k x^k.$$

In our case, we can use the orthogonality relations (4.9), (4.11) to reduce the first $n-1$ of these equations to

$$\sum_{\substack{k=\ell+2 \\ (k-\ell)\text{ even}}}^{n} (k-\ell)k(k+\ell)c_k = \frac{2}{\pi}\int_{-1}^{1} \frac{T_\ell(x)f(x)}{\sqrt{1-x^2}}\,\mathrm{d}x, \quad \ell = 0, \ldots, n-2. \tag{10.29}$$

The similarities between equations (10.29) and (10.13) are no coincidence. In fact, the right-hand sides of (10.13) are just what we obtain when we apply the Gauss–Chebyshev quadrature rule 1 of Theorem 8.4 (page 183) to the integrals on the right-hand sides of (10.29).

If we use this rule for evaluating the integrals, therefore, we get precisely the same algorithm as in the collocation method; in many contexts we may, however, have a better option of evaluating the integrals more accurately—or even exactly.

- If we use the alternative approximation

$$u_n(x) := (1-x^2)\sum_{k=0}^{n-2} \gamma_k U_k(x) + \frac{1-x}{2}a + \frac{1+x}{2}b,$$

as in (10.14), we are similarly led to the equations

$$\sum_{\substack{k=\ell \\ (k-\ell)\text{ even}}}^{n-2} (\ell^2 - 3k^2 - 6k - 4)\gamma_k = \frac{2}{\pi}\int_{-1}^{1} \frac{T_\ell(x)f(x)}{\sqrt{1-x^2}}\,\mathrm{d}x, \quad \ell = 0, \ldots, n-2, \tag{10.30}$$

and the same final remarks apply.

[1]Compare the tau method for approximating rational functions described in Section 3.6.

EXAMPLE 10.2: Taking same differential equation (10.18) as previously, and now applying the above projection method with $n = 10$, we get the results in Table 10.2, where for convenience we show the values of the exact and approximate solutions at same points as in Table 10.1. It will be seen that the results are slightly more accurate.

Table 10.2: Solution of (10.18) by projection, $n = 10$

x_j	$u(x_j)$	$u_n(x_j)$
0.0000	1.0000	1.0017
±0.3420	0.9600	0.9592
±0.6427	0.7344	0.7347
±0.8660	0.3505	0.3503
±0.9848	0.0449	0.0449

10.2.4 Error of the preceding projection method

We may carry out a backward error analysis just as we did for the collocation method in Section 10.2.2.

As before, let $u(x)$ denote the true solution to the problem (10.3) and $u_n(x)$ the approximate solution, let $f_n(x) = \mathrm{d}^2 u_n(x)/\,\mathrm{d}x^2$ and $\delta f(x) = f(x) - f_n(x)$. Then the error bounds (10.24) still apply.

Assume now that the integrals in (10.29) are evaluated exactly. The function $f_n(x)$ will again be a polynomial of degree $n-2$ but this time, instead of interpolating $f(x)$ at collocation points, it is determined by the integral relations

$$\int_{-1}^{1} \frac{T_\ell(x)\delta f(x)}{\sqrt{1-x^2}}\,\mathrm{d}x = 0, \quad \ell = 0, \ldots, n-2. \tag{10.31}$$

In other words, $f_n(x)$ is a weighted least-squares approximation to $f(x)$ with respect to the weight $w(x) = (1-x^2)^{-1/2}$; that is to say, it is the truncated Chebyshev series expansion

$$f_n(x) = {\sum_{k=0}^{n-2}}{}' d_k T_k(x)$$

of $f(x)$.

Then, for instance, applying the results of Section 5.7, we can say that if the analytic continuation $f(z)$ of $f(x)$ is regular on and within the ellipse E_r then

$$|\delta f(x)| = \left|f(x) - (S_{n-2}^T f)(x)\right| \le \frac{M}{r^{n-2}(r-1)}, \tag{10.32}$$

so that again

$$|\delta f(x)| = O(r^{-n}) \text{ as } n \to \infty, \tag{10.33}$$

and the projection solution converges as n increases at the same rate as the collocation solution based on the zeros of T_{n-1}.

10.3 The original Lanczos tau (τ) method

The original 'tau method' for ordinary differential equations as described by Lanczos (1938) approximated the unknown function by an ordinary polynomial rather than by a truncated Chebyshev expansion — Chebyshev polynomials made their appearance only in the residual. We illustrate this by a simple example.

EXAMPLE 10.3: Consider the equation

$$\frac{\mathrm{d}u}{\mathrm{d}x} + 4xu = 0 \tag{10.34}$$

on the interval $-1 \le x \le 1$, with the condition $u(0) = 1$, to which the solution is

$$u(x) = \mathrm{e}^{-2x^2}.$$

If we try the approximation

$$u_6(x) = 1 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6, \tag{10.35}$$

which satisfies the given condition, we come up with the residual

$$\begin{aligned}\frac{\mathrm{d}u_6}{\mathrm{d}x} + 4xu_6 &= a_1 + (4 + 2a_2)x + (4a_1 + 3a_3)x^2 + (4a_2 + 4a_4)x^3 + \\ &\quad + (4a_3 + 5a_5)x^4 + (4a_4 + 6a_6)x^5 + 4a_5x^6 + 4a_6x^7.\end{aligned}$$

The conventional technique (Frobenius method) for dealing with this residual would be to ignore the last two terms (the highest powers of x), and to equate the remaining terms to zero. This gives

$$a_2 = -2, \quad a_4 = 2, \quad a_6 = -\tfrac{4}{3}, \tag{10.36}$$

with all the odd-order coefficients vanishing.

Lanczos's approach in the same case would have been to equate the residual to

$$\tau_6T_6(x) + \tau_7T_7(x).$$

This gives $\tau_6 = 0$, and the odd-order coefficients again vanishing, while

$$\tau_7 = -\frac{4}{139}, \quad a_2 = -\frac{264}{139}, \quad a_4 = \frac{208}{139}, \quad a_6 = -\frac{64}{139}. \tag{10.37}$$

Thus the conventional approach gives the approximate solution

$$u_6(x) = 1 - 2x^2 + 2x^4 - \tfrac{4}{3}x^6 \tag{10.38}$$

while Lanczos's method gives

$$u_6(x) = 1 - \frac{264x^2 - 208x^4 + 64x^6}{139}. \tag{10.39}$$

The improvement is clear—compare Figures 10.1 and 10.2.

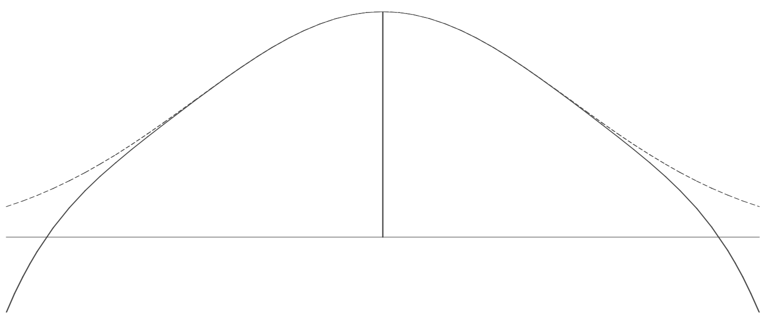


Figure 10.1: Power series solution (10.38) compared with true solution on $[-1, 1]$

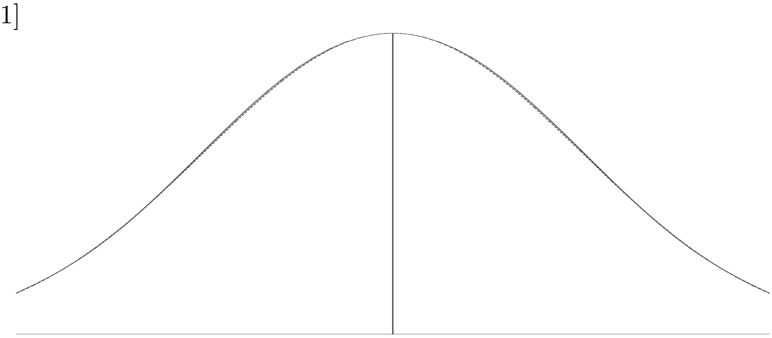


Figure 10.2: Lanczos tau solution (10.39) compared with true solution on $[-1, 1]$

10.4 A more general linear equation

The methods used to attack the simple equation of Section 10.2 may be applied with little alteration to the general linear two-point boundary-value problem

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) + q(x)\frac{\mathrm{d}}{\mathrm{d}x}u(x) + r(x)u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b, \tag{10.40}$$

where $q(x)$ and $r(x)$ are given continuous functions of x.

Approximating $u(x)$ again by the finite sum (10.9)

$$u_n(x) = \sum_{k=0}^{n}{}' c_k T_k(x),$$

using the formula (10.5)

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}T_k(x) = \sum_{\substack{r=0 \\ (k-r)\text{ even}}}^{k-2}{}' (k-r)k(k+r)T_r(x), \ (k \geq 2),$$

for $\frac{\mathrm{d}^2}{\mathrm{d}x^2}T_k(x)$ and the formula (2.49)

$$\frac{\mathrm{d}}{\mathrm{d}x}T_k(x) = \sum_{\substack{r=0 \\ (k-r)\text{ odd}}}^{k-1}{}' 2kT_r(x), \ (k \geq 1)$$

for $\frac{\mathrm{d}}{\mathrm{d}x}T_k(x)$, we get linear equations similar to (10.6) but with the first equation (10.6a) replaced by

$$\begin{aligned} &\sum_{k=2}^{n}\sum_{\substack{r=0 \\ (k-r)\text{ even}}}^{k-2}{}' (k-r)k(k+r)c_k T_r(x) + \\ &\qquad + q(x)\sum_{k=1}^{n}\sum_{\substack{r=0 \\ (k-r)\text{ odd}}}^{k-1}{}' 2kc_k T_r(x) + \\ &\qquad\qquad + r(x)\sum_{k=0}^{n}{}' c_k T_k(x) = f(x). \end{aligned} \tag{10.41}$$

10.4.1 Collocation method

If we substitute $x = x_1, \ldots, x = x_{n-1}$ in (10.41), we again get a system of linear equations for the coefficients $c_0, \ldots, c_n$, which we can go on to solve directly.

If $q(x)$ and $r(x)$ are polynomials in x, and therefore can be expressed as sums of Chebyshev polynomials, we can go on to use the multiplication formula (2.38) or that quoted in Problem 4 of Chapter 2 to reduce the products $q(x)T_r(x)$ and $r(x)T_k(x)$ in (10.41) to simple sums of Chebyshev polynomials $T_k(x)$. We can then use discrete orthogonality as before to simplify the equations to some extent.

Whether this is possible or not, however, collocation methods for linear problems are straightforward to apply.

It should be noted that the error analysis in Section 10.2.2 does not extend to this general case. The reason why it breaks down is that, where previously we could say that $f_n(x)$ was an interpolating polynomial of degree $n-2$, we now have the more complicated expression

$$f_n(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}u_n(x) + q(x)\frac{\mathrm{d}}{\mathrm{d}x}u_n(x) + r(x)u_n(x).$$

We can therefore no longer appeal to formulae for polynomial interpolation error.

10.4.2 Projection method

If $q(x)$ and $r(x)$ are polynomials in x, and we can therefore proceed as in the collocation method to reduce $q(x)T_r(x)$ and $r(x)T_k(x)$ in (10.41) to simple sums of Chebyshev polynomials $T_k(x)$, then we can use integral orthogonality relations (multiplying by $\frac{2}{\pi}(1-x^2)^{-\frac{1}{2}}T_\ell(x)$ and integrating, for $\ell = 0, \ldots, n-2$) to derive a set of linear equations to solve for the coefficients $\{c_k\}$.

In more general circumstances, however, we may need either to approximate $q(x)$ and $r(x)$ by polynomials or to estimate the integrals

$$\int_{-1}^{1} \frac{T_l(x)q(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x, \qquad \int_{-1}^{1} \frac{T_l(x)r(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x$$

numerically.

10.5 Pseudospectral methods — another form of collocation

In the collocation method we discussed earlier in Section 10.2.1, we approximated $u(x)$ by the truncated Chebyshev expansion $u_n(x)$ of (10.9), an nth degree polynomial. Instead of representing this polynomial by its $n+1$ Chebyshev coefficients, suppose now that we represent it by its values at the two boundary points (x_0 and x_n) and at $n-1$ internal collocation points (x_1, …, x_{n-1}); these $n+1$ values are exactly sufficient to define the polynomial uniquely. According to some writers, the coefficients yield a spectral and the values a pseudospectral representation of the polynomial. To make use of such a representation, we need formulae for the derivatives of such a polynomial in terms of these values.

10.5.1 Differentiation matrices

Suppose that we know the values of any nth degree polynomial $p(x)$ at $n+1$ points $x_0,\ldots,x_n$. Then these values determine the polynomial uniquely, and so determine the values of the derivatives $p'(x) = \mathrm{d}p(x)/\,\mathrm{d}x$ at the same $n+1$ points. Each such derivative can, in fact, be expressed as a fixed linear combination of the given function values, and the whole relationship written in matrix form:

$$\begin{pmatrix} p'(x_0) \\ \vdots \\ p'(x_n) \end{pmatrix} = \begin{pmatrix} d_{0,0} & \cdots & d_{0,n} \\ \vdots & \ddots & \vdots \\ d_{n,0} & \cdots & d_{n,n} \end{pmatrix} \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{pmatrix}. \tag{10.42}$$

We shall call $\mathbf{D} = \{d_{j,k}\}$ a *differentiation matrix*.

Suppose now that the points x_j are the $n+1$ zeros of some $(n+1)$st degree polynomial $P_{n+1}(x)$.

If for $k = 0,\ldots,n$ we let $p_k(x) = P_{n+1}(x)/(x - x_k)$, which is an nth degree polynomial since x_k is a zero of P_{n+1}, then we can show without much difficulty that

$$\begin{aligned} p_k(x_k) &= P'_{n+1}(x_k) \\ p_k(x_j) &= 0, \quad j \neq k \\ p'_k(x_k) &= \frac{1}{2} P''_{n+1}(x_k) \\ p'_k(x_j) &= \frac{P'_{n+1}(x_j)}{x_j - x_k}, \quad j \neq k. \end{aligned}$$

From this we can deduce (by setting $p(x) = p_k(x)$ in (10.42)) that the kth column of the differentiation matrix $\mathbf{D}$ must have elements

$$d_{k,k} = \frac{1}{2} \frac{P''_{n+1}(x_k)}{P'_{n+1}(x_k)}, \tag{10.43a}$$

$$d_{j,k} = \frac{P'_{n+1}(x_j)}{(x_j - x_k)P'_{n+1}(x_k)}, \quad j \neq k. \tag{10.43b}$$

Notice that if $p(x_0) = p(x_1) = \cdots = p(x_n) = 1$ then we must have $p(x) \equiv 1$ and $p'(x) \equiv 0$. It follows that each row of the matrix $\mathbf{D}$ must sum to zero, so that the matrix is singular, although it may not be easy to see this directly by looking at its elements.

Not only can we use the relationship

$$\begin{pmatrix} p'(x_0) \\ \vdots \\ p'(x_n) \end{pmatrix} = \mathbf{D} \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{pmatrix}$$

to connect the first derivatives with the function values, but we can repeat the process (since p' is an $(n-1)$st degree polynomial, which may be regarded as an nth degree polynomial with zero leading coefficient), to give us a similar relationship for the second derivatives,

$$\begin{pmatrix} p''(x_0) \\ \vdots \\ p'(x_n) \end{pmatrix} = \mathbf{D}^2 \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{pmatrix},$$

and so on.

10.5.2 Differentiation matrix for Chebyshev points

In particular, suppose that the $n+1$ points are the points $y_k = \cos\frac{k\pi}{n}$, which are the zeros of the polynomial $P_{n+1}(x) = (1-x^2)U_{n-1}(x)$ and the extrema in $[-1,1]$ of $T_n(x)$. Making the usual substitution $x = \cos\theta$ gives us

$$P_{n+1}(x) = \sin\theta\ \sin n\theta.$$

Differentiating with respect to x, we then have

$$P'_{n+1}(x) = -\frac{\cos\theta\ \sin n\theta + n\sin\theta\ \cos n\theta}{\sin\theta} \tag{10.44}$$

and

$$P''_{n+1}(x) = -\frac{\cos^2\theta\ \sin n\theta - n\sin\theta\ \cos\theta\ \cos n\theta + (1+n^2)\sin^2\theta\ \sin n\theta}{\sin^3\theta}. \tag{10.45}$$

However, we know that if $\theta_k = \frac{k\pi}{n}$ then $\sin n\theta_k = 0$ and $\cos n\theta_k = (-1)^k$. Therefore (10.44) gives us

$$P'_{n+1}(y_k) = \begin{cases} -(-1)^k n, & 0 < k < n, \\ -2n, & k = 0, \\ -2(-1)^n n, & k = n, \end{cases} \tag{10.46}$$

and (10.45) gives

$$P''_{n+1}(y_k) = \begin{cases} (-1)^k n\dfrac{y_k}{1-y_k^2}, & 0 < k < n, \\[2ex] -2n\dfrac{1+2n^2}{3}, & k = 0, \\[2ex] 2(-1)^n n\dfrac{1+2n^2}{3}, & k = n. \end{cases} \tag{10.47}$$

(In each case, the values for $k = 0$ and $k = n$ are obtained by proceeding carefully to the limit as $x \to 1$ and $x \to -1$, respectively.)

Substituting (10.46) and (10.47) in (10.43a) and (10.43b) gives the following as elements of the differentiation matrix $\mathbf{D}$:

$$\begin{aligned}
d_{j,k} &= \frac{(-1)^{k-j}}{y_j - y_k}, \quad 0 < j \neq k < n, & d_{k,k} &= -\tfrac{1}{2}\frac{y_k}{1-y_k^2}, \quad 0 < k < n, \\
d_{0,0} &= \tfrac{1}{6}(1+2n^2), & d_{n,n} &= -\tfrac{1}{6}(1+2n^2), \\
d_{0,k} &= 2\frac{(-1)^k}{1-y_k}, \quad 0 < k < n, & d_{k,0} &= -\tfrac{1}{2}\frac{(-1)^k}{1-y_k}, \quad 0 < k < n, \\
d_{k,n} &= \tfrac{1}{2}\frac{(-1)^{n-k}}{1+y_k}, \quad 0 < k < n, & d_{n,k} &= -2\frac{(-1)^{n-k}}{1+y_k}, \quad 0 < k < n, \\
d_{0,n} &= \tfrac{1}{2}(-1)^n, & d_{n,0} &= -\tfrac{1}{2}(-1)^n.
\end{aligned} \tag{10.48}$$

That is to say, we have $\mathbf{D} =$

$$\begin{pmatrix}
\tfrac{1}{6}(1+2n^2) & -2\frac{1}{1-y_1} & 2\frac{1}{1-y_2} & \cdots & 2\frac{(-1)^{n-1}}{1-y_{n-1}} & \tfrac{1}{2}(-1)^n \\
\tfrac{1}{2}\frac{1}{1-y_1} & -\tfrac{1}{2}\frac{y_1}{1-y_1^2} & -\frac{1}{y_1-y_2} & \cdots & \frac{(-1)^{n-2}}{y_1-y_{n-1}} & \tfrac{1}{2}\frac{(-1)^{n-1}}{1+y_1} \\
-\tfrac{1}{2}\frac{1}{1-y_2} & -\frac{1}{y_2-y_1} & -\tfrac{1}{2}\frac{y_2}{1-y_2^2} & \cdots & \frac{(-1)^{n-3}}{y_2-y_{n-1}} & \tfrac{1}{2}\frac{(-1)^{n-2}}{1+y_2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
-\tfrac{1}{2}\frac{(-1)^{n-1}}{1-y_{n-1}} & \frac{(-1)^{n-2}}{y_{n-1}-y_1} & \frac{(-1)^{n-3}}{y_{n-1}-y_2} & \cdots & -\tfrac{1}{2}\frac{y_{n-1}}{1-y_{n-1}^2} & -\tfrac{1}{2}\frac{1}{1+y_{n-1}} \\
-\tfrac{1}{2}(-1)^n & -2\frac{(-1)^{n-1}}{1+y_1} & -2\frac{(-1)^{n-2}}{1+y_2} & \cdots & 2\frac{1}{1+y_{n-1}} & -\tfrac{1}{6}(1+2n^2)
\end{pmatrix}. \tag{10.49}$$

For instance

$n = 1 \ \ (y_k = 1, -1)$

$$\mathbf{D} = \begin{pmatrix} \tfrac{1}{2} & -\tfrac{1}{2} \\ \tfrac{1}{2} & -\tfrac{1}{2} \end{pmatrix}, \quad \mathbf{D}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

$n = 2$ $(y_k = 1,\, 0,\, -1)$

$$\mathbf{D} = \begin{pmatrix} \frac{3}{2} & -2 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 2 & -\frac{3}{2} \end{pmatrix}, \quad \mathbf{D}^2 = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{pmatrix};$$

$n = 3$ $(y_k = 1,\, \frac{1}{2},\, -\frac{1}{2},\, -1)$

$$\mathbf{D} = \begin{pmatrix} \frac{19}{6} & -4 & \frac{4}{3} & -\frac{1}{2} \\ 1 & -\frac{1}{3} & -1 & \frac{1}{3} \\ -\frac{1}{3} & 1 & \frac{1}{3} & -1 \\ \frac{1}{2} & -\frac{4}{3} & 4 & -\frac{19}{6} \end{pmatrix}, \quad \mathbf{D}^2 = \begin{pmatrix} \frac{16}{3} & -\frac{28}{3} & \frac{20}{3} & -\frac{8}{3} \\ \frac{10}{3} & -\frac{16}{3} & \frac{8}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{8}{3} & -\frac{16}{3} & \frac{10}{3} \\ -\frac{8}{3} & \frac{20}{3} & -\frac{28}{3} & \frac{16}{3} \end{pmatrix};$$

$n = 4$ $(y_k = 1,\, 1/\sqrt{2},\, 0,\, -1/\sqrt{2},\, -1)$ (Fornberg 1996, p.164)[2]

$$\mathbf{D} = \begin{pmatrix} \frac{11}{2} & -4-2\sqrt{2} & 2 & -4+2\sqrt{2} & \frac{1}{2} \\ 1+1/\sqrt{2} & -1/\sqrt{2} & -\sqrt{2} & 1/\sqrt{2} & -1+1/\sqrt{2} \\ -\frac{1}{2} & \sqrt{2} & 0 & -\sqrt{2} & \frac{1}{2} \\ 1-1/\sqrt{2} & -1/\sqrt{2} & \sqrt{2} & 1/\sqrt{2} & -1-1/\sqrt{2} \\ -\frac{1}{2} & 4-2\sqrt{2} & -2 & 4+2\sqrt{2} & -\frac{11}{2} \end{pmatrix},$$

$$\mathbf{D}^2 = \begin{pmatrix} 17 & -20-6\sqrt{2} & 18 & -20+6\sqrt{2} & 5 \\ 5+3/\sqrt{2} & -14 & 6 & -2 & 5-3\sqrt{2} \\ -1 & 4 & -6 & 4 & -1 \\ 5-3/\sqrt{2} & -2 & 6 & -14 & 5+3\sqrt{2} \\ 5 & -20+6\sqrt{2} & 18 & -20-6\sqrt{2} & 17 \end{pmatrix}.$$

Notice that the matrices $\mathbf{D}$ and $\mathbf{D}^2$ are singular in each case, as expected.

10.5.3 Collocation using differentiation matrices

We return first to the simple example of (10.3)

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b, \tag{10.50}$$

[2]Fornberg numbers the nodes in the direction of increasing y_k—we have numbered them in order of increasing k and so of decreasing y_k.

and suppose that the collocation points $\{x_j\}$ are chosen so that $x_0 = +1$ and $x_n = -1$ (as will be the case if they are the Chebyshev points $\{y_j\}$ used in Section 10.5.2.)

We know that

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u_n(x_j) = \sum_{k=0}^{n} (\mathbf{D}^2)_{j,k} u_n(x_k).$$

The collocation equations thus become

$$\sum_{k=0}^{n} (\mathbf{D}^2)_{j,k} u_n(x_k) = f(x_j), \ j = 1, \ldots, n-1, \quad u_n(x_n) = a, \quad u_n(x_0) = b. \tag{10.51}$$

Partition the matrices $\mathbf{D}$ and $\mathbf{D}^2$ as follows, by cutting off the first and last rows and columns:

$$\mathbf{D} = \left(\begin{array}{c|c|c} \cdot & \cdots & \cdot \\ \hline \mathbf{e}_0 & \mathbf{E} & \mathbf{e}_n \\ \hline \cdot & \cdots & \cdot \end{array} \right), \mathbf{D}^2 = \left(\begin{array}{c|c|c} \cdot & \cdots & \cdot \\ \hline \mathbf{e}_0^{(2)} & \mathbf{E}^{(2)} & \mathbf{e}_n^{(2)} \\ \hline \cdot & \cdots & \cdot \end{array} \right) \tag{10.52}$$

and let $\mathbf{u}$ and $\mathbf{f}$ denote the vectors

$$\mathbf{u} = \begin{pmatrix} u_n(x_1) \\ \vdots \\ u_n(x_{n-1}) \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_{n-1}) \end{pmatrix}.$$

The collocation equations (10.51) can then be written in matrix notation as

$$u_n(x_0)\mathbf{e}_0^{(2)} + \mathbf{E}^{(2)}\mathbf{u} + u_n(x_n)\mathbf{e}_n^{(2)} = \mathbf{f},$$

or

$$\mathbf{E}^{(2)}\mathbf{u} = \mathbf{f} - b\mathbf{e}_0^{(2)} - a\mathbf{e}_n^{(2)}, \tag{10.53}$$

since the values of $u_n(x_0)$ and $u_n(x_n)$ are given by the boundary conditions. We have only to solve (10.53) to obtain the remaining values of $u_n(x_k)$.

In order to obtain the corresponding Chebyshev coefficients, we can then make use of discrete orthogonality relationships, as in Section 6.3.2.

Similarly, in the case of the more general equation (10.40)

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u(x) + q(x) \frac{\mathrm{d}}{\mathrm{d}x} u(x) + r(x)u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b, \tag{10.54}$$

if we let $\mathbf{Q}$ and $\mathbf{R}$ denote diagonal matrices with elements $q(x_k)$ and $r(x_k)$ $(k = 1, \ldots, n-1)$, the collocation equations can be written as

$$(\mathbf{E}^{(2)} + \mathbf{QE} + \mathbf{R})\mathbf{u} = \mathbf{f} - b(\mathbf{e}_0^{(2)} + \mathbf{Q}\mathbf{e}_0) - a(\mathbf{e}_n^{(2)} + \mathbf{Q}\mathbf{e}_0). \tag{10.55}$$

10.6 Nonlinear equations

We mention briefly that the techniques discussed in the preceding Sections can sometimes be extended to include nonlinear equations.

To take one simple example, using pseudospectral methods and following the principles of (10.51) and (10.52), the problem

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) = f(u(x)), \quad u(-1) = u(+1) = 0, \tag{10.56}$$

where $f(u)$ is an arbitrary function of u, gives rise to the system of equations

$$\sum_{k=0}^{n}(\mathbf{D}^2)_{j,k}u_n(x_k) = f(u_n(x_j)),\ j = 1,\ldots,n-1, \quad u_n(x_n) = u_n(x_0) = 0, \tag{10.57}$$

or, in matrix terms,

$$\mathbf{E}^{(2)}\mathbf{u} = \mathbf{f}(\mathbf{u}), \tag{10.58}$$

where $\mathbf{f}(\mathbf{u})$ denotes the vector with elements $\{f(u_n(x_j))\}$.

Equations (10.58) may or may not have a unique solution. If they do, or if we can identify the solution we require, then we may be able to approach it by an iterative procedure. For instance:

simple iteration Assume that we have a good guess $\mathbf{u}^{(0)}$ at the required solution of (10.58). Then we can generate the iterates $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$, and so on, by solving successive sets of linear equations

$$\mathbf{E}^{(2)}\mathbf{u}^{(k)} = \mathbf{f}(\mathbf{u}^{(k-1)}), \quad k = 1, 2, \ldots. \tag{10.59}$$

Newton iteration Provided that $f(u)$ is differentiable, let $\mathbf{f}'(\mathbf{u})$ denote the diagonal matrix with elements $\{f'(u_n(x_j))\}$, and again assume that we have a good guess $\mathbf{u}^{(0)}$ at the required solution. Then generate successive corrections $(\mathbf{u}^{(1)} - \mathbf{u}^{(0)})$, $(\mathbf{u}^{(2)} - \mathbf{u}^{(1)})$, and so on, by solving successive sets of linear equations

$$\left(\mathbf{E}^{(2)} - \mathbf{f}'(\mathbf{u}^{(k-1)})\right)(\mathbf{u}^{(k)} - \mathbf{u}^{(k-1)}) = \mathbf{f}(\mathbf{u}^{(k-1)}) - \mathbf{E}^{(2)}\mathbf{u}^{(k-1)}, \quad k = 1, 2, \ldots. \tag{10.60}$$

There is no general guarantee that either iteration (10.59) or (10.60) will converge to a solution — this needs to be studied on a case-by-case basis. If both converge, however, then the Newton iteration is generally to be preferred, since its rate of convergence is ultimately quadratic.

10.7 Eigenvalue problems

Similar techniques to those of the preceding sections may be applied to eigenvalue problems in ordinary differential equations. One would not, of course, think of applying them to the simplest such problem

$$\frac{d^2}{dx^2}u(x) + \lambda u(x) = 0, \quad u(\pm 1) = 0, \tag{10.61}$$

since its solutions

$$u(x) = \sin \tfrac{1}{2}k\pi(x+1), \quad \lambda = (\tfrac{1}{2}k\pi)^2, \quad k = 1, 2, \ldots, \tag{10.62}$$

can be written down analytically. This problem nevertheless has its uses as a test case.

Less trivial is the slightly more general problem

$$\frac{d^2}{dx^2}u(x) + q(x)u(x) + \lambda u(x) = 0, \quad u(\pm 1) = 0, \tag{10.63}$$

where $q(x)$ is some prescribed function of x.

10.7.1 Collocation methods

If we approximate $u(x)$ in the form (10.9)

$$u_n(x) := \sum_{k=0}^{n}{}' c_k T_k(x), \tag{10.64}$$

and select the zeros (10.11) of $T_{n-1}(x)$

$$x_j = \cos \frac{(j - \frac{1}{2})\pi}{n-1}, \quad j = 1, \ldots, n-1, \tag{10.65}$$

as collocation points, then the collocation equations for (10.63) become

$$\sum_{k=2}^{n} \sum_{\substack{r=0 \\ (k-r)\text{ even}}}^{k-2}{}' (k-r)k(k+r)c_k T_r(x_j) + $$

$$+ (q(x_j) + \lambda) \sum_{k=0}^{n}{}' c_k T_k(x_j) = 0,$$

$$j = 1, \ldots, n-1, \tag{10.66a}$$

$$\sum_{k=0}^{n}{}' (-1)^k c_k = 0, \tag{10.66b}$$

$$\sum_{k=0}^{n}{}' c_k = 0. \tag{10.66c}$$

Using discrete orthogonality wherever possible, as before, (10.66a) gives us the equations

$$\sum_{\substack{k=\ell+2 \\ (k-\ell)\text{ even}}}^{n} (k-\ell)k(k+\ell)c_k + \frac{2}{n-1}\sum_{j=1}^{n-1}\sum_{k=0}^{n}{}' q(x_j)T_\ell(x_j)T_k(x_j)c_k + \\ + \lambda c_\ell = 0, \\ \ell = 0,\ldots,n-2. \quad (10.67)$$

Using (10.66b) and (10.66c), we can make the substitutions

$$c_n = -\sum_{\substack{k=0 \\ (n-k)\text{ even}}}^{n-2}{}' c_k, \quad (10.68a)$$

$$c_{n-1} = -\sum_{\substack{k=0 \\ (n-k)\text{ odd}}}^{n-3}{}' c_k \quad (10.68b)$$

to reduce (10.67) to the form of a standard matrix eigenvalue equation of the form

$$\mathbf{A}\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-2} \end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-2} \end{pmatrix}, \quad (10.69)$$

which we may solve by standard algebraic techniques.

This will yield $n-1$ matrix eigenvalues $\{\lambda^{(j)}\}$, with corresponding eigenvectors $\{\mathbf{c}^{(j)}\}$. These eigenvalues should approximate the $n-1$ dominant eigenvalues of (10.63), with

$$\sum_{k=0}{}' n c_k^{(j)} T_k(x)$$

approximating the eigenfunction $u^{(j)}(x)$ corresponding to the eigenvalue approximated by $\lambda^{(j)}$. (Equations (10.68) are used to obtain the coefficients $c_n^{(j)}$ and $c_{n-1}^{(j)}$.)

We illustrate this with the trivial example (10.61) in which $q(x) \equiv 0$, when (10.67) becomes

$$\sum_{\substack{k=\ell+2 \\ (k-\ell)\text{ even}}}^{n} (k-\ell)k(k+\ell)c_k + \lambda c_\ell = 0, \\ \ell = 0,\ldots,n-2. \quad (10.70)$$

Table 10.3: Eigenvalues of (10.70) and differential equation (10.61)

Matrix eigenvalues	O.D.E. eigenvalues
2.4674	2.4674
9.8696	9.8696
22.2069	22.2066
39.6873	39.4784
62.5951	61.6850
119.0980	88.8624
178.5341	120.9026
1991.3451	157.9136
3034.1964	199.8595

Results for $n = 10$ are shown in Table 10.3.

Note, incidentally, that if $q(x)$ is an even function of x, so that

$$\sum_{j=1}^{n-1} q(x_j)T_\ell(x_j)T_k(x_j) = 0, \quad k - l \text{ even},$$

we can separate the odd-order coefficients c_k from the even-order ones, thus halving the dimensions of the matrices that we have to deal with.

10.7.2 Collocation using the differentiation matrix

The methods of Section 10.5.3 may be applied equally well to eigenvalue problems. For instance, the problem posed in (10.63)

$$\frac{d^2}{dx^2}u(x) + q(x)u(x) + \lambda u(x) = 0, \quad u(\pm 1) = 0,$$

is the same as (10.54)

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) + q(x)\frac{\mathrm{d}}{\mathrm{d}x}u(x) + r(x)u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b,$$

with $r(x) \equiv \lambda$, $f(x) \equiv 0$ and $a = b = 0$.

Thus equation (10.55) becomes

$$(\mathbf{E}^{(2)} + \mathbf{QE} + \lambda)\mathbf{u} = 0, \tag{10.71}$$

where

$$\mathbf{Q} = \begin{pmatrix} q(x_1) & & \\ & \ddots & \\ & & q(x_{n-1}) \end{pmatrix},$$

where $\mathbf{E}$ and $\mathbf{E}^{(2)}$ are defined as in (10.52) on page 250 and where

$$\mathbf{u} = \begin{pmatrix} u_n(x_1) \\ \vdots \\ u_n(x_{n-1}) \end{pmatrix}.$$

Equation (10.71) is another standard matrix eigenvalue equation.

The test case (10.61) gives the simple eigenvalue problem

$$\mathbf{E}^{(2)}\mathbf{u} + \lambda\mathbf{u} = 0. \tag{10.72}$$

We may use this to illustrate the accuracy of the computed eigenvalues. It will be seen from the last example in Section 10.5.2 that the matrix $\mathbf{E}^{(2)}$ will not always be symmetric, so that it could conceivably have some complex pairs of eigenvalues. However, Gottlieb & Lustman (1983) have shown that its eigenvalues are all real and negative in the case where the collocation points are taken as Chebyshev points (although not in the case where they are equally spaced). Whether the same is true for the matrix $\mathbf{E}^{(2)} + \mathbf{QE}$ depends on the size of the function $q(x)$.

Table 10.4: Eigenvalues of differentiation matrices and differential equation, Chebyshev abscissae

Matrix eigenvalues		O.D.E. eigenvalues
$n = 5$	$n = 10$	
-2.4668	-2.4674	2.4674
-9.6000	-9.8696	9.8696
-31.1332	-22.2060	22.2066
-40.0000	-39.5216	39.4784
	-60.7856	61.6850
	-97.9574	88.8624
	-110.8390	120.9026
	-486.2513	157.9136
	-503.3019	199.8595

We show in Table 10.4 the computed eigenvalues (which are indeed all real and negative) of the $(n-1) \times (n-1)$ matrix $\mathbf{E}^{(2)}$ corresponding to the Chebyshev collocation points $\{y_k\}$, for the cases $n = 5$ and $n = 10$, together with the dominant (smallest) eigenvalues of (10.61). We see that respectively three and five of these eigenvalues are quite closely approximated. For general values of n, in fact, the lowest $\lfloor 2n/\pi \rfloor$ eigenvalues are computed very accurately.

Table 10.5: Eigenvalues of differentiation matrices and differential equation, evenly spaced abscissae

Matrix eigenvalues		O.D.E. eigenvalues
$n = 5$	$n = 10$	
-2.4803	-2.4674	2.4674
-11.1871	-9.8715	9.8696
-15.7488	-22.3049	22.2066
-17.4587	-36.3672	39.4784
	-48.5199	61.6850
	$(-57.6718\ \pm$	88.8624
	$\pm\ 45.9830\ \mathrm{i})$	120.9026
	$(-58.2899\ \pm$	157.9136
	$\pm\ 62.5821\ \mathrm{i})$	199.8595

Table 10.5 displays the corresponding results when the Chebyshev points are replaced by points evenly spaced through the interval $[-1, 1]$. We see that not only are the lower eigenvalues less accurately computed, but higher eigenvalues can even occur in complex pairs.

The same phenomena are illustrated for $n = 40$ by Fornberg (1996, Figure 4.4-2).

10.8 Differential equations in one space and one time dimension

A general discussion of the application of Chebyshev polynomials to the solution of partial differential equations will be found in Chapter 11. A particular class of partial differential equations does, however, fall naturally within the scope of the present chapter—namely equations in two independent variables, the first of which (t, say) represents time and the second (x, say) runs over a finite fixed interval (which as usual we shall take to be the interval $[-1, 1]$). We may then try representing the solution at any fixed instant t of time in terms of Chebyshev polynomials in x.

For a specific example, we may consider the heat conduction equation

$$\frac{\partial}{\partial t}u(t,x) = q(x)\frac{\partial^2}{\partial x^2}u(t,x), \quad t \geq 0, \quad -1 \leq x \leq 1, \tag{10.73a}$$

where $q(x) > 0$, with the boundary conditions

$$u(t,-1) = u(t,+1) = 0, \quad t \geq 0, \tag{10.73b}$$

and the initial conditions

$$u(0,x) = u_0(x), \quad -1 \leq x \leq 1. \tag{10.73c}$$

10.8.1 Collocation methods

We try approximating $u(t,x)$ in the form

$$u_n(t,x) := \sum_{k=0}^{n}{}' c_k(t)T_k(x), \qquad (10.74)$$

and again select the zeros (10.11) of $T_{n-1}(x)$

$$x_j = \cos\frac{(j-\frac{1}{2})\pi}{n-1}, \quad j = 1,\ldots,n-1, \qquad (10.75)$$

as collocation points. The collocation equations for (10.73) become

$$\begin{aligned}
\sum_{k=0}^{n}{}' \frac{\mathrm{d}}{\mathrm{d}t}c_k(t)T_k(x_j) &= q(x_j)\sum_{k=2}^{n}\sum_{\substack{r=0\\(k-r)\text{ even}}}^{k-2}{}'(k-r)k(k+r)c_k(t)T_r(x_j),\\
& \qquad j = 1,\ldots,n-1, && (10.76\text{a})\\
\sum_{k=0}^{n}{}'(-1)^k c_k(t) &= 0, && (10.76\text{b})\\
\sum_{k=0}^{n}{}' c_k(t) &= 0. && (10.76\text{c})
\end{aligned}$$

Using discrete orthogonality as before, equations (10.76a) give

$$\frac{\mathrm{d}}{\mathrm{d}t}c_\ell(t) = \frac{2}{n-1}\sum_{k=2}^{n}\sum_{\substack{r=0\\(k-r)\text{ even}}}^{k-2}{}'(k-r)k(k+r)\sum_{j=1}^{n-1}q(x_j)T_r(x_j)T_\ell(x_j)c_k(t),$$
$$\ell = 0,\ldots,n-2. \qquad (10.77)$$

We again make the substitutions (10.68) for $c_{n-1}(t)$ and $c_n(t)$, giving us a system of linear differential equations for $c_0(t)$, ..., $c_{n-2}(t)$, of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} c_0(t)\\ c_1(t)\\ \vdots\\ c_{n-2}(t)\end{pmatrix} = \mathbf{A}\begin{pmatrix} c_0(t)\\ c_1(t)\\ \vdots\\ c_{n-2}(t)\end{pmatrix}. \qquad (10.78)$$

The matrix $\mathbf{A}$ in (10.78) is not the same as the one appearing in (10.69); however, if we can find its eigenvectors $\{\mathbf{c}^{(j)}\}$ and eigenvalues $\{\lambda^{(j)}\}$, then we can write down the general solution of (10.78) as a linear combination of terms $\{\exp\lambda^{(j)}t\,\mathbf{c}^{(j)}\}$, and hence find the solution corresponding to the given initial conditions (10.73c).

We shall not discuss here the question of how well, if at all, the solution to this system of differential equations approximates the solution to the original partial differential equation (10.73). In particular, we shall not examine the possibility that some $\lambda^{(j)}$ have positive real parts, in which case the approximate solution would diverge exponentially with time and therefore be unstable and completely useless.

10.8.2 Collocation using the differentiation matrix

Once more, we have an alternative approach by way of differentiation matrices. The heat conduction problem (10.73)

$$\frac{\partial}{\partial t}u(t,x) = q(x)\frac{\partial^2}{\partial x^2}u(t,x), \quad u(t,\pm 1) = 0,$$

is another that can be derived from (10.54)

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}u(x) + q(x)\frac{\mathrm{d}}{\mathrm{d}x}u(x) + r(x)u(x) = f(x), \quad u(-1) = a, \quad u(+1) = b,$$

by replacing $u(x)$ and $\frac{\mathrm{d}}{\mathrm{d}x}$ by $u(t,x)$ and $\frac{\partial}{\partial x}$, and setting $q(x) \equiv r(x) \equiv 0$, $a = b = 0$ and

$$f(x) \equiv \frac{1}{q(x)}\frac{\partial}{\partial t}u(t,x).$$

In place of equation (10.55) we thus find the system of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}(t) = \mathbf{Q}\mathbf{E}^{(2)}\mathbf{u}(t), \tag{10.79}$$

where

$$\mathbf{Q} = \begin{pmatrix} q(x_1) & & \\ & \ddots & \\ & & q(x_{n-1}) \end{pmatrix},$$

where $\mathbf{E}^{(2)}$ is defined as in (10.52) on page 250 and where

$$\mathbf{u}(t) = \begin{pmatrix} u_n(t,x_1) \\ \vdots \\ u_n(t,x_{n-1}) \end{pmatrix}.$$

As in the case of (10.78), we may write down the solution of this system of equations as a linear combination of terms $\{\exp\lambda^{(j)}t\,\mathbf{u}^{(j)}\}$, where the matrix $\mathbf{Q}\mathbf{E}^{(2)}$ has eigenvectors $\{\mathbf{u}^{(j)}\}$ and eigenvalues $\{\lambda^{(j)}\}$. In the special case where $q(x)$ is constant, $q(x) \equiv q > 0$ so that $\mathbf{Q}\mathbf{E}^{(2)} = q\mathbf{E}^{(2)}$, we recall from Section 10.7.2 that the eigenvalues will all be real and negative (Gottlieb & Lustman 1983), provided that collocation is at Chebyshev points; consequently all of these terms will decay exponentially with time and the approximate solution is stable.

10.9 Problems for Chapter 10

1. Show that the two collocation algorithms of Section 10.2.1 should lead to exactly the same result for any given value of n—likewise the two projection algorithms of Section 10.2.3.

2. (a) Consider the problem

$$(1-x)y' = 1 \text{ on } [0,1], \quad y(0) = 0.$$

Obtain polynomial approximations $y_n(x) = c_0 + c_1 + \cdots + c_n x^n$ to $y(x)$, of degrees $n = 1, 2, 3$, by including a term $\tau T_n^*(x)$ on the right-hand side of the equation. What is the true solution? Plot the errors in each case.

(b) Consider the slightly modified problem

$$(1-x)y' = 1 \text{ on } [0,\tfrac{3}{4}], \quad y(0) = 0.$$

How do we apply the tau method to this problem? Repeat the exercise of the previous example using $\tau T_n^*(4x/3)$.

(c) Repeat the exercises again for the intervals $[-1,1]$ (using $\tau T_n(x)$) and $[-\frac{3}{4},\frac{3}{4}]$ (using $\tau T_n(4x/3)$). What effect does extending the interval have on the approximate solution and the size of the error?

3. Where necessary changing the independent variable so that the interval becomes $[-1,1]$, formulate and solve a (first-order) differentiation matrix approximation to one of the parts of Problem 2.

4. Obtain a numerical solution to the differential equation (10.18) by the standard finite-difference method

$$(u_{j-1} - 2u_j + u_{j+1})/h^2 + 6\,|x_j| = 0, \quad u_{-n} = u_n = 0,$$

where $h = 1/n$, $x_j = j/n$ ($|j| \le n$) and u_j approximates $u(x_j)$.

For $n = 10$, say, how does the solution compare with the Chebyshev solutions in Tables 10.1 and 10.2?

5. Verify formulae (10.44) and (10.45).

6. Justify the limiting values given in (10.46) and (10.47).

7. Investigate the application of Chebyshev and conventional finite-difference methods to the solution of the differential equation

$$(1+25x^2)^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2} u(x) = 50(75x^2-1)u(x), \quad u(\pm 1) = \tfrac{1}{26},$$

whose exact solution is the function $1/(1+25x^2)$ used to illustrate the Runge phenomenon in Section 6.1.

8. Investigate similarly the non-linear equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u(x) + \frac{1}{1+u(x)^2} = 0, \quad u(\pm 1) = 0.$$

9. Verify the eigenvalues quoted in Tables 10.4, 10.5.

Chapter 11

Chebyshev and Spectral Methods for Partial Differential Equations

11.1 Introduction

Chebyshev polynomial applications to partial differential equations (PDEs)

$$Eu = 0 \text{ on a domain } S, \tag{11.1a}$$

subject to boundary conditions

$$Bu = 0 \text{ on } \partial S, \tag{11.1b}$$

where ∂S is the boundary of the domain S, are a natural progression of the work of Lanczos (1938) and Clenshaw (1957) on ordinary differential equations. However, the first formal publications in the topic of PDEs appear to be those of Elliott (1961), Mason (1965, 1967) and Fox & Parker (1968) in the 1960s, where some of the fundamental ideas for extending one-dimensional techniques to multi-dimensional forms and domains were first developed. Then in the 1970s, Kreiss & Oliger (1972) and Gottlieb & Orszag (1977) led the way to the strong development of so-called pseudo-spectral methods, which exploit the fast Fourier transform of Cooley & Tukey (1965), the intrinsic rapid convergence of Chebyshev methods, and the simplicity of differentiation matrices with nodal bases.

Another important early contribution was the expository paper of Finlayson & Scriven (1966), who set the new methods of the 1960s in the context of the established "method of weighted residuals" (MWR) and classified them formally into the categories of *Galerkin*, *collocation*, and *least squares* methods, as well as into the categories of *boundary*, *interior* and *mixed* methods.

Let us first clarify some of this nomenclature, as well as looking at early and basic approximation methods. We assume that the solution of (11.1a), (11.1b) is to be approximated in the form

$$u \simeq u_n = f(L_n) \tag{11.2}$$

where

$$L_n = \sum_{k=1}^{n} c_k \phi_k \tag{11.3}$$

is a linear combination of an appropriate basis of functions $\{\phi_k\}$ of the independent variables (x and y, say) of the problem and where f is a quasi-linear function

$$f(L) = A.L + B, \tag{11.4}$$

where A, B are specified functions (of x and y).

11.2 Interior, boundary and mixed methods

11.2.1 Interior methods

An *interior method* is one in which the approximation (11.2) exactly satisfies the boundary conditions (11.1b) for all choices of coefficients $\{c_i\}$. This is typically achieved by choosing each basis function ϕ_i appropriately. If Bu in (11.1b) is identically u, so that we have the homogeneous Dirichlet condition

$$u = 0 \text{ on } \partial S, \tag{11.5}$$

then we might well use the identity function for f, and choose a basis for which every ϕ_i vanishes on ∂S. For example, if S is the square domain with boundary

$$\partial S : x = 0,\ x = 1,\ y = 0,\ y = 1. \tag{11.6}$$

then one possibility would be to choose

$$\phi_k = \Phi_{ij} = \sin i\pi x \, \sin j\pi x \tag{11.7}$$

with

$$k = i + n(j - 1)$$

and

$$c_k = a_{ij},$$

say, so that the single index $k = 1, \ldots, n^2$ counts row by row through the array of n^2 basis functions corresponding to the indices $i = 1, \ldots, n$ and $j = 1, \ldots, n$. In practice we might in this case change notation from ϕ_k to Φ_{ij} and from u_{n^2}, L_{n^2} to u_{nn}, L_{nn}, setting

$$u \simeq u_{nn} = f(L_{nn})$$

where

$$L_{nn} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \Phi_{ij}(x, y). \tag{11.8}$$

It only remains to solve the interior problem (11.1a).

There is a generalisation of the above method, that is sometimes applicable to the general Dirichlet boundary conditions

$$u = B(x, y) \tag{11.9}$$

on the boundary

$$\Gamma : A(x, y) = 0,$$

where we know a formula $A = 0$ for the algebraic equation of Γ, as well as a formula $B = 0$ for the boundary data. Then we may choose

$$u \simeq u_{nn} = f(L_{nn}) = A(x, y)L_{nn} + B(x, y), \tag{11.10}$$

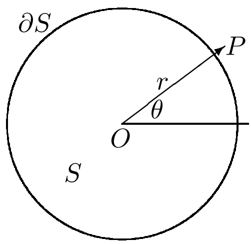


Figure 11.1:

which automatically satisfies (11.5), whatever we take for L_{nn}. See Mason (1967) for a successful application and early discussion of such techniques.

In the discussion that follows we assume unless otherwise stated that f is the identity, so that u_n and L_n are the same function.

11.2.2 Boundary methods

A *boundary method* is one in which the approximation (11.2) exactly satisfies the PDE (11.1a) for all choices of coefficients $\{c_i\}$. If the PDE is linear, for example, then this is achieved by ensuring that every basis function ϕ_k is a particular solution of (11.1a). This method is often termed the "method of particular solutions" and has a long history — see for example Vekua (1967) — and indeed the classical method of separation of variables for PDEs is typically of this nature. It remains to satisfy the boundary conditions approximately by suitable choice of coefficients $\{c_i\}$.

For example, consider Laplace's equation in (r, θ) coordinates:

$$\triangle u = r^2 \frac{\partial^2 u}{\partial r^2} + r \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0 \tag{11.11a}$$

in the disk $S : r \le 1$, together with

$$u = g(\theta) \tag{11.11b}$$

on $\partial S : r = 1$, where g is a known 2π-periodic function of the orientation θ of a general point, P say, on the boundary (Figure 11.1).

Then

$$u \simeq u_n(r, \theta) = {\sum_{k=0}^{n}}' [a_k(r^k \cos(k\theta)) + b_k(r^k \sin(k\theta))] \tag{11.12}$$

is an exact solution of (11.11a) for all $\{a_k, b_k\}$, since $r^k \cos(k\theta)$ and $r^k \sin(k\theta)$ are particular solutions of (11.11a), which may readily be derived by separation of variables in (11.11a) (see Problem 1).

Substituting (11.12) into (11.11b) gives

$$u = g(\theta) \simeq u_n(1,\theta) = \sum_{k=0}^{n}{}' [a_k \cos(k\theta) + b_k \sin(k\theta)]. \tag{11.13}$$

Clearly we require the latter trigonometric sum to approximate $g(\theta)$. This may theoretically be achieved by choosing a_k and b_k to be coefficients in the full Fourier series expansion of $g(\theta)$, namely

$$a_k = \pi^{-1} \int_0^{2\pi} g(\theta)\cos(k\theta)\,\mathrm{d}\theta, \quad b_k = \pi^{-1} \int_0^{2\pi} g(\theta)\sin(k\theta)\,\mathrm{d}\theta. \tag{11.14}$$

These integrals must be replaced by numerical approximations, which may be rapidly computed by the fast Fourier transform (FFT, see Section 4.7). The FFT computes an approximate integral transform, by "exactly" computing the discrete Fourier transform given by

$$a_k = n^{-1} \sum_{i=0}^{2n}{}'' g(\theta_i)\cos(k\theta_i), \quad b_k = n^{-1} \sum_{i=0}^{2n}{}'' g(\theta_i)\sin(k\theta_i), \tag{11.15}$$

where

$$\theta_i = i\pi/n \quad (i = 0, \ldots, 2n). \tag{11.16}$$

Here the periodic formulae (11.14) have been approximated by Filon's rule, namely the Trapezoidal rule for trigonometric functions, which is a very accurate substitute in this case.

Several examples of the method of particular solutions are given by Mason & Weber (1992), where it is shown that the method does not always converge! See also, however, Fox et al. (1967) and Mason (1969) where the "L-shaped membrane eigenvalue problem" is solved very rapidly and accurately by this method.

Boundary MWR methods are important because, when they are applicable, they effectively reduce the dimension of the problem by restricting it to the domain boundary. In consequence such methods can be very efficient indeed. Moreover, because they normally incorporate precise features of the solution behaviour, they are often very accurate too — see Mason (1969) where the first L-shaped membrane eigenvalue is computed correct to 13 significant figures for ($n =$)24 basis functions.

However, boundary MWR methods are not the only available techniques for in effect reducing the problem dimension. The method of fundamental solutions, which has been adopted prolifically by Fairweather & Karageorghis (1998), uses fundamental PDE solutions as a basis. These solutions typically have singularities at their centres, and so must be centred at points exterior to S. This method is closely related to the boundary integral equation (BIE) method and hence to the boundary element method (BEM) — for which there is a huge literature (Brebbia et al. 1984, for example), and indeed the boundary integral equation method adopts the same fundamental solutions, but as weight functions in integral equations. For example, functions behaving like $\log r$ occur in both the method of fundamental solutions and the boundary integral equation method for Laplace's equation in two dimensions.

Both the BIE method and the BEM convert a PDE on a domain into an integral equation over its boundary. They consequently have the possibility for considerable improvements in efficiency and accuracy over classical finite element methods for the original PDE, depending on the nature of the geometry and other factors.

11.2.3 Mixed methods

A *mixed method* is one in which both the PDE (11.1a) and its boundary conditions (11.1b) need to be approximated. In fact this is generally the case, since real-life problems are usually too complicated to be treated as boundary or interior problems alone. Examples of such problems will be given later in this chapter.

11.3 Differentiation matrices and nodal representation

An important development, which follows a contrasting procedure to that of the methods above, is to seek, as initial parameters, not the coefficients c_k in the approximation form L_n (11.3) but instead the values $u_n(x_i, y_j)$ of u_n at a suitable mesh of Chebyshev zeros. Derivatives can be expressed in terms of these u_n values also, and hence a system of (linear) algebraic equations can be formed for the required values of u_n. It is then possible, if required, to recover the coefficients c_k by a Chebyshev collocation procedure.

An example of the procedure was given in Chapter 10 (Section 10.5.1) for ordinary differential equations (ODEs). In the case of PDEs it should be noted that the procedure is primarily suited to rectangular regions.

11.4 Method of weighted residuals

11.4.1 Continuous MWR

The standard MWR, which we call the continuous MWR, seeks to solve an interior problem by finding an approximation of the form (11.2) which minimises, with respect to c_k $(k = 1, \ldots, n)$, the expression

$$\langle Eu_n \, , \, W_k \rangle^2 \equiv \left[\int_S (Eu_n).W_k \, \mathrm{d}S \right]^2 , \tag{11.17}$$

where W_k is a suitable weight function (Finlayson & Scriven 1966). Here we assume that E is a linear partial differential operator. More specifically :

(i) MWR is a *least squares method* if

$$W_k \equiv w.Eu_n, \quad (k = 1, \ldots, n), \tag{11.18}$$

where w is a fixed non-negative weight function. Then, on differentiating (11.17) with respect to c_k, we obtain the simpler form

$$\langle Eu_n\,,\, w.E\phi_k \rangle = 0, \quad (k = 1, \ldots, n). \tag{11.19}$$

This comprises a linear system of n equations for c_k.

(ii) MWR is a *Galerkin method* if

$$W_k \equiv w.\phi_k. \tag{11.20}$$

Note that, in this case, we can give a zero (minimum) value to (11.17) by setting

$$\langle Eu_n\,,\, w.\phi_k \rangle = 0, \quad (k = 1, \ldots, n), \tag{11.21}$$

again a system of linear equations for c_k. It follows from (11.21) that

$$\langle Eu_n\,,\, w.u_n \rangle = 0. \tag{11.22}$$

More generally, we can if we wish replace ϕ_k in (11.21) by any set of test functions ψ_k, forming a basis for u_k and solve

$$\langle Eu_n\,,\, w.\psi_k \rangle = 0, \quad (k = 1, \ldots, n). \tag{11.23}$$

(iii) MWR is a *collocation method (interpolation method)* at the points $P_1, \ldots, P_n$ if

$$W_k \equiv \delta(P_k), \tag{11.24}$$

where $\delta(P)$ is the Dirac delta function (which is infinite at the point P, vanishes elsewhere and has the property that $\langle u\,,\, \delta(P) \rangle = u(P)$ for any well-behaved function u). Then Eu_n in (11.17) will be set to zero at P_k, for every k.

11.4.2 Discrete MWR — a new nomenclature

It is also possible to define a discrete MWR, for each of the three types of methods listed above, by using a discrete inner product in (11.17). Commonly we do not wish, or are unable, to evaluate and integrate $Eu_n.W_k$ over a continuum, in which case we may replace the integral in (11.17) by the sum

$$\sum_{S_n} (Eu_n).W_k, \tag{11.25}$$

where S_n is a discrete point set representing S.

The discrete MWR, applied to an interior problem, is based on a discrete inner product. It seeks an approximation of the form (11.2) which solves

$$\min_{c_k} \left[\left(\sum_{j=1}^{p} Eu_n(\mathbf{x}_j) W_k(\mathbf{x}_j) \right)^2 \equiv (Eu_n, W_k)^2 \right], \tag{11.26}$$

where $\mathbf{x}_j$ $(j = 1, \ldots, p)$ are a discrete set of nodes in S, selected suitably from values of the vector $\mathbf{x}$ of independent variables, and W_k are appropriate weights.

(i) The discrete MWR is a *discrete least-squares method* if

$$W_k \equiv wEu_n. \tag{11.27}$$

This is commonly adopted in practice in place of (11.18) for convenience and to avoid integration.

(ii) The discrete MWR is a *discrete Galerkin method* if

$$W_k \equiv w\phi_k \tag{11.28}$$

or, equivalently,

$$(Eu_n, w\psi_k) = 0. \tag{11.29}$$

Note that the PDE operator Eu_n is directly orthogonal to every test function ψ_k, as well as to the approximation u_n, so that

$$(Eu_n, wu_n) = 0. \tag{11.30}$$

(iii) The discrete MWR is a *discrete collocation method* if (11.24) holds, where $\{P_k\}$ is contained within the discrete point set S_n.

11.5 Chebyshev series and Galerkin methods

The most basic idea in Chebyshev polynomial methods is that of expanding a solution in a (multiple) Chebyshev series expansion, and using the partial sum as an approximation. This type of approach is referred to as a *spectral method* by Gottlieb & Orszag (1977). This type of ODE/PDE method had previously, and still has, several other names, and it is known as (or is equivalent to) a Chebyshev series method, a Chebyshev–Galerkin method, and the tau method of Lanczos.

Before introducing PDE methods, we consider the Lanczos tau method: one of the earliest Chebyshev methods for solving a linear ODE

$$Ey = 0$$

in the approximate form y_n.

Lanczos (1938) and Ortiz and co-workers (Ortiz 1969, Freilich & Ortiz 1982, and many other papers) observed that, if y_n is expressed in the power form

$$y_n = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n, \tag{11.31}$$

then, for many important linear ODEs, Ey_n can be equated to a (finite) polynomial with relatively few terms, of the form

$$Ey_n = \tau_1 T_{q+1}(x) + \tau_2 T_{q+2}(x) + \cdots + \tau_s T_{q+s}(x), \tag{11.32}$$

where q and s are some integers dependent on E. The method involves substituting y_n (11.31) into the perturbed equation (11.32) and equating powers of x from x^0 to x^{q+s}. The t (say) boundary conditions are also applied to y_n, leading to a total of $q+s+t+1$ linear equations for $b_0, \ldots, b_n$, $\tau_1, \ldots, \tau_s$. We see that for the equations to have one and only one solution we must normally have

$$q + t = n. \tag{11.33}$$

The equations are solved by first expressing $b_0, \ldots, b_n$ in terms of $\tau_1, \ldots, \tau_s$, solving s equations for the τ values and hence determining the b values. Because of the structure of the resulting matrix and assuming s is small compared to n, the calculation can routinely reduce to one of $O(n)$ operations, and hence the method is an attractive one for suitable equations.

The above method is called the (Lanczos) *tau method* - with reference to the introduction by Lanczos (1938) of perturbation terms, with coefficients $\tau_1, \ldots, \tau_s$, on the right hand side of $Ey = 0$ to enable the ODE to be exactly solved in finite form. The nice feature of this approach is that the tau values give a measure of the sizes of the contributions that the perturbation terms make to the ODE — at worst,

$$|Ey_n| \le |\tau_1| + |\tau_2| + \cdots + |\tau_s|\,. \tag{11.34}$$

For some special cases, Lanczos (1957), Fox & Parker (1968), Mason (1965), Ortiz (1980, 1986, 1987), Khajah & Ortiz (1991) and many others were able to give quite useful error estimates based on the known form (11.32).

The tau method is also equivalent to a Galerkin method, since Ey_n is orthogonal with respect to $(1-x^2)^{-1/2}$ to all polynomials of degree up to q, as a consequence of (11.32). Note that the Galerkin method proper is more robust than equivalent tau or Chebyshev series methods, since, for example, it is unnecessary to introduce τ terms or to find and use the form (11.32). The Galerkin method directly sets up a linear system of equations for its coefficients. For example, if we wish to solve

$$u' - u = 0, \quad u(0) = 1 \tag{11.35}$$

by a Galerkin procedure using Legendre polynomials $P_i^*(x)$ appropriate to $[0,1]$, namely

$$u \sim u_n = \sum_{i=0}^{n} c_i P_i^*(x), \tag{11.36}$$

then we solve

$$\int_0^1 (u_n' - u_n).P_i^*(x)\,\mathrm{d}x = 0 \quad (i = 0, 1, \ldots, n-1) \tag{11.37}$$

and

$$\sum_{i=0}^{n} c_i P_i^*(0) = 1. \tag{11.38}$$

Here (11.37) and (11.38) comprise $n+1$ equations for $c_0, \ldots, c_n$. Note that a snag in the Galerkin method is the need to evaluate the various integrals that occur, which are likely to require a numerical treatment except in simple problems such as (11.35).

It is worth remembering that Chebyshev series are also transformed Fourier series, and so Chebyshev methods may be based on known methods for generating Fourier partial sums or Fourier transforms, based on integrals and expansions.

11.6 Collocation/interpolation and related methods

We have seen, in Sections 5.5 and 6.5, that a Chebyshev series partial sum of degree n of a continuous function is a near-minimax approximation on $[-1,1]$ within a relative distance of order $4\pi^{-2}\log n$, whereas the polynomial of degree n interpolating (collocating) the function at the $n+1$ zeros of $T_{n+1}(x)$ is near-minimax within a slightly larger relative distance of order $2\pi^{-1}\log n$. Thus, we may anticipate an error that is $\pi/2$ times as large in Chebyshev interpolation compared with Chebyshev series expansion. In practice, however, this is a very small potential factor, and polynomial approximations from the two approaches are virtually indistinguishable. Indeed, since collocation methods are simpler, more flexible and much more generally applicable, they offer a powerful substitute for the somewhat more mathematically orthodox but restrictive series methods.

The title *pseudo-spectral method* was introduced by Gottlieb & Orszag (1977), in place of *Chebyshev collocation method*, to put across the role of this method as a robust substitute for the spectral method. Both series (spectral) and collocation (pseudo-spectral) methods were rigorously described by Mason (1970) as near-minimax. Note that minimax procedures generally involve infinite procedures and are not practicably feasible, while spectral, and more particularly pseudo-spectral, methods are typically linear and very close to minimax and therefore provide an excellent and relatively very inexpensive substitute for a minimax approximation method.

It has long been realised that collocation for differential equations is almost identical to series expansion. Lanczos (1957) noted that the ODE error form adopted in his tau method (11.32) could conveniently be replaced with nearly identical results (though different τ coefficients) by

$$Ey_n = T_{q+1}(x).(\tau_1 + \tau_2 x + \cdots + \tau_s x^{s-1}), \tag{11.39}$$

where $q + s$ is the degree of Ey_n. Note that the error in the ODE vanishes at the zeros of $T_{q+1}(x)$, and so the method is equivalent to a collocation method (in the ODE). Lanczos called this method the *selected points method*, where the zeros of T_{q+1} are the points selected in this case. Lanczos sometimes also selected Legendre polynomial zeros instead, since in practice they sometimes give superior results.

We have already shown that the Chebyshev collocation polynomial, $f_n(x)$ of degree n to a given $f(x)$, may be very efficiently computed by adopting a discrete orthogonalisation procedure

$$f(x) \simeq f_n(x) = \sideset{}{'}\sum_{i=0}^{n} c_i T_i(x), \tag{11.40}$$

where

$$c_i = \frac{2}{N}\sum_{k=0}^{N} f(x_k)T_i(x_k) = \frac{2}{N}\sum_{k=0}^{N} f(\cos(\theta_k))\cos(i\theta_k), \tag{11.41}$$

with

$$x_k = \cos(\theta_k) = \cos\left(\frac{(2k+1)\pi}{2(N+1)}\right) \quad (k = 0, 1, \ldots, n). \tag{11.42}$$

For $N = n$, this yields the collocation polynomial, and this clearly mimics the Chebyshev series partial sum of order n, which has the form (11.40) with (11.41) replaced by

$$c_i = \frac{2}{\pi}\int_{-1}^{1}(1-x^2)^{-1/2} f(x)T_i(x)\,\mathrm{d}x = \frac{2}{\pi}\int_{0}^{\pi} f(\cos(\theta))\cos(i\theta)\,\mathrm{d}\theta. \tag{11.43}$$

with $x = \cos(\theta)$.

Note that the discrete Chebyshev transform in x and the discrete Fourier transform in θ, that appear in (11.41), also represent an excellent numerical method (Filon's rule for periodic integrands) for approximately computing the continuous Chebyshev transform and Fourier transform that appear in (11.43). The fast Fourier transform (FFT), which is of course a very efficient method of computing the Fourier transform, does in fact compute the discrete

Fourier transform instead. However, (11.43) is typically replaced by (11.41) for a value of N very much larger than n, say $N = 1024$ for $n = 10$. So there are really two different discrete Fourier transforms, one for $N = n$ (collocation) and one for $N \gg n$ (approximate series expansion).

11.7 PDE methods

We note that, for simplicity, the above discussions of nomenclature have been based on ODEs, for which boundary conditions apply at just a few points, usually only one or two. Typically these boundary conditions are imposed exactly as additional constraints on the approximation, with only a little effect on the number of coefficients remaining. For example, in the tau method for

$$Eu \equiv u' - u = 0, \quad u(0) = 1 \quad \text{in } [0, 1], \tag{11.44}$$

we determine

$$u \sim u_n = c_0 + c_1 x + \cdots + c_n x^n$$

by equating coefficients of $1, x, x^2, \ldots, x^n$ in

$$Eu_n \equiv (c_1 - c_0) + (2c_2 - c_1)x + (3c_3 - c_2)x^2 + \cdots + \\ + (nc_n - c_{n-1})x^{n-1} - c_n x^n = \tau T_n^*(x). \tag{11.45}$$

This yields $n+1$ linear equations for $c_0, \ldots, c_n, \tau$, and an additional equation is obtained by setting $c_0 = 1$ to satisfy the boundary (initial) condition.

In spectral and pseudo-spectral methods for PDEs, the boundary conditions play a much more crucial role than for ODEs, and it becomes important to decide whether to satisfy the boundary conditions implicitly, in the form chosen for the basis functions, or to apply the boundary conditions as additional constraints. For this reason, Gottlieb & Orszag (1977) and Canuto et al. (1988) differentiate between Galerkin and tau methods primarily in terms of their treatment of boundary conditions — whereas we have above viewed these methods as equivalent, one based on the orthogonality requirement and the other based on the form of the ODE (perturbation) error. Canuto et al. (1988) view a Galerkin method as a series method in which the boundary conditions are included implicitly in the chosen solution form, whereas a tau method is seen by them as a series method for which the boundary conditions are applied explicitly through additional constraints.

The distinction made by Canuto et al. (1988) between Galerkin and tau methods has virtues. In particular the number of free approximation coefficients needed to satisfy boundary conditions can be very large, whereas this may be a peripheral effect if the boundary can be treated implicitly. So a distinction is needed. But the words, Galerkin and tau, do not conjure up boundary issues, but rather an orthogonality technique and tau perturbation terms. A better terminology, we would suggest, would be to refer to methods which include/exclude boundary conditions from the approximation form as implicit/explicit methods respectively. We could alternatively use the titles interior/mixed methods, as discussed for the MWR above.

Nomenclature and methods become more complicated for PDEs in higher dimensions. In the following sections we therefore give a number of examples of problems and methods to illustrate the formalisms that result from approaches of the Galerkin, tau, collocation, implicit, explicit (&c.) variety. We do not view spectral and pseudo-spectral methods, unlike Galerkin and tau methods, as specifically definable methods, but rather as generic titles for the two main branches of methods (series and collocation). A generalisation of the Lanczos tau method might thus be termed a spectral explicit/mixed tau method.

11.7.1 Error analysis

Canuto et al. (1988) and Mercier (1989), among others, give careful attention to error bounds and convergence results. In particular, Canuto et al. (1988) address standard problems such as the Poisson problem, as well as individually addressing a variety of harder problems. In practice, however, the main advantage of a spectral method lies in the rapid convergence of the Chebyshev series; this in many cases makes feasible an error estimate based on the sizes of Chebyshev coefficients, especially where convergence is exponential.

11.8 Some PDE problems and various methods

It is simplest to understand, develop and describe spectral and pseudo-spectral methods by working through a selection of problems of progressively increasing complexity. This corresponds quite closely to the historical order of development, which starts, from a numerical analysis perspective, with the novel contributions of the 1960s and is followed by the fast (FFT-based) spectral methods of the 1970s. Early work of the 1960s did establish fundamental techniques and compute novel approximations to parabolic and elliptic PDEs, based especially on the respective forms

$$u(x,t) \sim u_n(x,t) = \sum_{i=0}^{n} c_i f_i(t) T_i(x) \quad (-1 \le x \le 1;\ t \ge 0) \tag{11.46}$$

for parabolic equations, such as $u_{xx} = u_t$, and

$$u(x,y) \sim u_n(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{ij} T_i(x) T_j(y) \quad (-1 \le x, y \le 1) \tag{11.47}$$

for elliptic problems, such as $\Delta u \equiv u_{xx} + u_{yy} = f$.

An early paper based on (11.46) was that of Elliott (1961), who determined $f_i(t)$ as approximate solutions of a system of ODEs, in the spirit of the "method of lines". Another early paper based on (11.47) was that of Mason (1967), which solves the membrane eigenvalue problem (see Section 11.8.2 below)

$$\Delta u + \lambda u = 0 \text{ in } S, \quad u = 0 \text{ on } \partial S, \tag{11.48}$$

for the classical problem of an L-shaped membrane (consisting of three squares co-joined), based on a preliminary conformal mapping of the domain and an approximation

$$u \simeq A(x,y).\phi_n(x,y), \tag{11.49}$$

where $A = 0$ is the algebraic equation of the mapped boundary. Mason (1969) also used an approximation of form (11.46) to solve a range of separable PDEs including (11.48). Indeed the leading eigenvalue of the L-membrane was computed to 13 significant figures by Mason (1969) ($\lambda = 9.639723844022$).

These early quoted papers are all essentially based on the collocation method for computing coefficients c_i or c_{ij}. It is also possible to develop tau/series methods for the form (11.46), based on the solution by the Lanczos tau method of the corresponding ODEs for $f_i(t)$; this has been carried out for very basic equations such as the heat equation and Poisson equation (Berzins & Dew 1987).

11.8.1 Power basis: collocation for Poisson problem

Consider the Poisson problem

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y) \text{ in } S, \quad u = 0 \text{ on } \partial S, \tag{11.50}$$

where S is the square with boundaries $x = \pm 1, y = \pm 1$. Suppose we approximate as

$$u \simeq u_{mn} = \phi(x,y) \sum_{i=0}^{m-2} \sum_{j=0}^{n-2} a_{ij} x^i y^j, \tag{11.51}$$

where we adopt the power basis $x^i y^j$ and include a multiplicative factor $\phi(x)$ such that $\phi = 0$ is the (combined) equation of the boundaries. In this case,

$$\phi(x,y) = (x^2 - 1)(y^2 - 1). \tag{11.52}$$

Then we may rewrite u_{mn} as

$$u_{mn} = \sum_{i=0}^{m-2} \sum_{j=0}^{n-2} a_{ij} (x^{i+2} - x^i)(y^{j+2} - y^j) \tag{11.53}$$

and hence, applying Δ, obtain

$$\begin{aligned}\Delta u_{mn} &= \sum_{i=0}^{m-2}\sum_{j=0}^{n-2} a_{ij}\left([(i+2)(i+1)x^2 - i(i-1)]x^{i-2}y^j(y^2-1)\right.\\ &\qquad \left. + [(j+2)(j+1)y^2 - j(j-1)]x^i(x^2-1)y^{j-2}\right). \end{aligned} \tag{11.54}$$

Now set Δu_{mn} equal to $f(x,y)$ at the $(m-1)(n-1)$ points (x_k, y_l) ($k = 1, \ldots, m-1$; $l = 1, \ldots, n-1$), where $\{x_k\}$, $\{y_l\}$ are the respective sets of zeros of $T_{m-1}(x)$, $T_{n-1}(y)$, respectively, namely the points

$$x_k = \cos\left(\frac{(2k-1)\pi}{2(m-1)}\right), \quad y_l = \cos\left(\frac{(2l-1)\pi}{2(n-1)}\right). \tag{11.55}$$

This leads to a full linear algebraic system of $(m-1)(n-1)$ equations for a_{ij}. It is relatively straightforward to code a computer procedure for the above algorithm.

We also observe that the approximation u_{mn} adopted in (11.53) above could equally well be replaced by the equivalent form

$$u_{mn} = \sum_{i=2}^{m}\sum_{j=2}^{n} a_{ij}(x^i - x^{i \bmod 2})(y^j - y^{j \bmod 2}), \tag{11.56}$$

where (i mod 2) is 0 or 1 according as i is even or odd, since x^2-1 and y^2-1 are in every case factors of u_{mn}. This leads to a simplification in Δu_{mn} (as in (11.54)), namely

$$\begin{aligned}\Delta u_{mn} &= \sum_{i=2}^{m}\sum_{j=2}^{n} a_{ij}\left[i(i-1)x^{i-2}(y^j - y^{j \bmod 2})\right.\\ &\qquad \left. + j(j-1)y^{j-2}(x^i - x^{i \bmod 2})\right]. \end{aligned} \tag{11.57}$$

The method then proceeds as before. However, we note that (11.51) is a more robust form for more general boundary shapes ∂S and more general boundary conditions $Bu = 0$, since simplifications like (11.56) are not generally feasible.

The above methods, although rather simple, are not very efficient, since no account has been taken of special properties of Chebyshev polynomials, such as discrete orthogonality. Moreover, (11.53) and (11.56) use the basis of power functions $x^i y^j$ which, for m and n sufficiently large, can lead to significant loss of accuracy in the coefficients a_{ij}, due to rounding error and poor conditioning in the resulting linear algebraic system. We therefore plan to follow up this discussion in a later Section by considering more efficient and well conditioned procedures based on the direct use of a Chebyshev polynomial product as a basis, namely $T_i(x)T_j(y)$.

However, before we return to the Poisson problem, let us consider a more difficult problem, where the form (11.51) is very effective and where a power basis is adequate for achieving relatively high accuracy.

11.8.2 Power basis: interior collocation for the L-membrane

Consider the eigenvalue problem

$$\Delta u + \lambda u = 0 \text{ in } S, \quad u = 0 \text{ on } \partial S, \tag{11.58}$$

where S is the L-shaped region shown (upside down for convenience) in Figure 11.2. It comprises three squares of unit sides placed together. To remove the re-entrant corner at O, we perform the mapping, adopted by Reid & Walsh (1965),

$$z' = z^{2/3} \quad (z' = x' + \mathrm{i}y', \quad z = x + \mathrm{i}y), \tag{11.59}$$

where x, y are coordinates in the original domain S (Figure 11.2) and x', y' are corresponding coordinates in the mapped domain S' (Figure 11.3).

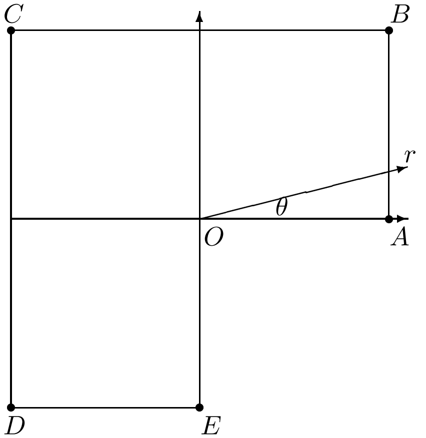


Figure 11.2: L-membrane

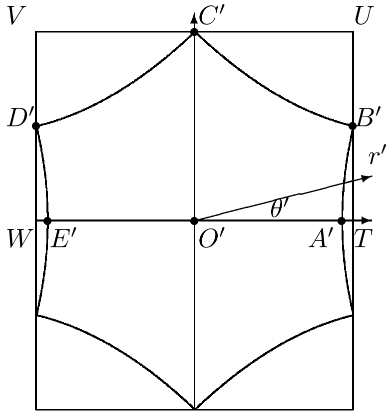


Figure 11.3: Mapped domain

Note that the domain S' is the upper half of a curved hexagon with corners of angle $\frac{\pi}{2}$ shown in Figure 11.3, where the vertices A', B', C', D', E' correspond to A, B, C, D, E in Figure 11.2. (The lower half of the hexagon does not concern us, but is included in Figure 11.3 to show the geometry and the symmetry.) Now, from the mapping,

$$r' = r^{2/3}, \quad \theta' = \tfrac{2}{3}\theta. \tag{11.60}$$

Then

$$\begin{aligned}
\Delta u + \lambda u &\equiv r^{-2}\left[r^2\frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} + \lambda u\right] \\
&= r^{-2}\left[r\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial \theta^2}\right] + \lambda u \\
&= (r')^{-3}\left[\tfrac{2}{3}\left(r'\frac{\partial}{\partial r'}\right)\tfrac{2}{3}\left(r'\frac{\partial u}{\partial r'}\right) + \tfrac{4}{9}\frac{\partial^2 u}{\partial \theta'^2}\right] + \lambda u. \qquad (11.61)
\end{aligned}$$

Hence

$$\Delta u + \lambda u = \tfrac{4}{9}(r')^{-1}.(\Delta' u + \tfrac{9}{4}r'\lambda u) = 0, \tag{11.62}$$

where dashes on Δ', r' indicate that dashed coordinates are involved. Thus the problem (11.58) has transformed, now dropping dashes on r, u, to

$$\Delta u + \tfrac{9}{4}r\lambda u = 0 \text{ in } S', \quad u = 0 \text{ on } \partial S'. \tag{11.63}$$

Before proposing a numerical method, we need to find the algebraic equation of the boundary $O'A'B'C'D'E'(O')$ in Figure 11.3. This boundary has two parts: the straight line $E'O'A'$, namely $y' = 0$, and the set of four curves $A'B'C'D'E'$ which correspond to $(x^2-1)(y^2-1) = 0$ in S. Thus the boundary equation is

$$\begin{aligned} 0 &= A(x,y) = 4y'(x^2-1)(y^2-1) = 4y'(x^2y^2 - r^2 + 1) \\ &= y'(r^4\sin^2(2\theta) - 4r^2 + 4) = y'\left[(r')^6\sin^2(3\theta') - 4(r')^3 + 4\right] \\ &= y'\left[(y')^2\{3(x')^2 - (y')^2\}^2 - 4\{(x')^2 + (y')^2\}^{3/2} + 4\right]. \end{aligned} \tag{11.64}$$

Dropping dashes again,

$$A(x,y) = y.\left[y^2(3x^2 - y^2)^2 - 4(x^2+y^2)^{3/2} + 4\right] = 0. \tag{11.65}$$

We now adopt as our approximation to u, using (11.65) for ϕ:

$$u \simeq u_{mn} = A(x,y).\sum_{i=0}^{m}\sum_{j=0}^{n} c_{ij}x^{2i+t}y^j, \tag{11.66}$$

where $t = 0$ or 1, according as we seek a solution which is symmetric or anti-symmetric about OC. For the leading (i.e., largest) λ, we choose $t = 0$.

The rectangle $TUVW$, which encloses the mapped membrane, has sides $O'T$, TU in the x, y directions, respectively, of lengths a, b, say, given by

$$\begin{aligned} a &= O'T = O'B'\cos(\pi/6) = 2^{1/3}3^{1/2}/2 = 2^{-2/3}3^{1/2}, \\ b &= TU = O'C' = (2^{1/2})^{2/3} = 2^{1/3}. \end{aligned} \tag{11.67}$$

An appropriate 'interior' collocation method is now simply constructed. We specify that the form of approximation (11.66) should satisfy the PDE (11.63) at the tensor product of $(m+1)(n+1)$ positive zeros of $T^*_{m+1}(x/a)T^*_{n+1}(y/b)$, where a, b are given in (11.67), namely the points

$$\{x,y\} = \left\{a.\cos^2\left(\frac{(2k-1)\pi}{4(m+1)}\right), \quad b.\cos^2\left(\frac{(2l-1)\pi}{4(n+1)}\right)\right\}$$
$$(k = 1,\ldots,m+1;\ l = 1,\ldots,n+1). \tag{11.68}$$

Table 11.1: Estimates of first 3 eigenvalues of L-membrane

$m = n$	λ	Rayleigh quotient
Functional form $A(x, y) \sum_0^m \sum_0^n x^{2i} y^j$		
4	9.6398	9.6723
6	9.6400	9.6397
8	9.6397	9.6397
Functional form $xA(x, y) \sum_0^m \sum_0^n x^{2i} y^{2j}$		
4	15.2159	
5	15.1978	15.1980
6	15.1974	15.1974
7	15.1974	15.1974
Functional form $A(x, y) \sum_0^m \sum_0^n x^{2i} y^{2j}$		
4	19.8054	
5	19.7394	
6	19.7392	19.7392
7	19.7392	19.7392

This leads to a homogeneous system of $(m+1)(n+1)$ linear equations, which we may write in matrix form as $\mathbf{A}.\mathbf{c} = 0$, for the determination of $\mathbf{c} = \{c_{ij}\}$, where $\mathbf{A}$ depends on λ. The determinant of $\mathbf{A}$ must vanish, thus defining eligible values of λ, corresponding to eigenvalues of the PDE. We have applied the secant method to find the λ nearest to a chosen guess. Results for the first three eigenvalues, taken from Mason (1965), are shown in Table 11.1 together with Rayleigh quotient estimates. Clearly the method is rather successful, and the application serves as an interesting model problem.

Strictly speaking, the collocation method above is not an interior method, since some collocation points are exterior to S although interior to the rectangle $TUVW$. However, the PDE solution does extend continuously across the problem boundaries to these exterior points.

In fairness we should point out that, although the above collocation method is probably at least as effective for this problem as the best finite difference method, such as that of Reid & Walsh (1965), it is not the best method of all. A better method for the present problem is the boundary method, based on separation of variables, due to Fox et al. (1967) and further extended by Mason (1969). This breaks down for regions with more than one re-entrant corner, however, on account of ill-conditioning; a better method is the finite-element/domain-decomposition method described by Driscoll (1997).

11.8.3 Chebyshev basis and discrete orthogonalisation

In the remaining discussion, we concentrate on the use of a Chebyshev polynomial basis for approximation and exploit properties such as discrete orthogonality and the FFT for efficiency. However, it is first appropriate to remind the reader that the classical method of separation of variables provides both a fast boundary method for Laplace's equation and a superposition method, combining interior and boundary methods for the Poisson equation with non-zero boundary conditions.

Separation of variables: basic Dirichlet problem

Consider the basic Dirichlet problem for Laplace's equation on a square, namely

$$\Delta u = 0 \text{ in } S, \tag{11.69a}$$

$$u = g(x, y) \text{ on } \partial S, \tag{11.69b}$$

where ∂S is the square boundary formed by $x = \pm 1, y = \pm 1$, S is its interior and g is defined only on ∂S. Then we may solve (11.69a) analytically for the partial boundary conditions

$$u = g(-1, y) \text{ on } x = -1; \quad u = 0 \text{ on } x = +1, y = -1, y = +1, \tag{11.70}$$

in the form

$$u = \sum_{k=1}^{\infty} a_k \sinh \tfrac{1}{2}k\pi(1 - x) \, \sin \tfrac{1}{2}k\pi(1 - y), \tag{11.71}$$

where a_k are chosen to match the Fourier sine series of $g(-1, y)$ on $x = -1$. Specifically

$$g(-1, y) = \sum_{k=1}^{\infty} b_k \sin \tfrac{1}{2}k\pi(1 - y), \tag{11.72}$$

where

$$b_k = 2 \int_{-1}^{1} g(-1, y) \, \sin \tfrac{1}{2}k\pi(1 - y) \, \mathrm{d}y, \tag{11.73}$$

and hence a_k is given by

$$a_k = b_k [\sinh k\pi]^{-1}. \tag{11.74}$$

Clearly we can define three more solutions of (11.69a), analogous to (11.71), each of which matches $g(x, y)$ on one side of ∂S and is zero on the remainder of ∂S. If we sum these four solutions then we obtain the analytical solution of (11.69a) and (11.69b). For an efficient numerical solution, the Fourier series should be truncated and evaluated by using the FFT [see Section 4.7].

Chebyshev basis: Poisson problem

The Poisson Problem can be posed in a slightly more general way than in Section 11.8.3, while still permitting efficient treatment. In particular we may introduce two general functions, f as the right-hand side of the PDE, and g as the boundary function, as follows.

$$\Delta u = f(x,y) \text{ in } S, \quad u = g(x,y) \text{ in } \partial S, \tag{11.75}$$

where S and ∂S denote the usual square $\{-1 \le x \le 1, -1 \le y \le 1\}$ and its boundary. Then we may eliminate g (and replace it by zero) in (11.75), by superposing two problem solutions

$$u = u_1 + u_2, \tag{11.76}$$

where u_1 is the solution of the Laplace problem ((11.75) with $f \equiv 0$) and u_2 is the solution of the simple Poisson problem ((11.75) with $g \equiv 0$), so that

$$\Delta u_1 = 0 \text{ in } S, \ \ u_1 = g(x,y) \text{ on } \partial S, \tag{11.77a}$$

$$\Delta u_2 = f(x,y) \text{ in } S, \ \ u_2 = 0 \text{ on } \partial S. \tag{11.77b}$$

We gave details above of the determination of u_1 from four Fourier sine series expansions based on separation of variables, as per (11.71) above. We may therefore restrict attention to the problem (11.77b) defining u_2, which we now rename u.

We now re-address (11.77b), which was discussed in Section 11.8.1 using a power basis, and, for greater efficiency and stability, we adopt a boundary method based on a Chebyshev polynomial representation

$$u_{mn} = (x^2-1)(y^2-1) \sum_{i=0}^{m-2}{}' \sum_{j=0}^{n-2}{}' c_{ij} T_i(x) T_j(y), \tag{11.78}$$

or equivalently, again to ensure that $u = 0$ on ∂S,

$$u_{mn} = \sum_{i=2}^{m} \sum_{j=2}^{n} a_{ij} [T_i(x) - T_{i \bmod 2}(x)]\, [T_j(y) - T_{j \bmod 2}(y)]. \tag{11.79}$$

Now, as in Problem 16 of Chapter 2,

$$\frac{\partial^2}{\partial x^2} T_i(x) = \sum_{\substack{r=0 \\ i-r \text{ even}}}^{i-2} (i-r)i(i+r) T_r(x), \tag{11.80}$$

and hence

$$\Delta u_{mn} = \sum_{i=2}^{m} \sum_{j=2}^{n} a_{ij} \left[\sum_{\substack{r=0 \\ i-r \text{ even}}}^{i-2} (i-r)i(i+r) T_r(x) \left(T_j(y) - T_{j \bmod 2}(y)\right) + \right.$$

$$
\left. + \sum_{\substack{s=0 \\ j-s \text{ even}}}^{j-2} (j-s)j(j+s)T_s(y)\left(T_i(x) - T_{i \bmod 2}(x)\right) \right]
$$

$$
= f \tag{11.81}
$$

Now define collocation points $\{x_k(k = 1, \ldots, m-1)\}$ and $\{y_l(l = 1, \ldots, n-1)\}$ to be, respectively, the zeros of $T_{m-1}(x)$ and $T_{n-1}(y)$. Then discrete orthogonality gives, for p, r less than $m - 1$ and q, s less than $n - 1$,

$$
2(m+1)^{-1} \sum_{k=1}^{m-1} T_p(x_k)T_r(x_k) = \begin{cases} 2, & p = r = 0, \\ 1, & p = r \neq 0, \\ 0, & p \neq r, \end{cases} \tag{11.82a}
$$

$$
2(n+1)^{-1} \sum_{l=1}^{n-1} T_q(y_l)T_s(y_l) = \begin{cases} 2, & q = s = 0, \\ 1, & q = s \neq 0, \\ 0, & q \neq s. \end{cases} \tag{11.82b}
$$

Evaluating (11.81) at (x_k, y_ℓ), multiplying by $4[(m-1)(n-1)]^{-1}$ and by $T_p(x_k)T_q(y_\ell)$ for $p = 0, \ldots, m-2; q = 0, \ldots, n-2$, summing over k, ℓ, and using discrete orthogonality, we obtain

$$
A_{pq} + B_{pq} = 4[(m-1)(n-1)]^{-1} \sum_{k=1}^{m-1} \sum_{\ell=1}^{n-1} f(x_k, y_\ell)T_p(x_k)T_q(y_\ell), \tag{11.83}
$$

where

$$
A_{pq} = \begin{cases} \displaystyle\sum_{\substack{i=2 \\ i-p \text{ even}}}^{m} a_{iq}(i-p)i(i+p), & q \geq 2, \\ \displaystyle -\sum_{\substack{i=2 \\ i-p \text{ even}}}^{m} \sum_{\substack{j=3 \\ j \text{ odd}}}^{n} a_{ij}(i-p)i(i+p), & q = 1, \\ \displaystyle -2\sum_{\substack{i=2 \\ i-p \text{ even}}}^{m} \sum_{\substack{j=2 \\ j \text{ even}}}^{n} a_{ij}(i-p)i(i+p), & q = 0, \end{cases}
$$

$$
B_{pq} = \begin{cases} \displaystyle\sum_{\substack{j=2 \\ j-q \text{ even}}}^{n} a_{pj}(j-q)j(j+q), & p \geq 2, \\ \displaystyle -\sum_{\substack{j=2 \\ j-q \text{ even}}}^{n} \sum_{\substack{i=3 \\ i \text{ odd}}}^{m} a_{ij}(j-q)j(j+q), & p = 1, \\ \displaystyle -2\sum_{\substack{j=2 \\ j-q \text{ even}}}^{n} \sum_{\substack{i=2 \\ i \text{ even}}}^{m} a_{ij}(j-q)j(j+q), & p = 0. \end{cases}
$$

$$
\tag{11.84}
$$

This system of linear equations for a_{ij} is very sparse, having between 2 and $(m+n-2)/2$ entries in each row of the matrix for $p, q \geq 2$. It is only the equations where "boundary effects" enter (for $p = 0, 1$; $q = 0, 1$), that fill out the matrix entries into alternate rows and/or columns. Note also that all right-hand sides are discrete Chebyshev transforms, which could be evaluated by adopting FFT techniques.

The border effects can be neatly avoided for this particular Poisson problem, by adopting instead a matrix method based on differentiation matrices, in which the unknowns of the problem become the solution values at Chebyshev nodes, rather than the solution coefficients. This approach was introduced in Section 10.5.3 of Chapter 10 for ODEs and is particularly convenient for some relatively simple problems. We now see its advantages for the present problem.

11.8.4 Differentiation matrix approach: Poisson problem

To illustrate this approach, we follow the treatment of Trefethen (2000), setting $m = n$ and adopting as collocation points the tensor product of the zeros of $(1 - x^2)U_{n-1}(x)$ and the zeros of $(1 - y^2)U_{n-1}(y)$. The reader is referred to Section 10.5.2 for a detailed definition of the $(n+1) \times (n+1)$ differentiation matrix $\mathbf{D} \equiv \mathbf{D}_n$, which transforms all u values at collocation points into approximate derivative values at the same points, by forming linear combinations of the u values. The problem is significantly simplified by noting that the border elements of $\mathbf{D}$, namely the first and last rows and columns of $\mathbf{D}_n$, correspond to zero boundary values and may therefore be deleted to give an active $(n-1) \times (n-1)$ matrix $\tilde{\mathbf{D}}_n$.

For the Poisson problem

$$\Delta u = f(x, y) \text{ in } S : \{-1 \leq x \leq 1, -1 \leq y \leq 1\}, \tag{11.85a}$$

$$u = 0 \text{ on } \partial S : \{x = \pm 1, y = \pm 1\}, \tag{11.85b}$$

the method determines a vector $\mathbf{u}$ of approximate solutions at the interior collocation points (compare (10.52) with $\mathbf{e}_0 = \mathbf{e}_n = \mathbf{0}$) by solving

$$\mathbf{E}_n \mathbf{u} = \mathbf{f}_n \tag{11.86}$$

where

$$\mathbf{E}_n = \mathbf{I} \otimes \tilde{\mathbf{D}}_n^2 + \tilde{\mathbf{D}}_n^2 \otimes \mathbf{I} \tag{11.87}$$

and $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker (tensor) product, illustrated by the example

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \left(\begin{array}{cc|cc} a\alpha & a\beta & b\alpha & b\beta \\ a\gamma & a\delta & b\gamma & b\delta \\ \hline c\alpha & c\beta & d\alpha & d\beta \\ c\gamma & c\delta & d\gamma & d\delta \end{array}\right).$$

For example, for $n = 3$, this gives (see (10.49) *et seq.*)

$$
\begin{aligned}
\mathbf{E}_3 &= \mathbf{I} \otimes \tilde{\mathbf{D}}_3^2 \;+\; \tilde{\mathbf{D}}_3^2 \otimes \;\mathbf{I} \\
&= \begin{pmatrix}
-14 & 6 & -2 & & & & & & \\
4 & -6 & 4 & & & & & & \\
-2 & 6 & -14 & & & & & & \\
& & & -14 & 6 & -2 & & & \\
& & & 4 & -6 & 4 & & & \\
& & & -2 & 6 & -14 & & & \\
& & & & & & -14 & 6 & -2 \\
& & & & & & 4 & -6 & 4 \\
& & & & & & -2 & 6 & -14
\end{pmatrix} + \\
&\quad + \begin{pmatrix}
-14 & & & 6 & & & -2 & & \\
& -14 & & & 6 & & & -2 & \\
& & -14 & & & 6 & & & -2 \\
4 & & & -6 & & & 4 & & \\
& 4 & & & -6 & & & 4 & \\
& & 4 & & & -6 & & & 4 \\
-2 & & & 6 & & & -14 & & \\
& -2 & & & 6 & & & -14 & \\
& & -2 & & & 6 & & & -14
\end{pmatrix} .
\end{aligned}
\tag{11.88}
$$

Clearly the matrix $\mathbf{E}_n$ is sparse and easily calculable. Hence (11.86) may readily be solved by Gauss elimination, using for example Matlab's efficient system. The right-hand side of (11.86) is simply obtained by evaluating f at the interior collocation points. The shape (location of the non-zero terms) of

Figure 11.4: Shape of matrix $\mathbf{E}_n$ for $n = 3$

$$
\left(\begin{array}{ccc|ccc|ccc}
\times & \times & \times & \times & & & \times & & \\
\times & \times & \times & & \times & & & \times & \\
\times & \times & \times & & & \times & & & \times \\
\hline
\times & & & \times & \times & \times & \times & & \\
& \times & & \times & \times & \times & & \times & \\
& & \times & \times & \times & \times & & & \times \\
\hline
\times & & & \times & & & \times & \times & \times \\
& \times & & & \times & & \times & \times & \times \\
& & \times & & & \times & \times & \times & \times
\end{array}\right)
$$

the matrix $\mathbf{E}_n$ is illustrated in Figure 11.4 and consists of a diagonal of square matrices, flanked by diagonal matrices. The matrix is very sparse in all rows, having $2n - 3$ non-zero entries out of $(n - 1)^2$.

This Differentiation Matrix method is very attractive and efficient for this problem, and should always be given consideration in problems that respond to it. We now turn our attention to a more general problem, with non-zero boundary conditions.

11.8.5 Explicit collocation for the quasilinear Dirichlet problem: Chebyshev basis

We now continue to exploit the better conditioning of a Chebyshev polynomial basis, but we also consider the greater generality of a Dirichlet problem for a quasilinear elliptic equation on a square, namely

$$Lu \equiv a.u_{xx} + b.u_{xy} + c.u_{yy} + d.u_x + e.u_y = f \text{ in } D : |x| \le 1, |y| \le 1, \quad (11.89a)$$

$$u = g(x, y) \text{ on } \partial D : \{x = \pm 1, y = \pm 1\}, \quad (11.89b)$$

where a, b, c, d, e, f are functions of x and y defined in D, $g(x, y)$ is defined on ∂D only, and where, to ensure ellipticity,

$$a.c \ge b^2 \text{ for all } (x, y) \text{ in } D. \quad (11.90)$$

This is an extension of recent work by Mason & Crampton (2002).

For generality we do not attempt to include the boundary conditions (11.89b) implicitly in the form of approximation, but rather we represent them by a set of constraints at a discrete set of selected points, namely Chebyshev zeros on the boundary. Moreover we adopt a Chebyshev polynomial basis for u:

$$u \simeq u_{mn} = \sum_{i=0}^{m}{}' \sum_{j=0}^{n}{}' a_{ij} T_i(x) T_j(y). \quad (11.91)$$

As it happens, we find that the apparently most logical collocation procedure, similar to that of Section 11.8.1 above, for approximately solving (11.89a), (11.89b) in the form (11.91), leads to a singular matrix and requires modification. More details about this follow as the method develops. The fundamental idea that we use is that, since u_{mn}, given by (11.91), has $(m + 1)(n + 1)$ undetermined coefficients, we expect to be able to generate an appropriate set of equations for a_{ij} if we form $(m - 1)(n - 1)$ equations by collocating (11.89a) at a tensor product of Chebyshev zeros and a further $2m + 2n$ equations by collocating (11.89b) at Chebyshev zeros on the boundary. It is in the formation of the latter boundary equations that difficulties arise, and so we consider these equations first, noting that they are completely independent of the specification $Lu = f$ of the elliptic equation (11.89a).

To form the $2m + 2n$ boundary equations for a_{ij}, we set u_{mn} equal to g at the zeros, X_k $(k = 1, \ldots, m)$ and Y_ℓ $(\ell = 1, \ldots, n)$ of $T_m(x)$ and $T_n(y)$, respectively, on $y = \pm 1$ and $x = \pm 1$, respectively. This gives the two pairs of equations

$$\begin{aligned}
\sum_{i=0}^{m}{}' \sum_{j=0}^{n}{}' a_{ij} T_i(X_k) T_j(\pm 1) &= g(X_k, \pm 1), \\
\sum_{i=0}^{m}{}' \sum_{j=0}^{n}{}' a_{ij} T_i(\pm 1) T_j(Y_\ell) &= g(\pm 1, Y_\ell).
\end{aligned} \tag{11.92}$$

If we add/subtract these pairs of equations, noting that $T_j(1) = 1$ and that $T_j(-1) = (-1)^j$, we deduce that

$$\begin{aligned}
\sum_{i=0}^{m}{}' \sum_{\substack{j=0 \\ j \text{ even}}}^{n}{}' a_{ij} T_i(X_k) &= G_{k0} \equiv \tfrac{1}{2}(g(X_k, 1) + g(X_k, -1)), (k = 1, \ldots, m) \\
\sum_{i=0}^{m}{}' \sum_{\substack{j=1 \\ j \text{ odd}}}^{n} a_{ij} T_i(X_k) &= G_{k1} \equiv \tfrac{1}{2}(g(X_k, 1) - g(X_k, -1)), (k = 1, \ldots, m) \\
\sum_{\substack{i=0 \\ i \text{ even}}}^{m}{}' \sum_{j=0}^{n}{}' a_{ij} T_j(Y_\ell) &= H_{\ell 0} \equiv \tfrac{1}{2}(g(1, Y_\ell) + g(-1, Y_\ell)), (\ell = 1, \ldots, n) \\
\sum_{\substack{i=1 \\ i \text{ odd}}}^{m} \sum_{j=0}^{n}{}' a_{ij} T_j(Y_\ell) &= H_{\ell 1} \equiv \tfrac{1}{2}(g(1, Y_\ell) - g(-1, Y_\ell)), (\ell = 1, \ldots, n)
\end{aligned} \tag{11.93}$$

where the arrays G_{kp}, $H_{\ell q}$ are defined above for $p = 0, 1$; $q = 0, 1$.

Now, defining w_i to be $2/(i+1)$ for all i, multiplying the first pair of equations in (11.93) by $w_m T_r(X_k)$ and summing over k, multiplying the second pair of equations by $w_n T_s(Y_\ell)$ and summing over ℓ, and exploiting discrete orthogonality, it follows that

$$\begin{aligned}
R_{r0} \equiv \sum_{\substack{j=0 \\ j \text{ even}}}^{n}{}' a_{rj} &= J_{r0} \equiv w_m \sum_{k=1}^{m} T_r(X_k) G_{k0}, (r = 0, \ldots, m-1) \\
R_{r1} \equiv \sum_{\substack{j=1 \\ j \text{ odd}}}^{n} a_{rj} &= J_{r1} \equiv w_m \sum_{k=1}^{m} T_r(X_k) G_{k1}, (r = 0, \ldots, m-1) \\
C_{s0} \equiv \sum_{\substack{i=0 \\ i \text{ even}}}^{m}{}' a_{is} &= K_{s0} \equiv w_n \sum_{\ell=1}^{n} T_s(Y_\ell) H_{\ell 0}, (s = 0, \ldots, n-1)
\end{aligned}$$

$$C_{s1} \equiv \sum_{\substack{i=1 \\ i \text{ odd}}}^{m} a_{is} = K_{s1} \equiv w_n \sum_{\ell=1}^{n} T_s(Y_\ell) H_{\ell 1}, (s = 0, \ldots, n-1) \quad (11.94)$$

where R, C, J, K are defined to form left-hand sides (R, C) or right-hand sides (J, K) of the relevant equations. In addition each R or C is a linear sum of alternate elements in a row or column, respectively, of the matrix $\mathbf{A} = [a_{ij}]$.

Now we claim that this set of $2m+2n$ linear equations (11.94) in $a_{00}, \ldots, a_{mn}$ is not of rank $2m + 2n$ but rather of rank $2m + 2n - 1$. Indeed, it is easy to verify that a sum of alternate rows of $\mathbf{A}$ equals a sum of alternate columns; specifically

$$\sum_{\substack{i=0 \\ n-i \text{ odd}}}^{n-1}{}' R_{ip} = \sum_{\substack{j=0 \\ m-j \text{ odd}}}^{m-1}{}' C_{jq} = \sum_{\substack{i=0 \\ m-i \text{ odd}}}^{m-1}{}' \sum_{\substack{j=0 \\ n-j \text{ odd}}}^{n-1}{}' a_{ij}, \quad (11.95)$$

where $p = 0, 1$ for $m - 1$ even, odd, respectively, and $q = 0, 1$ for $n - 1$ even, odd, respectively. For example, for $m = n = 4$,

$$R_{11} + R_{31} = C_{11} + C_{31} = a_{11} + a_{13} + a_{31} + a_{33}, \quad (11.96)$$

and, for $m = n = 3$,

$$\tfrac{1}{2} R_{00} + R_{20} = \tfrac{1}{2} C_{00} + C_{20} = \tfrac{1}{2}(\tfrac{1}{2} a_{00} + a_{02} + a_{20}) + a_{22}. \quad (11.97)$$

Clearly we must delete one equation from the set (11.94) and add an additional independent equation in order to restore full rank. For simplicity we shall only discuss the cases where m, n are both even or both odd, leaving the even/odd and odd/even cases to the reader.

For m, n both even, we delete $C_{11} = K_{11}$ from (11.94) and add an averaged "even/even" boundary collocation equation

$$\begin{aligned} &\tfrac{1}{4}[u(1, 1) + u(-1, 1) + u(-1, -1) + u(1, -1)] = \\ &\lambda_{00} := \tfrac{1}{4}[g(1, 1) + g(-1, 1) + g(-1, -1) + g(1, -1)]. \end{aligned} \quad (11.98)$$

This simplifies to

$$\sum_{\substack{i=0 \\ i \text{ even}}}^{m}{}' R_{i0} = \lambda_{00} \quad (11.99)$$

where R_{m0} is defined by extending the definition (11.94) of R_{r0} to $r = m$ and where λ_{00} is as defined in (11.98). We may eliminate every R except R_{m0} from this equation, by using (11.94), to give a simplified form for the extra equation

$$R_{m0} = J_{m0} = \lambda_{00} - \sum_{\substack{i=0 \\ i \text{ even}}}^{m-2}{}' J_{i0} \quad (11.100)$$

where the right-hand side J_{m0} is defined as shown.

For m, n both odd, we delete $C_{00} = K_{00}$ from (11.94) and add an averaged "odd/odd" boundary collocation equation

$$\tfrac{1}{4}[u(1,1) - u(-1,1) + u(-1,-1) - u(1,-1)] =$$
$$\lambda_{11} := \tfrac{1}{4}[g(1,1) - g(-1,1) + g(-1,-1) - g(1,-1)]. \tag{11.101}$$

This simplifies to

$$\sum_{\substack{j=1 \\ j \text{ odd}}}^{n} C_{j1} = \lambda_{11} \tag{11.102}$$

where C_{n1} is defined by extending the definition (11.94) of C_{s1} to $s = n$ and where λ_{11} is as defined in (11.101).

We may eliminate every C except C_{n1} from this equation, by using (11.94), to give a simplified form for the extra equation

$$C_{n1} = K_{n1} \equiv \lambda_{11} - \sum_{\substack{j=1 \\ j \text{ odd}}}^{n-2} K_{j1} \tag{11.103}$$

where the right-hand side K_{n1} is defined as shown.

We now have $2m + 2n$ non-singular equations for the coefficients a_{ij}, and it remains to handle the elliptic equation by collocation at $(m-1)(n-1)$ suitable Chebyshev points in D.

For a general quasilinear equation we should set $Lu = f$ at a tensor product of $(m-1)\times(n-1)$ Chebyshev zeros, giving the same number of linear algebraic equations for $\{a_{ij}\}$, and these equations together with the $2m+2n$ boundary collocation equations would then be solved as a full system.

For simplicity, and so that we can give fuller illustrative details, we concentrate on the Poisson equation, as a special example of (11.89a), corresponding to the form

$$Lu \equiv \Delta u \equiv u_{xx} + u_{yy} = f(x,y) \text{ in } D. \tag{11.104}$$

Now second derivatives of Chebyshev sums are readily seen (see Chapter 2) to take the form

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} T_k(x) = \sideset{}{'}\sum_{\substack{r=0 \\ k-r \text{ even}}}^{k-2} (k-r)k(k+r)T_r(x) \quad (k \geq 2), \tag{11.105a}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2} T_\ell(y) = \sideset{}{'}\sum_{\substack{s=0 \\ \ell-s \text{ even}}}^{\ell-2} (\ell-s)\ell(\ell+s)T_s(y) \quad (\ell \geq 2). \tag{11.105b}$$

Hence, on substituting (11.91) into (11.104), we obtain

$$\begin{aligned}\Delta u_{mn} &= \sum_{k=2}^{m} \sum_{\ell=0}^{n}{}' a_{k\ell} T_\ell(y) \sum_{\substack{r=0\\k-r \text{ even}}}^{k-2}{}' (k-r)k(k+r)T_r(x) \\ &\quad + \sum_{\ell=2}^{n} \sum_{k=0}^{m}{}' a_{k\ell} T_k(x) \sum_{\substack{s=0\\\ell-s \text{ even}}}^{\ell-2}{}' (\ell-s)\ell(\ell+s)T_s(y) \\ &= A_{mn}, \text{ say.}\end{aligned} \tag{11.106}$$

Setting Δu_{mn} equal to $f(x, y)$ at the abscissae x_i, y_j , where x_i are zeros of $T_{m-1}(x)$ and y_j are zeros of $T_{n-1}(y)$ (for $i = 1, \ldots, m-1$; $j = 1, \ldots, n-1$), multiplying by $T_p(x_i)T_q(y_j)$, and summing over i, j, we deduce that, for every $p = 0, \ldots, m-2$; $q = 0, \ldots, n-2$:

$$\begin{aligned}E_{pq} &\equiv \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} A_{mn}(x_i, y_j) T_p(x_i) T_q(yj) \\ &= \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} f(x_i, y_j) T_p(x_i) T_q(y_j) \\ &\equiv f_{pq},\end{aligned} \tag{11.107}$$

where f_{pq} represents the discrete Chebyshev transform of f with respect to $T_p(x)T_q(y)$. Substituting for A_{mn},

$$\begin{aligned}E_{pq} &= \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sum_{k=2}^{m} \sum_{\ell=0}^{n}{}' a_{k\ell} T_\ell(y_j) T_q(y_j) \sum_{r=0}^{k-2}{}' (k-r)k(k+r) T_r(x_i) T_p(x_i) \\ &\quad + \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sum_{\ell=2}^{n} \sum_{k=0}^{n}{}' a_{k\ell} T_k(x_i) T_p(x_i) \sum_{s=0}^{\ell-2}{}' (\ell-s)\ell(\ell+s) T_s(y_j) T_q(j_j) \\ &= \sum_{k=2}^{m} \sum_{\ell=0}^{n}{}' \sum_{j=1}^{n-1} T_\ell(y_j) T_q(y_j) a_{k\ell} \sum_{\substack{r=0\\k-r \text{ even}}}^{k-2}{}' (k-r)k(k+r) \sum_{i=1}^{m-1} T_r(x_i) T_p(x_i) \\ &\quad + \sum_{\ell=2}^{n} \sum_{k=0}^{m}{}' \sum_{i=1}^{m-1} T_k(x_i) T_p(x_i) a_{k\ell} \sum_{\substack{s=0\\\ell-s \text{ even}}}^{\ell-2}{}' (\ell-s)\ell(\ell+s) \sum_{j=1}^{n-1} T_s(y_j) T_q(y_j).\end{aligned} \tag{11.108}$$

Using the discrete orthogonality property that, for example,

$$\sum_{j=1}^{n-1} T_\ell(y_j)T_q(y_j) = \begin{cases} 0, & \ell \neq q \\ (n-1)/2, & \ell = q \neq 0 \\ n-1, & \ell = q = 0 \end{cases},$$

we deduce that

$$\begin{aligned} E_{pq} = & \left[\sum_{\substack{k=p+2 \\ k-p \text{ even}}}^{m} \tfrac{1}{2}(n-1)a_{kq} + \sum_{j=1}^{n-1} T_{n-1}(y_j)T_q(y_j)a_{k,n-1} \right. \\ & \left. + \sum_{j=1}^{n-1} T_n(y_j)T_q(y_j)a_{kn} \right] \tfrac{1}{2}(m-1)(k-p)k(k+p) \\ & + \left[\sum_{\substack{\ell=q+2 \\ \ell-q \text{ even}}}^{n} \tfrac{1}{2}(m-1)a_{p\ell} + \sum_{i=1}^{m-1} T_{m-1}(x_i)T_p(x_i)a_{m-1,\ell} \right. \\ & \left. + \sum_{i=1}^{m-1} T_m(x_i)T_p(x_i)a_{m\ell} \right] \tfrac{1}{2}(n-1)(\ell-q)\ell(\ell+q). \end{aligned} \tag{11.109}$$

Now, by the definition of x_i, y_j, we know that $T_{m-1}(x_i)$ and $T_{n-1}(y_j)$ are zero. Also, using the three-term recurrence at x_i,

$$T_m(x_i) = 2xT_{m-1}(x) - T_{m-2}(x_i) = -T_{m-2}(x_i), \quad T_n(y_j) = -T_{n-2}(y_j). \tag{11.110}$$

Substituting these values into (11.109), using discrete orthogonality, and using the Kronecker delta notation

$$\delta_{rs} = 1 \ (r = s), \quad \delta_{rs} = 0 \ (r \neq s), \tag{11.111}$$

we deduce that

$$\begin{aligned} E_{pq} \equiv & \tfrac{1}{4}(m-1)(n-1) \left[\sum_{\substack{k=p+2 \\ k-p \text{ even}}}^{m} \left(a_{kq} - \delta_{q,n-2}a_{kn}\right)(k-p)k(k+p) + \right. \\ & \left. + \sum_{\substack{l=q+2 \\ l-q \text{ even}}}^{n} \left(a_{pl} - \delta_{p,m-2}a_{ml}\right)(l-q)l(l+q) \right] \\ = & f_{pq} \quad (p = 0, \ldots, m-2;\ q = 0, \ldots, n-2). \end{aligned} \tag{11.112}$$

For example, for $m = n = 3$ we have this set of four collocation equations:

$$\begin{aligned}
E_{11} &\equiv \frac{4}{4}[(a_{31} - a_{33}) + (a_{13} - a_{33})]2.3.4 \\
&= 24(a_{13} + a_{31} - 2a_{33}) = f_{11} = 24F_{11}, \\
E_{10} &\equiv a_{30}2.3.4 + (a_{12} - a_{32})2.2.2 \\
&= 8(a_{12} + 3a_{30} - a_{32}) = f_{10} = 8F_{10}, \\
E_{01} &\equiv (a_{21} - a_{23})2.2.2 + a_{03}2.3.4 \\
&= 8(3a_{03} + a_{21} - a_{23}) = f_{01} = 8F_{01}, \\
E_{00} &\equiv a_{20}2.2.2 + a_{02}2.2.2 \\
&= 8(a_{02} + a_{20}) = f_{00} = 8F_{00}, \qquad (11.113)
\end{aligned}$$

where F_{ij} are defined as shown by scaling f_{ij}. For $m = n = 4$, the system (11.112) gives the following nine equations, where we leave the reader to confirm the details:

$$\begin{aligned}
E_{22} &\equiv 108(a_{24} + a_{42} - 2a_{44}) \\
&= f_{22} = 108F_{22}, \\
E_{21} &\equiv 54(a_{23} + 2a_{41} - a_{43}) \\
&= f_{21} = 54F_{21}, \\
E_{20} &\equiv 18(a_{22} + 8a_{24} + 6a_{40} - a_{42} - 8a_{44}) \\
&= f_{20} = 18F_{20}, \\
E_{12} &\equiv 54(2a_{14} + a_{32} - a_{34}) \\
&= f_{12} = 54F_{12}, \\
E_{11} &\equiv 54(a_{13} + a_{31}) \\
&= f_{11} = 54F_{11}, \\
E_{10} &\equiv 18(a_{12} + 8a_{12} + 3a_{30}) \\
&= f_{10} = 18F_{10}, \\
E_{02} &\equiv 18(6a_{04} + a_{22} - a_{24} + 8a_{42} - 8a_{44}) \\
&= f_{02} = 18F_{02}, \\
E_{01} &\equiv 18(3a_{03} + a_{21} + 8a_{41}) \\
&= f_{01} = 18F_{01}, \\
E_{00} &\equiv 18(a_{02} + 8a_{04} + a_{20} + 8a_{40}) \\
&= f_{00} = 18F_{00}. \qquad (11.114)
\end{aligned}$$

For $m = n = 4$, the complete set of 25 collocation equations, 16 boundary equations and 9 interior PDE equations, namely (11.94) for $m = n = 4$ and

(11.114), may be written in the matrix form

$$\mathbf{M}\mathbf{a} = \mathbf{b}, \tag{11.115}$$

where $\mathbf{M}$ is the matrix of equation entries and $\mathbf{a}$ is the column vector of approximation coefficients

$$\mathbf{a} = (a_{00}, a_{01}, \ldots, a_{04}, a_{10}, a_{11}, \ldots, a_{14}, a_{20}, \ldots, a_{30}, \ldots, a_{40}, \ldots, a_{44})' \tag{11.116}$$

and $\mathbf{b}$ is the set of right-hand sides, either boundary or PDE terms, in appropriate order. In Table 11.2 we display the matrices $\mathbf{M}$, $\mathbf{a}$, $\mathbf{b}$ for $m = n = 4$, blank entries denoting zeros. The column of symbols to the left of $\mathbf{M}$ indicates which equation has been used to construct each row. Note that we have ordered the equations to give a structure in $\mathbf{M}$ as close to lower triangular as possible. The order used is based on:

$$R_{4*}, R_{3*}, E_{2*}, R_{2*}, E_{1*}, R_{1*}, E_{0*}, R_{0*}, C_{3*}, C_{2*}, C_{1*}, C_{0*} \tag{11.117}$$

where $*$ is a wild subscript, E indicates a PDE term, and R, C indicate boundary conditions.

On studying Table 11.2, some important facts emerge. The coefficients a_{ij} appearing in any equation are exclusively in one of the four symmetry classes: i, j both even, i, j both odd, i odd and j even, and i even and j odd. Thus the set of 25 equations can be separated into four wholly independent subsystems, respectively involving 4 subsets of a_{ij}. These four subsystems are shown in Table 11.3 for $m = n = 4$, and they consist of 8,5,6,6 equations in 9,4,6,6 coefficients a_{ij}, respectively.

This immediately confirms that we have a surplus equation in the odd/odd subsystem (for $a_{11}, a_{13}, a_{31}, a_{33}$) and one too few equations in the even/even subsystem. As proposed in earlier discussions, we therefore delete equation C_{11} and replace it by equation R_{40}, as indicated in Table 11.2.

The extremely sparse nature of the matrix $\mathbf{M}$ is clear from Table 11.2, and moreover the submatrices formed from even and/or odd subsystems remain relatively sparse, as is illustrated in Tables 11.3 to 11.6.

The odd/odd subsystem (for $m = n = 4$) (in Table 11.6) is remarkably easy to solve in the case $g \equiv 0$ of zero boundary conditions, when $J_{**} = K_{**} = 0$. The solution is then

$$-a_{11} = a_{13} = a_{31} = -a_{33} = \tfrac{1}{2} F_{11}. \tag{11.118}$$

In Table 11.7, we also show the full algebraic system for the odd degrees $m = n = 3$, and in Tables 11.8 to 11.11 the equations are separated into their four even/odd subsystems. The equation C_{00} is noted and is to be deleted, while the equation C_{31} has been added. Equations are again ordered so as to optimise sparsity above the diagonal. The $m = n = 3$ subsystems are easily

Table 11.2: Full collocation matrix—Poisson problem: $m = n = 4$

$$
\begin{array}{c}
R_{40} \\ \hline
R_{31} \\ R_{30} \\ E_{22} \\ E_{21} \\ E_{20} \\ R_{21} \\ R_{20} \\ E_{12} \\ E_{11} \\ E_{10} \\ R_{11} \\ R_{10} \\ E_{02} \\ E_{01} \\ E_{00} \\ R_{01} \\ R_{00} \\ C_{30} \\ C_{31} \\ C_{21} \\ C_{20} \\ C_{10} \\ \hline
C_{11} \\ \hline
C_{01} \\ C_{02}
\end{array}
\left[\begin{array}{ccccccccccccccccccccccccc}
&&&& &&&&& &&&&& &&&&& \frac{1}{2}&0&1&0&1 \\ \hline
&&&& &&&&& &&&&& 0&1&0&1&0 &&&&& \\
&&&& &&&&& &&&&& \frac{1}{2}&0&1&0&1 &&&&& \\
&&&& &&&&& 0&0&0&0&1 &&&&& 0&0&1&0&-2 \\
&&&& &&&&& 0&0&0&1&0 &&&&& 0&2&0&-1&0 \\
&&&& &&&&& 0&0&1&0&8 &&&&& 6&0&-1&0&-8 \\
&&&& &&&&& 0&1&0&1&0 &&&&& &&&& \\
&&&& &&&&& \frac{1}{2}&0&1&0&1 &&&&& &&&& \\
&&&& & 0&0&0&0&2 &&&&& 0&0&1&0&-1 &&&&& \\
&&&& & 0&0&0&1&0 &&&&& 0&1&0&0&0 &&&&& \\
&&&& & 0&0&1&0&8 &&&&& 3&0&0&0&0 &&&&& \\
&&&& & 0&1&0&1&0 &&&&& &&&&& &&&& \\
&&&& & \frac{1}{2}&0&1&0&1 &&&&& &&&&& &&&& \\
0&0&0&0&6 &&&&& 0&0&1&0&-1 &&&&& 0&0&8&0&-8 \\
0&0&0&3&0 &&&&& 0&1&0&0&0 &&&&& 0&8&0&0&0 \\
0&0&1&0&8 &&&&& 1&0&0&0&0 &&&&& 8&0&0&0&0 \\
0&1&0&1&0 &&&&& &&&&& &&&&& &&&& \\
\frac{1}{2}&0&1&0&1 &&&&& &&&&& &&&&& &&&& \\
0&0&0&\frac{1}{2}&0 &&&&& 0&0&0&1&0 &&&&& 0&0&0&1&0 \\
&&&& & 0&0&0&1&0 &&&&& 0&0&0&1&0 &&&&& \\
&&&& & 0&0&1&0&0 &&&&& 0&0&1&0&0 &&&&& \\
0&0&\frac{1}{2}&0&0 &&&&& 0&0&1&0&0 &&&&& 0&0&1&0&0 \\
0&\frac{1}{2}&0&0&0 &&&&& 0&1&0&0&0 &&&&& 0&1&0&0&0 \\ \hline
&&&& & 0&1&0&0&0 &&&&& 0&1&0&0&0 &&&&& \\ \hline
&&&& & 1&0&0&0&0 &&&&& 1&0&0&0&0 &&&&& \\
\frac{1}{2}&0&0&0&0 &&&&& 1&0&0&0&0 &&&&& 1&0&0&0&0
\end{array}\right]
\left[\begin{array}{c}
a_{00} \\ a_{01} \\ a_{02} \\ a_{03} \\ a_{04} \\ a_{10} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{20} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \\ a_{30} \\ a_{31} \\ a_{32} \\ a_{33} \\ a_{34} \\ a_{40} \\ a_{41} \\ a_{42} \\ a_{43} \\ a_{44}
\end{array}\right]
=
\left[\begin{array}{c}
J_{40} \\ \hline
J_{31} \\ J_{30} \\ F_{22} \\ F_{21} \\ F_{20} \\ J_{21} \\ J_{20} \\ F_{12} \\ F_{11} \\ F_{10} \\ J_{11} \\ J_{10} \\ F_{02} \\ F_{01} \\ F_{00} \\ J_{01} \\ J_{00} \\ K_{30} \\ K_{31} \\ K_{21} \\ K_{20} \\ K_{10} \\ \hline
K_{11} \\ \hline
K_{01} \\ K_{02}
\end{array}\right]
$$

Table 11.3: $m = n = 4$, partial system odd/even in x/y

$$
\begin{matrix} R_{30} \\ E_{12} \\ E_{10} \\ R_{10} \\ C_{21} \\ C_{01} \end{matrix}
\begin{bmatrix}
 & & & \frac{1}{2} & 1 & 1 \\
0 & 0 & 2 & 0 & 1 & -1 \\
0 & 1 & 8 & 3 & 0 & 0 \\
\frac{1}{2} & 1 & 1 & & & \\
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}
\begin{bmatrix} a_{10} \\ a_{12} \\ a_{14} \\ a_{30} \\ a_{32} \\ a_{34} \end{bmatrix}
=
\begin{bmatrix} J_{30} \\ F_{12} \\ F_{10} \\ J_{10} \\ K_{21} \\ K_{01} \end{bmatrix}
$$

Table 11.4: $m = n = 4$, partial system even/odd in x/y

$$
\begin{matrix} E_{21} \\ R_{21} \\ E_{01} \\ R_{01} \\ C_{30} \\ C_{10} \end{matrix}
\begin{bmatrix}
 & & 0 & 1 & 2 & -1 \\
 & & 1 & 1 & & \\
0 & 3 & 1 & 0 & 8 & 0 \\
1 & 1 & & & & \\
0 & \frac{1}{2} & 0 & 1 & 0 & 1 \\
\frac{1}{2} & 0 & 1 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix} a_{01} \\ a_{03} \\ a_{21} \\ a_{23} \\ a_{41} \\ a_{43} \end{bmatrix}
=
\begin{bmatrix} F_{21} \\ J_{21} \\ F_{01} \\ J_{01} \\ K_{30} \\ K_{10} \end{bmatrix}
$$

Table 11.5: $m = n = 4$, partial system even/even in x/y

$$
\begin{array}{c} R_{40} \\ \hline E_{22} \\ E_{20} \\ R_{20} \\ E_{02} \\ E_{00} \\ R_{00} \\ C_{20} \\ C_{00} \end{array}
\left[\begin{array}{ccccccccc}
 & & & & & & \frac{1}{2} & 1 & 1 \\ \hline
 & & & 0 & 0 & 1 & 0 & 1 & -2 \\
 & & & 0 & 1 & 8 & 6 & -1 & -8 \\
 & & & \frac{1}{2} & 1 & 1 & & & \\
0 & 0 & 6 & 0 & 1 & 1 & 0 & 8 & -8 \\
0 & 1 & 8 & 1 & 0 & 0 & 8 & 0 & 0 \\
\frac{1}{2} & 1 & 1 & & & & & & \\
0 & \frac{1}{2} & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
\frac{1}{2} & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0
\end{array}\right]
\begin{bmatrix} a_{00} \\ a_{02} \\ a_{04} \\ a_{20} \\ a_{22} \\ a_{24} \\ a_{40} \\ a_{42} \\ a_{44} \end{bmatrix}
=
\left[\begin{array}{c} J_{40} \\ \hline F_{22} \\ F_{20} \\ J_{20} \\ F_{02} \\ F_{00} \\ J_{00} \\ K_{20} \\ K_{00} \end{array}\right]
$$

Table 11.6: $m = n = 4$, partial system odd/odd in x/y

$$
\begin{array}{c} R_{31} \\ E_{11} \\ R_{11} \\ C_{31} \\ \hline C_{11} \end{array}
\left[\begin{array}{cccc}
 & & 1 & 1 \\
0 & 1 & 1 & 0 \\
1 & 1 & & \\
0 & 1 & 0 & 1 \\ \hline
1 & 0 & 1 & 0
\end{array}\right]
\begin{bmatrix} a_{11} \\ a_{13} \\ a_{31} \\ a_{33} \end{bmatrix}
=
\left[\begin{array}{c} J_{31} \\ F_{11} \\ J_{11} \\ K_{31} \\ \hline K_{11} \end{array}\right]
$$

Table 11.7: Full collocation matrix—Poisson problem: $m = n = 3$ (blank spaces contain zero entries)

	a_{00}	a_{01}	a_{02}	a_{03}	a_{10}	a_{11}	a_{12}	a_{13}	a_{20}	a_{21}	a_{22}	a_{23}	a_{30}	a_{31}	a_{32}	a_{33}	
C_{31}					0	0	0	1					0	0	0	1	K_{31}
R_{21}									0	1	0	1					J_{21}
R_{20}									$\frac{1}{2}$	0	1	0					J_{20}
E_{11}					0	0	0	1					0	1	0	-2	F_{11}
E_{10}					0	0	1	0					3	0	-1	0	F_{10}
R_{11}					0	1	0	1									J_{11}
R_{10}					$\frac{1}{2}$	0	1	0									J_{10}
E_{01}	0	0	0	3					0	1	0	-1					F_{01}
E_{00}	0	0	1	0					1	0	0	0					F_{00}
R_{01}	0	1	0	1													J_{01}
R_{00}	$\frac{1}{2}$	0	1	0													J_{00}
C_{20}	0	0	$\frac{1}{2}$	0					0	0	1	0					K_{20}
C_{21}					0	0	1	0					0	0	1	0	K_{21}
C_{11}					0	1	0	0					0	1	0	0	K_{11}
C_{10}	0	$\frac{1}{2}$	0	0					0	1	0	0					K_{10}
C_{00}	$\frac{1}{2}$	0	0	0					1	0	0	0					K_{00}
C_{01}					1	0	0	0					1	0	0	0	K_{01}

solved to give explicit formulae in the case $g \equiv 0$, as we now show. We leave it as an exercise to the reader to generate a set of tables for the case $m = n = 5$.

We may readily determine formulae for all coefficients a_{ij} for $m = n = 3$ by eliminating variables in Tables 11.8 to 11.11, and we leave this as an exercise to the reader (Problem 7).

We deduce from Table 11.8 that, for $g \equiv 0$, and hence $J_{**} = K_{**} = 0$, the even/even coefficients are

$$-a_{00} = 2a_{02} = 2a_{20} = -4a_{22} = F_{00}. \tag{11.119}$$

From Table 11.9, for $g \equiv 0$, the odd/odd coefficients are

$$-a_{11} = a_{13} = a_{31} = -a_{33} = \tfrac{1}{4}F_{11}. \tag{11.120}$$

From Table 11.10, for $g \equiv 0$, the even/odd coefficients are

$$-a_{01} = a_{03} = 2a_{21} = -2a_{23} = \tfrac{1}{4}F_{01}. \tag{11.121}$$

From Table 11.11, for $g \equiv 0$, the odd/even coefficients are

$$-a_{10} = 2a_{12} = a_{30} = -2a_{32} = \tfrac{1}{4}F_{10}. \tag{11.122}$$

Table 11.8: $m = n = 3$, partial system even/even in x/y

$$
\begin{array}{c} R_{20} \\ E_{00} \\ R_{00} \\ C_{20} \\ \hline C_{00} \end{array}
\left[\begin{array}{cccc} & & \frac{1}{2} & 1 \\ 0 & 1 & 1 & 0 \\ \frac{1}{2} & 1 & & \\ 0 & \frac{1}{2} & 0 & 1 \\ \hline \frac{1}{2} & 0 & 1 & 0 \end{array}\right]
\begin{bmatrix} a_{00} \\ a_{02} \\ a_{20} \\ a_{22} \end{bmatrix}
=
\left[\begin{array}{c} J_{20} \\ F_{00} \\ J_{00} \\ K_{20} \\ \hline K_{00} \end{array}\right]
$$

Table 11.9: $m = n = 3$, partial system odd/odd in x/y

$$
\begin{array}{c} C_{31} \\ \hline E_{11} \\ R_{11} \\ C_{11} \end{array}
\left[\begin{array}{cccc} 0 & 1 & 0 & 1 \\ \hline 0 & 1 & 1 & -2 \\ 1 & 1 & & \\ 1 & 0 & 1 & 0 \end{array}\right]
\begin{bmatrix} a_{11} \\ a_{13} \\ a_{31} \\ a_{33} \end{bmatrix}
=
\left[\begin{array}{c} K_{31} \\ \hline F_{11} \\ J_{11} \\ K_{11} \end{array}\right]
$$

Table 11.10: $m = n = 3$, partial system even/odd in x/y

$$
\begin{array}{c} R_{21} \\ E_{01} \\ R_{01} \\ C_{10} \end{array}
\left[\begin{array}{cccc} & & 1 & 1 \\ 0 & 3 & 1 & -1 \\ 1 & 1 & & \\ \frac{1}{2} & 0 & 1 & 0 \end{array}\right]
\begin{bmatrix} a_{01} \\ a_{03} \\ a_{21} \\ a_{23} \end{bmatrix}
=
\begin{bmatrix} J_{21} \\ F_{01} \\ J_{01} \\ K_{10} \end{bmatrix}
$$

Table 11.11: $m = n = 3$, partial system odd/even in x/y

$$
\begin{array}{c} E_{10} \\ R_{10} \\ C_{21} \\ C_{01} \end{array}
\left[\begin{array}{cccc} 0 & 1 & 3 & -1 \\ \frac{1}{2} & 1 & & \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right]
\begin{bmatrix} a_{10} \\ a_{12} \\ a_{30} \\ a_{32} \end{bmatrix}
=
\begin{bmatrix} F_{10} \\ J_{10} \\ K_{21} \\ K_{01} \end{bmatrix}
$$

Thus for zero boundary conditions, the approximate solution u is given very simply for $m = n = 3$. Indeed we see, not surprisingly, (Problem 8) that u_{mn} can be written exactly in the form

$$u_{mn} = (x^2 - 1)(y^2 - 1)(a + bx + cy + dxy). \qquad (11.123)$$

If J_{**} and K_{**} are not both zero, then no simplification such as (11.123) occurs, but we still obtain four separate sparse subsystems to solve for the coefficients a_{ij} for all m, n.

An alternative but closely related approach to the special case of the Poisson problem is given by Haidvogel & Zang (1979).

11.9 Computational fluid dynamics

One of the most important PDE problems in computational fluid dynamics is the *Navier–Stokes equation*

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \nu \Delta \mathbf{v}, \quad \nabla \cdot \mathbf{v} = 0 \qquad (11.124)$$

where $\mathbf{v}$ is the velocity vector, p is the pressure divided by the (constant) density, ν is the kinematic viscosity and Δ denotes the Laplacian. This problem is studied in detail in the lecture notes by Deville (1984), as well as in Canuto et al. (1988). Deville considers, as preparatory problems, the Helmholtz equation, the Burgers equation and the Stokes problem. We shall here briefly discuss the Burgers equation.

The Burgers equation is the nonlinear equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}, \qquad (11.125)$$

which we shall take to have the boundary and initial conditions

$$u(-1, t) = u(1, t) = 0, \quad u(x, 0) = u_0(x). \qquad (11.126)$$

The general procedure for solution is to discretise (11.125) into a first-order system of ordinary differential equations in t, which is solved by a scheme that is explicit as regards the nonlinear part and implicit for the linear part. Using Chebyshev collocation at the $n + 1$ points $\{y_j\}$, the discretisation can be written (Canuto et al. 1988) as

$$\mathbf{Z}_n \left(\frac{\partial \mathbf{u}_n}{\partial t} + \mathbf{U}_n \mathbf{D}_n \mathbf{u}_n - \nu \mathbf{D}^2 \mathbf{u}_n \right) = \mathbf{0}, \qquad (11.127)$$

where $\mathbf{D}_n$ is the appropriate Chebyshev collocation differentiation matrix, $\mathbf{u}_n$ is a vector with elements $u_n(y_j, t)$, $\mathbf{U}_n$ is a diagonal matrix with the elements of $\mathbf{u}_n$ on its diagonal and $\mathbf{Z}_n$ is a unit matrix with its first and last elements replaced by zeros. The boundary conditions are imposed by requiring $u_n(y_0, t) = u_n(y_n, t) = 0$. The method as it stands involves $O(n^2)$ operations at each time step for the implicit term, but this can be reduced to $O(n \log n)$ operations by using FFT techniques.

A Chebyshev tau method may instead be applied, defining

$$\langle f\,,\, T_k \rangle = \frac{2}{\pi c_k} \int_{-1}^{1} \frac{f(x) T_k(x)}{\sqrt{1-x^2}}\, \mathrm{d}x. \tag{11.128}$$

Then, defining

$$u_n(x,t) = {\sum_{k=0}^{n}}' a_k(t) T_k(x),$$

we have

$$\left\langle \frac{\partial u_n}{\partial t} + u_n \frac{\partial u_n}{\partial x} - \nu \frac{\partial^2 u_n}{\partial x^2}\,,\, T_k \right\rangle = 0, \tag{11.129}$$

which reduces to

$$\frac{\mathrm{d}a_k}{\mathrm{d}t} + \left\langle u_n \frac{\partial u_n}{\partial x}\,,\, T_k \right\rangle - \nu a_k^{(2)} = 0. \tag{11.130}$$

Again a mixed explicit/implicit method may be adopted for each time step, the inner product being evaluated explicitly.

For discussion of the Stokes and Navier–Stokes equations, the reader is referred to Deville (1984), Canuto et al. (1988), Fornberg (1996), and Gerritsma & Phillips (1998, 1999).

11.10 Particular issues in spectral methods

It is important to remember that the key advantages of spectral and pseudospectral methods lie in

1. the rapid (e.g., exponential) convergence of the methods for very smooth data and PDEs, which makes Chebyshev methods so powerful;

2. the use of discrete orthogonality, which greatly simplifies collocation equations;

3. the use of the FFT, which speeds up computations typically from $O(n^2)$ to $O(n \log n)$ operations;

4. the possibility of a matrix representation of derivatives, which simplifies the representation of the solution and boundary conditions in certain problems.

For the reasons above, the method will always be restricted to somewhat special classes of problems if it is to compete with more general methods like the finite element method. However, the spectral method shares with the finite element method a number of common features, including the pointwise and continuous representation of its solution (as in the differentiation matrix method) and the possibility of determining good preconditioners (Fornberg & Sloan 1994).

We now raise some further important issues that arise in spectral/pseudo-spectral methods. We do not have the scope to illustrate these issues in detail but can at least make the reader aware of their significance.

Aliasing (see Section 6.3.1) is an interesting feature of trigonometric and Chebyshev polynomials on discrete meshes. There is a potential for ambiguity of definition when a Chebyshev or Fourier series attempts to match a PDE on too coarse a grid. Fortunately, aliasing is not generally to be regarded as threatening, especially not in linear problems, but nonlinear problems do give cause for some concern on account of the possible occurrence of high-frequency modes which may be misinterpreted as low-frequency ones. Canuto et al. (1988, p.85) note that aliases may be removed by phase shifts, which can eliminate special relationships between low and high frequency modes.

Preconditioners are frequently used in finite-element methods to improve the conditioning of linear equations. Their use with finite differences for Chebyshev methods is discussed for example by Fornberg (1996), Fornberg & Sloan (1994) and Phillips et al. (1986). The idea is, for example, to take a system of linear equations whose matrix is neither diagonally dominant nor symmetric, and to find a multiplying matrix that yields a result that is strictly diagonally dominant, and therefore amenable to Gauss–Seidel iteration. More broadly, it improves the conditioning of the system matrix.

Basis functions in spectral methods may be not only Chebyshev polynomials, but also Legendre polynomials or trigonometric polynomials (Canuto et al. 1988). Legendre polynomials are sometimes preferred for Galerkin methods and Chebyshev polynomials for collocation methods (because of discrete orthogonality). Trigonometric polynomials are versatile but normally suitable for periodic functions only, because of the Gibbs phenomenon (see page 118, footnote). Clearly we are primarily interested in Chebyshev polynomials here, and so shall leave discussion of Legendre polynomials and other possible bases to others.

11.11 More advanced problems

The subject of partial differential equations is a huge one, and we cannot in this broadly-based book do full credit to spectral and pseudospectral methods. We have chosen to illustrate some key aspects of the methods, mainly for linear and quasilinear problems, and to emphasise some of the technical ideas that need to be exploited.

For discussion of other problems and, in particular, more advanced PDE problems including nonlinear problems, the reader is referred to such books as:

- Canuto et al. (1988) — for many fluid problems of varying complexity and solution structures, as well as an abundance of background theory;
- Trefethen (2000) — for a very useful collection of software and an easy-to-read discussion of the spectral collocation approach;
- Boyd (2000) — for a modern treatment including many valuable results;
- Guo Ben-yu (1998) — for an up-to-date and very rigorous treatment;
- Fornberg (1996) — as it says, for a practical guide to pseudospectral methods;
- Deville (1984) — for a straightforward introduction mainly to fluid problems;
- Gottlieb & Orszag (1977) — for an early and expository introduction to the spectral approach.

11.12 Problems for Chapter 11

1. Apply the method of separation of variables in (r, θ) coordinates to

$$\Delta u = r^2 \frac{\partial^2 u}{\partial r^2} + r \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0$$

   (see (11.11a) above) in the disc $S : r \leq 1$, where $u(1, \theta) = g(\theta)$ on $\partial S : r = 1$, and $g(\theta)$ is a known 2π-periodic function of the orientation of a point P of the boundary. Determine the solution as a series in the cases in which

   (a) $g(\theta) = \pi^2 - \theta^2$;

   (b) $g(\theta) = \begin{cases} -1, & -\pi \leq \theta \leq 0 \\ +1, & 0 \leq \theta \leq \pi. \end{cases}$

2. In addition to satisfying $(m-1)(n-1)$ specified linear conditions in the interior of the square domain $D : \{|x| < 1, |y| < 1\}$, the form

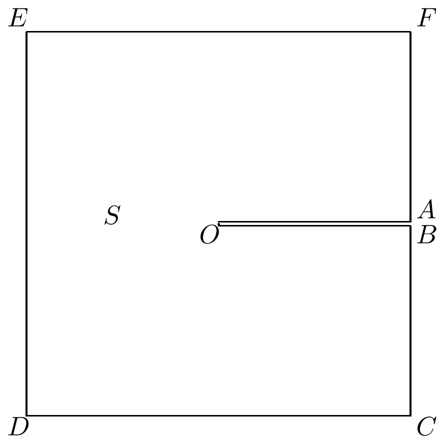
$$\sum_{i=0}^{m+1}{}' \sum_{j=0}^{n+1}{}' a_{ij} T_i(x) T_j(y)$$

is collocated to a function $g(x, y)$ at $2(m+n)$ points on its boundary ∂D. The latter points are chosen at the zeros of $(1-x^2)U_{m-2}(x)$ on $y = \pm 1$ and at the zeros of $(1-y^2)U_{n-2}(y)$ on $x = \pm 1$, where each of the four corners of the boundary (which occur in both sets of zeros) is only counted once. Investigate whether or not the resulting linear system is singular and determine its maximum rank.

(This question is an analogue of a result in Section 11.8.5, where the zeros of $T_m(x)$ and $T_n(y)$ were adopted.)

3. The diagram shows a square membrane with a slit from the midpoint A of one side to the centre O. We wish to determine solutions of the eigenvalue problem

$$\Delta u + \lambda u = 0 \text{ in } S,$$
$$u = 0 \text{ on } \partial S.$$



Follow the style of Section 11.8.2 to transform the domain and problem into one which may be approximated by Chebyshev collocation. Use the mapping $z' = z^{\frac{1}{2}}$ about O to straighten the cut AOB, determine equations for the mapped (curved) sides of the domain, determine the mapped PDE and deduce the form of approximation u_{mn} to u. Describe the method of solution for λ and u.

[Note: The boundary equation is again $y'(x^2y^2 - r^2 + 1) = 0$ before mapping.]

4. Investigate whether or not there is any gain of efficiency or accuracy in practice in using the Chebyshev form $\sum\sum c_{ij} T_{2i+t}(x/a) T_j(y/b)$ rather than $\sum\sum c_{ij} x^{2i+t} y^j$ in the L-membrane approximation of Section 11.8.2 and, similarly, for the relevant forms in the method for Problem 3 above. Is it possible, for example, to exploit discrete orthogonality in the collocation equations?

5. As a variant on the separation of variables method, consider the solution of

$$\Delta u = f(x, y) \text{ in the ellipse } D : \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1, \qquad (A)$$

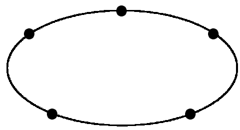
$$u = g(x, y) \text{ on } \partial D : \phi(x, y) \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad (B)$$

where $f \equiv 0$ and g is given explicitly on ∂D.

Then the form

$$u_n = a_0 + \sum_{k=1}^{n} (a_k \cos k\theta + b_k \sin k\theta) r^k,$$

where $x = r\cos\theta$ and $y = r\sin\theta$, satisfies (A) for all coefficients a_0, a_k, b_k. Compute $a_0, \ldots, a_n$, $b_1, \ldots, b_n$ so that (B) is collocated at $2n + 1$ suitably chosen points of ∂D. It is suggested that equal angles of θ should be used on $[0, 2\pi]$; discuss some of the possible choices. What set of points would remain distinct as $b \to 0$, if the ellipse has a small eccentricity?



[Hint: Start at $\theta = \frac{1}{2}\pi/(2n+1)$; the nodes for $n = 2$ are then chosen as in the figure and occur at $\pi/10$, $5\pi/10$, $9\pi/10$, $13\pi/10$, $17\pi/10$. Choose simple non-polynomial functions for g; e.g., $g(x, y) = \cosh(x + y)$.]

6. Repeat Problem 5, but with $g \equiv 0$ and f given explicitly on D, using the Chebyshev polynomial approximation

$$u_{mn} = \phi(x, y). \sum_{i=0}^{m}{}' \sum_{j=0}^{n}{}' a_{ij} T_i(x) T_j(y)$$

and collocating the PDE at a tensor product of the zeros of $T_{m+1}(x/a)$ and $T_{n+1}(y/b)$ on the rectangle

$$R : \{-a \le x \le a; -b \le y \le b\}.$$

Compute results for small choices of m, n.

[Note: This is a method which might be extended to more general boundary $\phi(x, y)$, and ϕ does not need to be a polynomial in x, y.]

7. Generate a set of tables similar to Tables 11.7–11.11 for the odd/odd case $m = n = 5$, showing the 36×36 linear algebraic system for $\{a_{ij}\}$ and the four subsystems derived from this.

8. For $m = n = 3$ (see Section 11.8 above) show that the approximate solution u_{mn} of (11.75) with $g \equiv 0$, given by (11.91) with coefficients (11.119)–(11.122), may be simplified exactly into the form

$$u_{mn} = (1 - x^2)(1 - y^2)[a + bx + cy + dxy].$$

What are the values of a, b, c, d?

Derive u_{mn} directly in this form by collocating the PDE at the Chebyshev zeros. (Note that this method cannot be applied unless $g(x, y) \equiv 0$.)

9. For $m = n = 3$ in (11.75), in the case where g is *not* identically zero, obtain formulae for the coefficients a_{ij} in u_{mn} from Tables 11.8–11.11, namely the four linear subsystems that define them.

CHAPTER 12

Conclusion

In concluding a book that is at least as long as originally intended, we are painfully aware of all that we have left out, even though we have tried to include as much extra material in the Problem sets as we reasonably could. The reader is therefore referred to other texts for more advanced material, including the excellent book of Rivlin (1990), who includes fascinating material on various topics, such as the density of the locations of the zeros of Chebyshev polynomials of all degrees. We were also inspired by the book of Fox & Parker (1968), not only because it made the most up-to-date statement of its time but also, and even more, because it was extremely well written and stimulated major research. We hope that we are following in this tradition, and that there are numerous places in this book that future researchers might take as a starting point.

Bibliography

Abramowitz, M. & Stegun, I. A., Eds. (1964), *Handbook of Mathematical Functions*, number 55 in Applied Mathematics Series, National Bureau of Standards, Washington.

Achieser, N. I. (1956), *Theory of Approximation*, Ungar, New York. (Translated from the Russian).

Adam, G. (1987), Alternative formulation of the extended Clenshaw–Curtis quadrature rules for oscillating integrals, *Rev. Roumaine de Physique* **32**, 813–826.

Adam, G. & Nobile, A. (1991), Product integration rules at Clenshaw–Curtis and related points — a robust implementation, *IMA J. Numer. Analysis* **11**, 271–296.

Anderson, I. J., Cox, M. G. & Mason, J. C. (1995), Tensor-product spline interpolation to data near a family of lines, *Numerical Algorithms* **5**, 193–204.

Atkinson, K. E. (1989), *An Introduction to Numerical Analysis*, 2nd ed., John Wiley, New York.

Baltensprenger, R. & Berrut, J.-P. (1999), The errors in calculating the pseudospectral differentiation matrices for Chebyshev–Gauss–Lobatto points, *Comput. Math. Appl.* **37**, 41–48.

Bayliss, A., Class, A. & Matkowsky, B. J. (1994), Roundoff error in computing derivatives using the Chebyshev differentiation matrix, *J. Comput. Phys.* **116**, 380–383.

Bennell, R. P. (1996), Continuous Approximation Methods for Data Smoothing and Fredholm Integral Equations of the First Kind when the Data are Noisy, Ph.D. thesis, Cranfield University (RMCS).

Bennell, R. P. & Mason, J. C. (1989), Continuous approximation methods for the regularisation and smoothing of integral transforms, *Rocky Mountain J. Math.* **19**, 51–66.

Bennell, R. P. & Mason, J. C. (1991), Bivariate orthogonal polynomial approximation to curves of data, in C. Brezinski et al., Eds., *Orthogonal Polynomials and their Applications*, IMACS, Baltzer, Bussum, Netherlands, pp. 177–183.

Bernardi, C. & Maday, Y. (1992), *Approximations spectrales de problèmes aux limites elliptiques*, Springer-Verlag, Berlin.

Berzins, M. & Dew, P. M. (1987), A note on $\mathcal{C}_0$ Chebyshev methods for parabolic PDEs, *IMA J. Numer. Anal.* **7**, 15–37.

Borel, E. (1905), *Leçons sur les Fonctions de Variables Réelles*, Gautier–Villars, Paris.

Boyd, J. P. (2000), *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover, New York.

Brebbia, C. A., Telles, J. C. F. & Wrobel, L. C. (1984), *Boundary Element Techniques*, Springer-Verlag, Berlin.

Brutman, L. (1978), On the Lebesgue function for polynomial interpolation, *SIAM J. Numer. Anal.* **15**, 694–704.

Brutman, L. (1988), Alternating polynomials associated with the Chebyshev extrema nodes, *J. Approx. Theory* **53**, 33–40.

Brutman, L. (1993), An application of the generalised alternating polynomials to the numerical solution of Fredholm integral equations, *Numerical Algorithms* **5**, 437–442.

Brutman, L. (1997), Lebesgue functions for polynomial interpolation—a survey, *Ann. Numer. Math.* **4**, 111–127.

Burrus, C. S. & Eschenbacher, P. W. (1981), An in-place, in-order prime factor FFT algorithm, *IEEE Trans. on ASSP* **29**, 806–817.

Butzer, P. L. & Jongmans, F. (1999), P. L. Chebyshev [1821–1894], a guide to his life and work, *J. Approx. Theory* **96**, 111–138.

Canuto, C., Hussaini, M. Y., Quarteroni, A. & Zang, T. A. (1988), *Spectral Methods in Fluid Dynamics*, Springer-Verlag, Berlin.

Chalmers, B. L. & Mason, J. C. (1984), Minimal $\mathcal{L}_p$ projections by Fourier, Taylor and Laurent series, *J. Approx. Theory* **40**, 293–297.

Chawla, M. M. (1968), Error estimates for the Clenshaw–Curtis quadrature, *Math. Comp.* **22**, 651–656.

Chebyshev, P. L. (1854), Théorie des mécanismes connus sous le nom de parallélogrammes, *Mémoires présentés à l'Académie Impériale des Sciences de St-Pétersbourg* **VII**, 539–568. [Reprinted in Chebyshev (1899, Vol.I), pages 111–143].

Chebyshev, P. L. (1859), Sur les questions de minima qui se rattachent à la représentation approximative des fonctions, *Mémoires présentés à l'Académie Impériale des Sciences de St-Pétersbourg* **VII**, 199–291. [Reprinted in Chebyshev (1899, Vol.I), pages 273–378].

Chebyshev, P. L. (1874), Sur les fonctions qui différent le moins possible de zéro, *J. Math. Pures et Appl.* **19**, 189–215. [Reprinted in Chebyshev (1899, Vol.II), pages 189–215].

Chebyshev, P. L. (1899), *Oeuvres de P. L. Tchebychef*, St Petersburg.

Cheney, E. W. (1966), *Introduction to Approximation Theory*, McGraw–Hill, New York. (2nd ed. Chelsea, New York, 1982).

Cheney, E. W. & Price, K. H. (1970), Minimal projections, in A. Talbot, Ed., *Approximation Theory (Lancaster, 1969)*, Academic Press, New York, pp. 261–289.

Chisholm, J. S. R. & Common, A. K. (1980), Generalisations of Padé approximation for Chebyshev and Fourier series, in *Proc. 1979 International Christoffel Symposium*, pp. 212–231.

Clenshaw, C. W. (1955), A note on the summation of Chebyshev series, *Math. Tab. Wash.* **9**, 118–120.

Clenshaw, C. W. (1957), The numerical solution of linear differential equations in Chebyshev series, *Proc. Camb. Phil. Soc.* **53**, 134–149.

Clenshaw, C. W. (1959/1960), Curve fitting with a digital computer, *Comput. J.* **2**, 170–173.

Clenshaw, C. W. (1962), *Chebyshev Series for Mathematical Functions*, Vol. 5 of *NPL Math. Tables*, HMSO, London.

Clenshaw, C. W. & Curtis, A. R. (1960), A method for numerical integration on an automatic computer, *Numer. Math.* **2**, 197–205.

Clenshaw, C. W. & Hayes, J. G. (1965), Curve and surface fitting, *J. IMA* **1**, 164–183.

Clenshaw, C. W. & Lord, K. (1974), Rational approximations from Chebyshev series, in B. K. P. Skaife, Ed., *Studies in Numerical Analysis*, Academic Press, New York, pp. 95–113.

Clenshaw, C. W. & Norton, H. J. (1963), The solution of nonlinear ordinary differential equations in Chebyshev series, *Computer J.* **6**, 88–92.

Cooley, J. W. & Tukey, J. W. (1965), An algorithm for the machine calculation of complex Fourier series, *Math. Comput.* **19**, 297–301.

Davis, P. J. (1961), *Interpolation and Approximation*, Blaisdell, New York. (Reprinted Dover, New York, 1975).

Delves, L. M. & Walsh, J., Eds. (1974), *Numerical Solution of Integral Equations*, Clarendon Press, Oxford.

Deville, M. (1984), Recent developments of spectral and pseudospectral methods in fluid dynamics, Lecture notes on computational fluid dynamics, Université Catholique de Louvain, 2 Rue du Levant, B-1348 Louvain-la-neuve, Belgium.

Driscoll, T. A. (1997), Eigenmodes of isospectral drums, *SIAM Rev.* **39**, 1–17.

El-Daou, M. K. & Khajah, H. G. (1997), Iterated solutions of linear operator equations with the tau method, *Math. Comp.* **217**, 207–213.

Elliott, D. (1961), A method for numerical integration of the one-dimensional heat equation using Chebyshev series, *Proc. Camb. Phil. Soc.* **43**, 823–832.

Elliott, D. (1963), A Chebyshev series method for the numerical solution of Fredholm integral equations, *Comp. J.* **6**, 102–111.

Elliott, D. (1965), Truncation errors in two Chebyshev series approximations, *Math. Comp.* **19**, 234–248.

Elliott, D. (1979), The approximate solution of singular integral equations, in M. A. Goldberg, Ed., *Solution Methods for Integral Equations*, Plenum, New York, pp. 83–107.

Elliott, D. (1989), A comprehensive approach to the approximate solution of singular integral equations over the arc $(-1, 1)$, *J. Integral Eq. and Appl.* **2**, 59–94.

Erdélyi, A., Magnus, W., Oberhettinger, F. & Tricomi, F. G. (1953), *Higher Transcendental Functions*, McGraw–Hill, New York.

Erdélyi, A., Magnus, W., Oberhettinger, F. & Tricomi, F. G. (1954), *Tables of Integral Transforms*, McGraw–Hill, New York.

Erdös, P. & Feldheim, E. (1936), Sur le mode de convergence pour l'interpolation de Lagrange, *Comptes Rendus* **203**, 913–915.

Erdös, P. & Turán, P. (1937), On interpolation, *Annals of Math.* **38**, 142–155.

Evans, G. A., Hyslop, J. & Morgan, A. P. G. (1981), Iterative solution of Volterra integral equations using Clenshaw–Curtis quadrature, *J. Comp. Phys.* **40**, 64–76.

Fairweather, D. & Karageorghis, A. (1998), The method of fundamental solutions for elliptic boundary value problems, *Adv. Comp. Math.* **9**, 69–95.

Favati, P., Lotti, G. & Romani, F. (1993), Bounds on the error of Fejér and Clenshaw–Curtis type quadrature for analytic functions, *Appl. Math. Letters* **6**, 3–8.

Fejér, L. (1904), Untersuchen über Fouriersche Reihen, *Math. Annalen* **58**, 51–69.

Filippi, S. (1964), Angenäherte Tschebyscheff-Approximation einer Stammfunktion — eine Modifikation des Verfahrens von Clenshaw und Curtis, *Numer. Math.* **6**, 320–328.

Finlayson, B. A. & Scriven, L. E. (1966), The method of weighted residuals — a review, *Appl. Mech. Rev.* **19**, 735–748.

Fornberg, B. (1996), *A Practical Guide to Pseudospectral Methods*, CUP, Cambridge.

Fornberg, B. & Sloan, D. M. (1994), A review of pseudospectral methods for solving partial differential equations, in A. Iserles, Ed., *Acta Numerica 1994*, Cambridge, pp. 203–267.

Forsythe, G. E. (1957), Generation and use of orthogonal polynomials for data-fitting with a digital computer, *J. SIAM* **5**, 74–88.

Fox, L. (1962), Chebyshev methods for ordinary differential equations, *Comp. J.* **4**, 318–331.

Fox, L. (1966*a*), Chebyshev least-squares approximation, in D. C. Handscomb, Ed., *Methods of Numerical Approximation*, Pergamon, Oxford, pp. 39–46.

Fox, L. (1966*b*), Determination and properties of Chebyshev expansions, in D. C. Handscomb, Ed., *Methods of Numerical Approximation*, Pergamon, Oxford, pp. 47–60.

Fox, L., Henrici, P. & Moler, C. (1967), Approximations and bounds for eigenvalues of elliptic operators, *SIAM J. Numer. Anal.* **4**, 89–102.

Fox, L. & Parker, J. B. (1968), *Chebyshev Polynomials in Numerical Analysis*, OUP, Oxford. [Revised 2nd edition, 1972].

Fraser, W. & Wilson, M. W. (1966), Remarks on the Clenshaw–Curtis quadrature scheme, *SIAM Review* **8**, 322–327.

Freilich, J. H. & Mason, J. C. (1971), Best and near-best $\mathcal{L}_1$ approximations by Fourier series and Chebyshev series, *J. Approx. Theory* **4**, 183–193.

Freilich, J. H. & Ortiz, E. L. (1982), Numerical solution of systems of ordinary differential equations with the tau method: An error analysis, *Math. Comp.* **39**, 467–479.

Fromme, J. A. & Golberg, M. A. (1979), Numerical solution of a class of integral equations arising in two-dimensional aerodynamics, in M. A. Golberg, Ed., *Solution Methods for Integral Equations*, Plenum, New York, pp. 109–146.

Fromme, J. A. & Golberg, M. A. (1981), Convergence and stability of a collocation method for the generalized airfoil equation, *Appl. Math. Comput.* **8**, 281–292.

Gautschi, W. (1984), Questions of numerical condition related to polynomials, in *Studies in Numerical Analysis*, number 24 in MAA Stud. Math., Math. Assoc. America, pp. 140–177.

Gautschi, W. (1992), On mean convergence of extended Lagrange interpolation, *J. Comp. Appl. Math.* **43**, 19–35.

Geddes, K. O. (1978), Near-minimax polynomial approximation in an elliptical region, *SIAM J. Numer. Anal.* **15**, 1228–1233.

Gentleman, W. M. (1972), Implementing Clenshaw–Curtis quadrature, *Comm. ACM* **15**, 337–346.

Gentleman, W. M. & Sande, G. (1966), Fast Fourier transforms for fun and profit, in *Proceedings of the AFIPS 1966 Fall Joint Computer Conference*, pp. 563–578.

Gerasoulis, A. (1986), The singular value decomposition of the Gauss–Lobatto and Lobatto–Chebyshev methods for Cauchy singular integral equations, *Computers and Maths with Applications* **12A**, 895–907.

Gerritsma, M. I. & Phillips, T. N. (1998), Discontinuous spectral element approximations for the velocity–pressure–stress formulation of the Stokes problem, *J. Numer. Methods Engrg.* **43**, 1401–1419.

Gerritsma, M. I. & Phillips, T. N. (1999), Compatible spectral approximations for the velocity–pressure–stress formulation of the Stokes problem, *SIAM J. Sci. Comp.* **20**, 1530–1550.

Gladwell, G. M. L. & England, A. H. (1977), Orthogonal polynomial solutions to some mixed boundary-value problems in elasticity theory, *Quart. J. Mech. Appl. Math.* **30**, 175–185.

Golberg, M. A., Ed. (1990), *Numerical Solution of Integral Equations*, Plenum, New York.

Golub, G. H. & van Loan, C. F. (1983), *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore.

Gottlieb, D. & Lustman, L. (1983), The spectrum of the Chebyshev collocation operator for the heat equation, *SIAM J. Numer. Anal.* **20**, 909–921.

Gottlieb, D. & Orszag, S. A. (1977), *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM, Philadelphia.

Gragg, W. B. (1977), Laurent, Fourier and Chebyshev–Padé tables, in E. B. Saff & R. S. Varga, Eds., *Padé and Rational Approximation*, Academic Press, New York, pp. 61–72.

Gragg, W. B. & Johnson, G. D. (1974), The Laurent–Padé table, in *Information Processing 1974*, North–Holland, Amsterdam, pp. 632–637.

Groetsch, C. W. (1984), *The Theory of Tikhonov Regularisation for Fredholm Integral Equations of the First Kind*, Vol. 105 of *Research Notes in Mathematics*, Pitman, London.

Grunwald, G. (1941), On a convergence theorem for the Lagrange interpolation polynomials, *Bull. Amer. Math. Soc.* pp. 271–275.

Guo Ben-yu (1998), *Spectral Methods and their Applications*, World Scientific, Singapore.

Guo Ben-yu (1999), Error estimation for Hermite spectral method for nonlinear partial differential equations, *Math. Comp.* **68**, 1067–1078.

Guo Ben-yu & He Li-ping (1998), The fully discrete Legendre spectral approximation of two-dimensional unsteady incompressible fluid flow in stream function form, *SIAM J. Numer. Anal.* **35**, 146–176.

Guo Ben-yu & Jie Shen (2000), Laguerre–Galerkin method for nonlinear partial differential equations on a semi-infinite interval, *Numer. Math.* **86**, 635–654.

Gutknecht, M. H. & Trefethen, L. N. (1982), Real polynomial Chebyshev approximation by the Carathéodory-Fejér method, *SIAM J. Numer. Anal.* **19**, 358–371.

Haidvogel, D. B. & Zang, T. (1979), The accurate solution of Poisson's equation by expansion in Chebyshev polynomials, *J. Comp. Phys.* **30**, 167–180.

Handscomb, D. C. (1966), Functions of many variables, in D. C. Handscomb, Ed., *Methods of Numerical Approximation*, Pergamon, Oxford, chapter 23, pp. 191–194.

Handscomb, D. C. (1973), The relative sizes of the terms in Chebyshev and other ultraspherical expansions, *J. Inst. Maths Applics* **11**, 241–246.

Hardy, G. H., Littlewood, J. E. & Pólya, G. (1952), *Inequalities*, second ed., CUP, Cambridge.

Hasegawa, T. & Torii, T. (1991), An automatic quadrature for Cauchy principal value integrals, *Math. Comp.* **56**, 741–754.

Henrici, P. (1974–1986), *Applied and Computational Complex Analysis*, Vol. 1–3, John Wiley, New York.

Hobson, E. W. (1926), *The Theory of Functions of a Real Variable and the Theory of Fourier Series*, Vol. 2, 2nd ed., CUP, Cambridge.

Huang, W.-Z. & Sloan, D. M. (1994), The pseudospectral method for solving differential eigenvalue equations, *J. Comput. Phys.* **111**, 399–409.

Jaswon, M. A. & Symm, G. (1977), *Integral Equation Methods in Potential Theory and Elastostatics*, Academic Press, New York.

Jorgens, K. (1970), *Lineare Integraloperatoren*, Teubner, Stuttgart.

Jovanovski, V. (1999), Three-dimensional Imaging and Analysis of the Morphology of Oral Structures from Co-ordinate Data, Ph.D. thesis, Royal London School of Dentistry, Queen Mary & Westfield College, London.

Khajah, H. G. (1997), Tau method approximation of the Gauss hypergeometric function, *Comp. Rend. de l'Acad. Bulgare des Sci.* **50**, 13–16.

Khajah, H. G. & Ortiz, E. L. (1991), Upper and lower error estimation for the tau method and related polynomial techniques, *Comput. Math. Appl.* **3**, 81–87.

Knibb, D. & Scraton, R. E. (1979), A note on the numerical solution of nonlinear parabolic equations in Chebyshev series, *Int. J. Comput. Maths* **7**, 217–225.

Kolba, D. P. & Parks, T. W. (1977), A prime factor FFT algorithm using high speed convolution, *IEEE Trans. on ASSP* **25**, 281–294.

Kreiss, H.-O. & Oliger, J. (1972), Comparison of accurate methods for the integration of hyperbolic systems, *Tellus* **24**, 199–215.

Lakin, W. D. (1986), Differentiation matrices for arbitrarily spaced grid points, *J. Numer. Methods Engrg.* **23**, 209–218.

Lanczos, C. (1938), Trigonometric interpolation of empirical and analytical functions, *J. Math. Phys.* **17**, 123–199.

Lanczos, C. (1952), Introduction, in (NBS 1952).

Lanczos, C. (1957), *Applied Analysis*, Prentice–Hall, Englewood Cliffs, NJ.

Light, W. A. (1978), A comparison between Chebyshev and ultraspherical expansions, *J. Inst. Maths Applics* **21**, 455–460.

Light, W. A. (1979), Some optimality conditions for Chebyshev expansions, *J. Approx. Theory* **27**, 113–126.

Light, W. A. (1980), Nonlinear families of projections on $\mathcal{C}[-1, 1]$, *J. Approx. Theory* **30**, 197–202.

Linz, P. (1977), An analysis of a method for solving singular integral equations, *BIT* **17**, 329–337.

Locher, F. (1969), Fehlerabschätzung für das Quadraturverfahren von Clenshaw und Curtis, *Computing* **4**, 304–315.

Luttman, F. W. & Rivlin, T. J. (1965), Some numerical experiments in the theory of polynomial interpolation, *IBM J. Develop.* **9**, 187–191.

Maehly, H. J. (1960), Rational approximations for transcendental functions, in *Information Processing*, Butterworth, London, pp. 57–62.

Martin, P. A. (1991), End point behaviour of solutions to hypersingular integral equations, *Proc. Roy. Soc. (London)* **432**, 301–320.

Martin, P. A. (1992), Exact solution of a simple hypersingular integral equation, *J. Int. Eqns and Applics* **4**, 197–204.

Mason, J. C. (1965), Some new Approximations for the Solution of Differential Equations, D.phil. thesis, Oxford University.

Mason, J. C. (1967), Chebyshev polynomial approximations for the L-membrane eigenvalue problem, *SIAM J. Appl. Math.* **15**, 171–186.

Mason, J. C. (1969), Chebyshev methods for separable partial differential equations, in *Information Processing 68*, Vol. 1, North-Holland, Amsterdam, pp. 164–169.

Mason, J. C. (1970), Orthogonal polynomial approximation methods in numerical analysis, in A. Talbot, Ed., *Approximation Theory (Lancaster, 1969)*, Academic Press, New York, pp. 17–33.

Mason, J. C. (1978), Near-best $\mathcal{L}_\infty$ and $\mathcal{L}_1$ approximations to analytic functions on two-dimensional regions, in D. C. Handscomb, Ed., *Multivariate Approximation*, Academic Press, New York, pp. 115–135.

Mason, J. C. (1980), Near-best multivariate approximation by Fourier series, Chebyshev series and Chebyshev interpolation, *J. Approx. Theory* **28**, 349–358.

Mason, J. C. (1982), Minimal projections and near-best approximations by multivariate polynomial expansion and interpolation, in W. Schempp & K. Zeller, Eds., *Multivariate Approximation 2*, number 17 in ISNM, Birkhäuser, Basel, pp. 241–254.

Mason, J. C. (1983*a*), Minimal $\mathcal{L}_p$ projections and near-best $\mathcal{L}_p$ approximations, in C. K. Chui, L. L. Schumaker & J. D. Ward, Eds., *Approximation Theory 4*, Academic Press, New York, pp. 605–610.

Mason, J. C. (1983*b*), Near-best $\mathcal{L}_p$ approximation by real and complex Chebyshev series, *IMA J. Numer. Analysis* **3**, 493–504.

Mason, J. C. (1984), Some properties and applications of Chebyshev polynomial and rational approximations, in P. R. Graves-Morris, E. B. Saff & R. S. Varga, Eds., *Rational Approximation and Interpolation*, number 1105 in Lecture Notes in Mathematics, Springer-Verlag, Berlin, pp. 27–48.

Mason, J. C. (1988), The $\mathcal{L}_p$ minimality and near-minimality of orthogonal polynomial approximation and integration methods, in M. Alfaro et al., Eds., *Orthogonal Polynomials and their Applications*, number 1329 in Lecture notes in Mathematics, Springer-Verlag, Berlin, pp. 291–299.

Mason, J. C. (1993), Chebyshev polynomials of the second, third and fourth kinds in approximation, indefinite integration and integral transforms, *J. Comp. and Appl. Maths* **49**, 169–178.

Mason, J. C. (1995), Minimality properties and applications of four kinds of Chebyshev polynomials, in W. M. Müller, M. Felten & D. H. Mache, Eds., *Approximation Theory*, Akademic Verlag, Berlin, pp. 231–250.

Mason, J. C. & Chalmers, B. L. (1984), Near-best $\mathcal{L}_p$ approximation by Fourier, Taylor and Laurent series, *IMA J. Numer. Analysis* **4**, 1–8.

Mason, J. C. & Crampton, A. (2002), Applications of orthogonalisation procedures for Gaussian radial basis functions and Chebyshev polynomials, in J. Levesley, I. J. Anderson & J. C. Mason, Eds., *Algorithms for Approximation 4*, University of Huddersfield, Huddersfield, UK. (To appear).

Mason, J. C. & Elliott, G. H. (1993), Near-minimax complex approximation by four kinds of Chebyshev polynomial expansion, *J. Comp. and Appl. Maths* **46**, 291–300.

Mason, J. C. & Elliott, G. H. (1995), Constrained near-minimax approximations by weighted expansion and interpolation using Chebyshev polynomials of the second, third and fourth kinds, *Numerical Algorithms* **9**, 39–54.

Mason, J. C. & Upton, N. K. (1989), Linear algorithms for transformed linear forms, in C. K. Chui, L. L. Schumaker & J. D. Ward, Eds., *Approximation Theory 6*, Vol. 2, Academic Press, New York, pp. 417–420.

Mason, J. C. & Venturino, E. (1996), Integration methods of Clenshaw–Curtis type, based on four kinds of Chebyshev polynomials, in G. Nürnberger, J. W. Schmidt & G. Walz, Eds., *Multivariate Approximation and Splines*, Birkhäuser, Basel, pp. 153–165.

Mason, J. C. & Venturino, E. (1997), Numerical analysis of a boundary hypersingular integral equation, in L. Elliott, D. B. Ingham & D. Lesnic, Eds., *First UK Conference on Boundary Integral Methods*, Leeds Univ. Press, Leeds, pp. 224–235.

Mason, J. C. & Venturino, E. (2002), A fully-discrete Chebyshev–Galerkin method for Hadamard finite-part integral equations, *J. Integral Equations* . (To appear).

Mason, J. C. & Weber, R. O. (1992), Particular solution methods for free and moving boundary problems, *Proc. Centre for Maths and its Appl., Australian National University* **30**, 87–107.

Mayers, D. F. (1966), Convergence of polynomial interpolation, in D. C. Handscomb, Ed., *Methods of Numerical Approximation*, Pergamon, Oxford, chapter 3, pp. 15–26.

Mercier, B. (1989), *An Introduction to the Numerical Analysis of Spectral Methods*, number 318 in Lecture Notes in Physics, Springer-Verlag, Berlin.

Mhaskar, H. N. & Pai, D. V. (2000), *Fundamentals of Approximation Theory*, Narosa Publishing Co., Delhi. (Available from Alpha Beta Publishing Co. in Europe and from CRC Press in the U.S.A.).

NBS (1952), *Tables of Chebyshev Polynomials $S_n(x)$ and $C_n(x)$*, number 9 in Applied Mathematics Series, National Bureau of Standards, Washington.

Novak, E. & Ritter, K. (1996), High-dimensional integration of smooth functions over cubes, *Num. Math.* **75**, 79–97.

NPL (1961), *Modern Computing Methods*, number 16 in National Physical Laboratory: Notes on Applied Science, HMSO, London.

O'Hara, H. & Smith, F. S. (1968), Error estimation in the Clenshaw–Curtis quadrature formula, *Computer J.* **11**, 213–219.

Ortiz, E. L. (1969), The tau method, *SIAM J. Numer. Anal.* **6**, 480–492.

Ortiz, E. L. (1980), Polynomial and rational approximation of boundary layer problems with the tau method, in J. J. H. Miller, Ed., *Boundary and Interior Layers — Computational and Asymptotic Methods*, Boole Press, Dun Laoghaire, pp. 387–391.

Ortiz, E. L. (1987), Singularity treatment in the bidimensional tau method with an application to problems defined on L-shaped domains, in J. Ballmann, Ed., *Effiziente numerische Verfahren für partielle Differentialgleichungen*, Teubner, Stuttgart.

Ortiz, E. L. & Pun, K.-S. (1986), A bi-dimensional tau-elements method for the numerical solution of nonlinear partial differential equations with an application to Burgers' equation, *Computers and Maths with Applications* **12**, 1225–1240.

Padé, H. (1892), Sur la representation approchée d'une fonction par des fonctions rationelles, *Ann. Ec. Norm. Sup.* **9**, 1–93.

Peherstorfer, F. (1997), Minimal polynomials on several intervals with respect to the maximum norm—a survey, in *Complex methods in approximation theory*, number 2 in Monogr. Cienc. Tecnol., Univ. Almeria, pp. 137–159.

Peyret, R. (1986), *Introduction to Spectral Methods*, Von Karman Institute, Rhode-St-Genèse, Belgium.

Phillips, T. N. & Owens, R. G. (1997), A mass-conserving multi-domain spectral collocation method for the Stokes problem, *Computers and Fluids* **26**, 825–840.

Phillips, T. N., Zang, T. A. & Hussaini, M. Y. (1986), Preconditioners for the spectral multigrid method, *IMA J. Numer. Anal.* **6**, 273–292.

Piessens, R. & Branders, M. (1983), Modified Clenshaw–Curtis method for the computation of Bessel-function integrals, *BIT* **23**, 370–381.

Piessens, R. & Branders, M. (1992), On the computation of Fourier transforms of singular functions, *J. Comp. Appl. Math.* **43**, 159–169.

Powell, M. J. D. (1967), On the maximum errors of polynomial approximations defined by interpolation and by least squares criteria, *Comput. J.* **9**, 404–407.

Reid, J. K. & Walsh, J. E. (1965), An elliptic eigenvalue problem for a reentrant region, *SIAM J. Appl. Math.* **13**, 837–850.

Rice, J. R. (1964), *The Approximation of Functions I — Linear Theory*, Addison–Wesley, Reading, MA.

Rivlin, T. J. (1974), *The Chebyshev Polynomials*, John Wiley, New York.

Rivlin, T. J. (1990), *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*, John Wiley, New York. (2nd ed. of Rivlin (1974)).

Rodriguez, G. & Seatzu, S. (1990), Numerical solutions of the finite moment problem in a reproducing kernel Hilbert space, *J. Comp. Appl. Math.* **33**, 233–244.

Sansone, G. (1959), *Orthogonal Functions*, Interscience, New York.

Shabat, G. B. & Voevodskii, V. A. (1990), Drawing curves over number fields, in *The Grothendieck Festschrift, Volume III*, number 88 in Progr. Math., Birkhäuser, Basel, pp. 199–227.

Shabat, G. B. & Zvonkin, A. (1994), Plane trees and algebraic numbers, in *Jerusalem Combinatorics '93*, number 178 in Contemp. Math., Amer. Math. Soc., Providence, RI, pp. 233–275.

Sloan, I. H. & Smith, W. E. (1978), Product integration with the Clenshaw–Curtis and related points: Convergence properties, *Numer. Math.* **30**, 415–428.

Sloan, I. H. & Smith, W. E. (1980), Product integration with the Clenshaw–Curtis points: Implementation and error estimates, *Numer. Math.* **34**, 387–401.

Sloan, I. H. & Smith, W. E. (1982), Properties of interpolatory integration rules, *SIAM J. Numer. Anal.* **19**, 427–442.

Sloan, I. H. & Stephan, E. P. (1992), Collocation with Chebyshev polynomials for Symm's integral equation on an interval, *J. Austral. Math. Soc. Ser. B* **34**, 199–211.

Smith, H. V. (1982), Global error bounds for the Clenshaw–Curtis quadrature formula, *BIT* **22**, 395–398.

Smith, W. E. & Paget, D. (1992), Optimal nodes for interpolatory product integration, *SIAM J. Numer. Anal.* **29**, 586–600.

Smolyak, S. A. (1963), Quadrature and interpolation formulas for tensor products of certain classes of functions, *Soviet Math. Dokl.* **4**, 240–243.

Snyder, M. A. (1966), *Chebyshev Methods in Numerical Approximation*, Prentice–Hall, Englewood Cliffs, NJ.

Srivastav, R. P. (1983), Numerical solution of singular integral equations using Gauss-type formulae — I. Quadrature and collocation on Chebyshev nodes, *IMA J. Numer. Analysis* **3**, 305–318.

Srivastav, R. P. & Jen, E. (1983), Numerical solution of singular integral equations using Gauss-type formulae — II. Lobatto–Chebyshev quadrature and collocation on Chebyshev nodes, *IMA J. Numer. Analysis* **3**, 319–325.

Symm, G. (1966), An integral equation method in conformal mapping, *Num. Math.* pp. 250–258.

Tikhonov, A. N. (1963*a*), Regularisation of incorrectly posed problems, *Soviet Math. Dokl.* **4**, 1624–1627.

Tikhonov, A. N. (1963*b*), Solution of incorrectly formulated problems and the regularisation method, *Soviet Math. Dokl.* **4**, 1035–1038.

Timan, A. F. (1963), *Theory of Approximation of Functions of a Real Variable*, Pergamon, Oxford. (Translated from Russian original of 1960).

Trefethen, L. N. (2000), *Spectral Methods in MATLAB*, SIAM, Philadelphia.

Trefethen, L. N. & Gutknecht, M. H. (1987), Padé, stable Padé and Chebyshev–Padé approximation, in J. C. Mason & M. G. Cox, Eds., *Algorithms for Approximation*, OUP, Oxford, pp. 227–264.

Trefethen, L. N. & Trummer, M. R. (1987), An instability phenomenon in spectral methods, *SIAM J. Numer. Anal.* **24**, 1008–1023.

Trefethen, L. N. & Weideman, J. A. C. (1991), Two results on polynomial interpolation in equally spaced points, *J. Approx. Theory* **65**, 247–260.

Tricomi, F. G. (1957), *Integral Equations*, Interscience.

Turetskii, A. H. (1940), The bounding of polynomials prescribed at equally distributed points, *Proc. Pedag. Inst. Vitebsk* **3**, 117–127. (Russian).

Urabe, M. (1966), Galerkin's procedure for nonlinear periodic systems and its extension to multipoint boundary-value problems for general nonlinear systems, in D. Greenspan, Ed., *Numerical Solution of Nonlinear Differential Equations*, John Wiley, New York, pp. 297–327.

van Loan, C. (1992), *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia.

Vandeven, H. (1990), On the eigenvalues of second-order spectral differentiation operators, *Comput. Methods Appl. Mech. Engrg.* **80**, 313–318.

Varga, R. S. (1962), *Matrix Iterative Analysis*, Prentice–Hall, Englewood Cliffs, NJ.

Vekua, I. N. (1967), *New Methods for Solving Elliptic Equations*, North–Holland, Amsterdam.

Venturino, E. (1986), Recent developments in the numerical solution of singular integral equations, *J. Math. Anal. & Appl.* **115**, 239–277.

Venturino, E. (1992), Unconventional solution of simple integral equations, *J. Integral Eq. and Appl.* **4**, 443–463.

Venturino, E. (1993), Simple quadrature for singular integrals, *J. Integral Eq. and Appl.* **5**, 257–276.

Voigt, R. G., Gottlieb, D. & Hussaini, M. Y., Eds. (1984), *Spectral Methods for Partial Differential Equations*, SIAM, Philadelphia.

Wahba, G. (1977), Practical approximate solutions to linear operator equations when the data are noisy, *SIAM J. Numer. Anal.* **14**, 652–667.

Weideman, J. A. C. & Trefethen, L. N. (1988), The eigenvalues of second-order spectral differentiation matrices, *SIAM J. Numer. Anal.* **25**, 1279–1298.

Youschkevitch, A. P. (1981), 'Chebyshev', in *Dictionary of Scientific Biography*, Vol. 3, Scribner, New York, pp. 222–232.

Zygmund, A. (1959), *Trigonometric Series*, 2nd ed., CUP, Cambridge.

Appendix A

Biographical Note

Pafnuty Lvovich Chebyshev was born in Okatovo in the Kaluga region of Russia on 16th May [4th May, Old Style] 1821.

He studied mathematics at Moscow University from 1837 to 1846, then moved to St Petersburg (as it then was and has now again become), where he became an assistant professor at Petersburg University in 1847 and a full professor in 1851, in which post he remained until 1882. It is he who was principally responsible for founding, directing and inspiring the 'Petersburg school' of mathematical research, noted for its emphasis on drawing its problems for study from practical necessities rather than from mere intellectual curiosity. He was elected a foreign associate of the Institut de France in 1874, and a fellow of the Royal Society of London in 1877.

He worked in many fields outside approximation theory, including number theory (the distribution of primes), integration of algebraic functions, geometric theory of hinge mechanisms (the subject which led to his special interest in minimax approximation of functions), the moment problem, quadrature formulae and probability theory (limit theorems).

The Chebyshev polynomials T_n which now bear his name (the symbol 'T' deriving from its continental transliterations as 'Tchebycheff', 'Tschebyscheff' &c.) were first introduced by him in a paper on hinge mechanisms (Chebyshev 1854) presented to the St Petersburg Academy in 1853. They were discussed in more mathematical depth in a second paper (Chebyshev 1859) presented in 1857; see also (Chebyshev 1874). Somewhat surprisingly, in the light of what seems today the obvious connection with Fourier theory, his discussion makes no use of the substitution $x = \cos\theta$.

He died in St Petersburg on 8th December [26th November, Old Style] 1894.

A much more extensive biography, from which these facts were extracted, is to be found in the *Dictionary of Scientific Biography* (Youschkevitch 1981). See also a recent article by Butzer & Jongmans (1999).

Appendix B

Summary of Notations, Definitions and Important Properties

B.1 Miscellaneous notations

$\sum'$	finite or infinite summation with first (T_0) term halved, $\sum_{r=0}^{\infty}{}' a_r T_r = \frac{1}{2}a_0 T_0 + a_1 T_1 + a_2 T_2 + \cdots$
$\sum''$	finite summation with first and last terms halved, $\sum_{r=0}^{n}{}'' a_r T_r = \frac{1}{2}a_0 T_0 + a_1 T_1 + \cdots + a_{n-1} T_{n-1} + \frac{1}{2}a_n T_n$
$\sum^*$	finite summation with last term halved, $\sum_{r=1}^{n}{}^{*} a_r P_r = a_1 P_1 + \cdots + a_{n-1} P_{n-1} + \frac{1}{2}a_n P_n$
$\oint$	integral round a closed contour
$⨍$	Cauchy principal value integral
$\lfloor \cdots \rfloor$	largest integer $\leq \cdots$
$\lVert \cdot \rVert$	a norm (see page 43)
$\langle \cdot\,, \cdot \rangle$	an inner product (see pages 72, 97)
$\mathcal{A}(D)$	the linear space of functions analytic on the (complex) domain D and continuous on its closure $\bar{D}$
$B_n f$	the minimax nth degree polynomial approximation to f on the interval $[-1, 1]$
$\mathcal{C}[a, b]$	the linear space of functions continuous on the interval $[a, b]$
$\mathcal{C}^n[a, b]$	the linear space of functions continuous and having n continuous derivatives on the interval $[a, b]$
$\mathcal{C}^0_{2\pi}$	the linear space of continuous periodic functions with period 2π
$\mathcal{C}^0_{2\pi,\mathrm{e}}$	the subspace of $\mathcal{C}^0_{2\pi}$ consisting of even functions only
C_r	the circular contour $\{w : \lvert w \rvert = r\}$ in the complex plane
D_r	the elliptic domain $\{z : 1 \leq \lvert z + \sqrt{z^2 - 1} \rvert < r\}$
E_r	the elliptic contour $\{z : \lvert z + \sqrt{z^2 - 1} \rvert = r\}$

= the image of C_r under $z = \frac{1}{2}(w + w^{-1})$

$J_n f$	the nth degree polynomial interpolating f at $n+1$ given points
$\mathcal{L}_p[a, b]$	the linear space of functions on $[a, b]$ on which the norm $\\|\cdot\\|_p$ can be defined
Π_n	the linear space of polynomials of degree n
$S_n^F f$	the nth partial sum of the Fourier expansion of f
$S_n^{FC} f$	the nth partial sum of the Fourier cosine expansion of f
$S_n^{FS} f$	the nth partial sum of the Fourier sine expansion of f
$S_n^T f$	the nth partial sum of the first-kind Chebyshev expansion of f
λ_n	Lebesgue constant (see page 125)
$\omega(\delta)$	the modulus of continuity of a function (see page 119)
∂S	the boundary of the two-dimensional domain S

B.2 The four kinds of Chebyshev polynomial

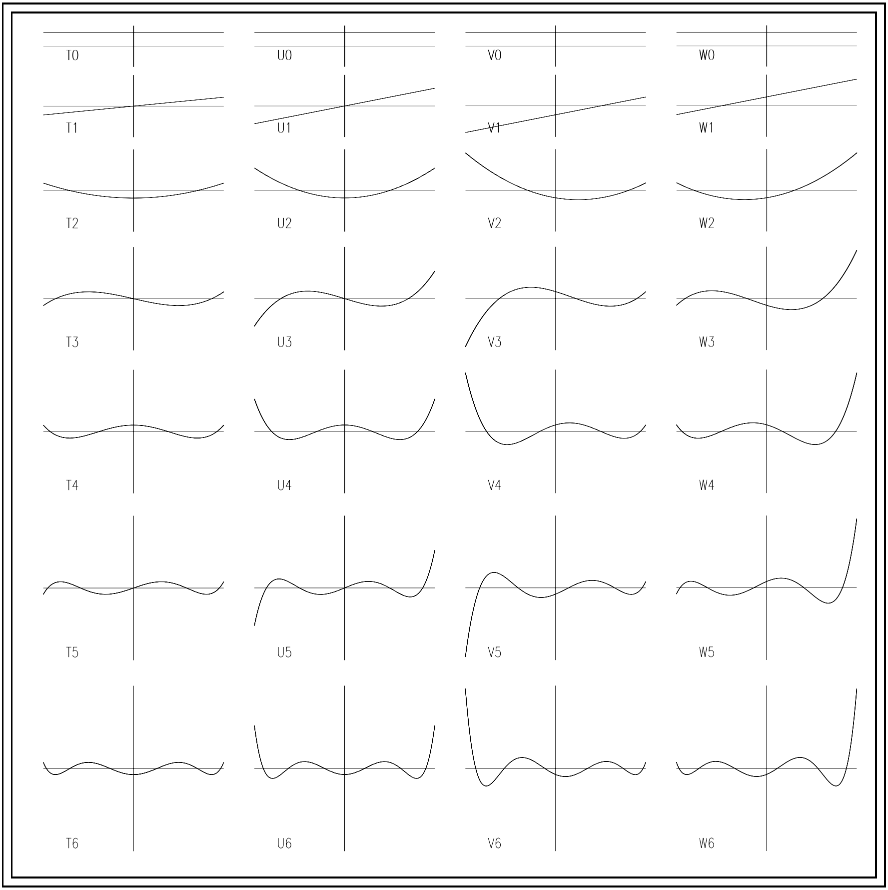


Figure B.1: Plots of the four kinds of Chebyshev polynomial: $T_n(x)$, $U_n(x)$, $V_n(x)$, $W_n(x)$ for values of x in the range $[-1, 1]$ and n running from 0 to 6

Table B.2: Key properties of the four kinds of Chebyshev polynomial

kind	1st	2nd	3rd	4th
$P_n =$	T_n	U_n	V_n	W_n
$P_n(\cos(\theta)) =$	$\cos n\theta$	$\dfrac{\sin(n+1)\theta}{\sin\theta}$	$\dfrac{\cos(n+\frac{1}{2})\theta}{\cos\frac{1}{2}\theta}$	$\dfrac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}$
$P_n(\frac{1}{2}(w+w^{-1})) =$	$\frac{1}{2}(w^n+w^{-n})$	$\dfrac{w^{n+1}-w^{-n-1}}{w-w^{-1}}$	$\dfrac{w^{n+\frac{1}{2}}+w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}}+w^{-\frac{1}{2}}}$	$\dfrac{w^{n+\frac{1}{2}}-w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}}-w^{-\frac{1}{2}}}$
$P_0(x) =$	1			
$P_1(x) =$	x	$2x$	$2x-1$	$2x+1$
recurrence	$P_n(x) = 2xP_{n-1}(x) - P_{n-2}(x)$			
x^n coefficient	$2^{n-1}\ (n>0)$	2^n		
zeros	$x_{k,n} := \cos\dfrac{(k-\frac{1}{2})\pi}{n}$	$\cos\dfrac{k\pi}{n+1}$	$\cos\dfrac{(k-\frac{1}{2})\pi}{n+\frac{1}{2}}$	$\cos\dfrac{k\pi}{n+\frac{1}{2}}$
extrema	$y_{k,n} := \cos\dfrac{k\pi}{n}$	no closed form		
$\Vert P_n \Vert_\infty =$	1	$n+1$	$2n+1$	

Table B.3: Orthogonality properties of the four kinds of Chebyshev polynomial

kind	1st	2nd	3rd	4th
$P_n =$	T_n	U_n	V_n	W_n
weight $w(x) =$	$\dfrac{1}{\sqrt{1-x^2}}$	$\sqrt{1-x^2}$	$\sqrt{\dfrac{1+x}{1-x}}$	$\sqrt{\dfrac{1-x}{1+x}}$
orthogonality	$\langle P_m\,,\ P_n \rangle = \int_{-1}^{1} w(x) P_m(x) P_n(x)\,\mathrm{d}x = 0 \ (m \neq n)$			
$\langle P_n\,,\ P_n \rangle =$	$\frac{1}{2}\pi \ \ (n > 0)$	$\frac{1}{2}\pi$	π	
contour orthogonality	$\langle P_m\,,\ P_n \rangle = \oint_{E_r} P_m(z)\overline{P_n(z)}\,\lvert w(z)\,\mathrm{d}z \rvert = 0 \ (m \neq n)$ $[E_r =$ locus of $\frac{1}{2}(r\mathrm{e}^{\mathrm{i}\theta} + r^{-1}\mathrm{e}^{-\mathrm{i}\theta})]$			
$\langle P_n\,,\ P_n \rangle =$	$\frac{1}{2}\pi(r^{2n} + r^{-2n}) \ \ (n > 0)$	$\frac{1}{2}\pi(r^{2n+2} + r^{-2n-2})$	$\pi(r^{2n+1} + r^{-2n-1})$	

Table B.4: Discrete orthogonality of the four kinds of Chebyshev polynomial

kind	1st	2nd	3rd	4th
$P_n =$	T_n	U_n	V_n	W_n
weight $w(x) =$	$\dfrac{1}{\sqrt{1-x^2}}$	$\sqrt{1-x^2}$	$\sqrt{\dfrac{1+x}{1-x}}$	$\sqrt{\dfrac{1-x}{1+x}}$
abscissae	$x_{k,N+1} = \cos\{(k-\frac{1}{2})\pi/(N+1)\}$			
discrete orthogonality	$\langle P_m\,,\,P_n\rangle = \sum_{k=1}^{N+1} P_m(x_{k,N+1})P_n(x_{k,N+1})w(x_{k,N+1})\sqrt{1-x_{k,N+1}^2} = 0\ (m \neq n \leq N)$			
$\langle P_n\,,\,P_n\rangle =$	$\frac{1}{2}(N+1)\ \ (0 < n \leq N)$	$\frac{1}{2}(N+1)$	$(N+1)$	
abscissae	$y_{k,N} = \cos\{k\pi/N\}$			
discrete orthogonality	$\langle P_m\,,\,P_n\rangle = {\sum_{k=0}^{N}}'' P_m(y_{k,N})P_n(y_{k,N})w(y_{k,N})\sqrt{1-y_{k,N}^2} = 0\ (m \neq n \leq N)$			
$\langle P_n\,,\,P_n\rangle =$	$\frac{1}{2}N\ \ (0 < n < N)$	$\frac{1}{2}N$	N	

Appendix C

Tables of Coefficients

Each of the following five Tables may be used in two ways, to give the coefficients of two different kinds of shifted or unshifted polynomials.

Table C.1: Coefficients of x^k in $V_n(x)$ and of $(-1)^{n+k}x^k$ in $W_n(x)$

$n =$	0	1	2	3	4	5	6	7	8	9	10
$k = 0$	1	−1	−1	1	1	−1	−1	1	1	−1	−1
1		2	−2	−4	4	6	−6	−8	8	10	−10
2			4	−4	−12	12	24	−24	−40	40	60
3				8	−8	−32	32	80	−80	−160	160
4					16	−16	−80	80	240	−240	−560
5						32	−32	−192	192	672	−672
6							64	−64	−448	448	1792
7								128	−128	−1024	1024
8									256	−256	−2304
9										512	−512
10											1024

Table C.2a: Coefficients of x^{2k} in $T_{2n}(x)$ and of x^k in $T_n^*(x)$

$n =$	0	1	2	3	4	5	6	7	8	9	10
$k = 0$	1	−1	1	−1	1	−1	1	−1	1	−1	1
1		2	−8	18	−32	50	−72	98	−128	162	−200
2			8	−48	160	−400	840	−1568	2688	−4320	6600
3				32	−256	1120	−3584	9408	−21504	44352	−84480
4					128	−1280	6912	−26880	84480	−228096	549120
5						512	−6144	39424	−180224	658944	−2050048
6							2048	−28672	212992	−1118208	4659200
7								8192	−131072	1105920	−6553600
8									32768	−589824	5570560
9										131072	−2621440
10											524288

Table C.2b: Coefficients of x^{2k+1} in $T_{2n+1}(x)$ and of x^k in $V_n^*(x)$

$n =$	0	1	2	3	4	5	6	7	8	9	10
$k = 0$	1	−3	5	−7	9	−11	13	−15	17	−19	21
1		4	−20	56	−120	220	−364	560	−816	1140	−1540
2			16	−112	432	−1232	2912	−6048	11424	−20064	33264
3				64	−576	2816	−9984	28800	−71808	160512	−329472
4					256	−2816	16640	−70400	239360	−695552	1793792
5						1024	−13312	92160	−452608	1770496	−5870592
6							4096	−61440	487424	−2723840	12042240
7								16384	−278528	2490368	−15597568
8									65536	−1245184	12386304
9										262144	−5505024
10											1048576

Table C.3a: Coefficients of x^{2k} in $U_{2n}(x)$ and of x^k in $W_n^*(x)$

$n =$	0	1	2	3	4	5	6	7	8	9	10
$k = 0$	1	−1	1	−1	1	−1	1	−1	1	−1	1
1		4	−12	24	−40	60	−84	112	−144	180	−220
2			16	−80	240	−560	1120	−2016	3360	−5280	7920
3				64	−448	1792	−5376	13440	−29568	59136	−109824
4					256	−2304	11520	−42240	126720	−329472	768768
5						1024	−11264	67584	−292864	1025024	−3075072
6							4096	−53248	372736	−1863680	7454720
7								16384	−245760	1966080	−11141120
8									65536	−1114112	10027008
9										262144	−4980736
10											1048576

Table C.3b: Coefficients of x^{2k+1} in $U_{2n+1}(x)$ and of x^k in $2U_n^*(x)$

$n =$	0	1	2	3	4	5	6	7	8	9	10
$k = 0$	2	−4	6	−8	10	−12	14	−16	18	−20	22
1		8	−32	80	−160	280	−448	672	−960	1320	−1760
2			32	−192	672	−1792	4032	−8064	14784	−25344	41184
3				128	−1024	4608	−15360	42240	−101376	219648	−439296
4					512	−5120	28160	−112640	366080	−1025024	2562560
5						2048	−24576	159744	−745472	2795520	−8945664
6							8192	−114688	860160	−4587520	19496960
7								32768	−524288	4456448	−26738688
8									131072	−2359296	22413312
9										524288	−10485760
10											2097152