THE USE OF CHEBYSHEV POLYNOMIALS
IN NUMERICAL ANALYSIS

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Donnie R. Forisha, B. S.
Denton, Texas
December, 1975

The purpose of this paper is to investigate the nature and practical uses of Chebyshev polynomials. Chapter I gives recognition to mathematicians responsible for studies in this area. Chapter II enumerates several mathematical situations in which the polynomials naturally arise and suggests reasons for the pursuance of their study.

Chapter III includes: Chebyshev polynomials as related to "best" polynomial approximation, Chebyshev series, and methods of producing polynomial approximations to continuous functions.

Chapter IV discusses the use of Chebyshev polynomials to solve certain differential equations and Chebyshev-Gauss quadrature.
# TABLE OF CONTENTS

LIST OF GRAPHS ........................................ iv

Chapter

I. INTRODUCTION AND HISTORICAL NOTES .... 1

II. CHEBYSHEV POLYNOMIALS: THEIR DEFINITION, GENERATION, AND GENERAL PROPERTIES .... 6

III. APPROXIMATION AND REPRESENTATION OF FUNCTIONS BY CHEBYSHEV POLYNOMIALS ........ 34

IV. ORDINARY DIFFERENTIAL EQUATIONS AND A QUADRATURE FORMULA .... 68

BIBLIOGRAPHY ........................................... 89
LIST OF GRAPHS

Graph  Page

I. Graph I  .................  8
CHAPTER I

INTRODUCTION AND HISTORICAL NOTES

In numerical analysis, it is often important to find the polynomial \( p_n \), of specified maximum degree \( n \), such that, for a function \( f(x) \) defined on the interval \([a, b]\), \( |f(x) - p_n(x)| \) has as small a maximum value over \([a, b]\) as possible. Usually this polynomial \( p_n \) is rather difficult to produce, but a very good approximation to \( p_n \) can be found with relative ease. This approximation takes the form of a linear combination of a special group of polynomials, the Chebyshev* polynomials.

In Chapter II of this paper, the definition and general nature of these polynomials will be presented. The polynomials' properties which are developed in that chapter will then be used in Chapter III to describe the polynomials' role in making approximations to functions which are defined in some explicit form. Chapter IV outlines further the use of Chebyshev polynomials in numerical analysis by showing how they are used to estimate the solutions of certain types of differential equations and by employing them to estimate some definite integrals.

*Several different transliterations of the name, Chebyshev, into English are used in mathematical literature. Other common renderings are Tchebycheff and Tchebytcheff.
The functions which are treated in this paper as typical functions are assumed to be continuous and sufficiently integrable over the interval under consideration, unless otherwise noted.

Chebyshev polynomials are named as they are in honor of a prolific mathematician Pafnutii Lvovich Chebyshev, who was born on May 14, 1821, and died on November 26, 1894. He is one of the most prominent representatives of the Russian mathematical school. Even as a small boy he was greatly interested in mechanical inventions and it is said that in his first lesson in geometry he was able to see the subject's relevance to mechanics and therefore resolved to master it. He received his diploma from the University of Moscow when he was twenty years old, having already received a medal for a work on the numerical solution of algebraic equations of higher order.

Chebyshev's father was a Russian nobleman, but after the famine of 1840 the estate was so reduced that Chebyshev was forced to practice extreme economy for the rest of his life. He spent freely for nothing except the mechanical models of his inventions. He never married, but instead devoted himself solely to science.

In 1847, Chebyshev moved to the University of St. Petersburg where he served as professor of mathematics until 1882. During different periods of this time, he taught
analytic geometry, higher algebra, number theory, integral calculus, probability theory, the calculus of finite differences, the theory of elliptic functions, and the theory of definite integrals. Besides his great contribution to the new establishment of the Moscow school, he played a leading role in fortifying the tradition of the Petersburg school founded by Bunyakovskii and Ostrogradskii. His biographers are agreed that the quality of his teaching was no less remarkable than that of his research as he produced such notable pupils as Korkin, Lyapunov, and Markov.

Chebyshev collaborated with Bunyakovskii in 1847 to edit the mammoth manuscripts left by Euler, and, as a result, he grew many-sided and began to cultivate a certain hunch for great problems. He made important contributions to number theory, theory of least squares, interpolation theory, calculus of variations, infinite series, and probability theory, and published nearly a hundred memoirs on these and other mathematical topics, being best known for his work on primes. On the very day before his death, he received his friends as usual and discoursed upon the subject of a simple rule he had discovered for the rectification of a curve.

The polynomials whose properties and applications are discussed in this paper were discovered almost a century ago by Chebyshev. However, their importance for practical computation was rediscovered about thirty-five years ago by
C. Lanczos, who must be regarded as the modern father of this branch of numerical mathematics. Digital computers have given further emphasis to this development and in recent years the research literature of numerical mathematics has contained many papers on applications of Chebyshev polynomials and the theory and practice of Chebyshev approximation.
REFERENCES


CHAPTER II
CHEBYSHEV POLYNOMIALS: THEIR DEFINITION,
GENERATION, AND GENERAL PROPERTIES

In mathematical literature, usually the term "Chebyshev polynomials" actually refers to the Chebyshev polynomials of the first kind.

**Definition 1.1.** The Chebyshev polynomials of the first kind are given by
\[
T_n(x) = \cos n\theta, \quad x = \cos \theta, \quad -1 \leq x \leq 1,
\]
where \( n = 0, 1, 2, \ldots \) . The Chebyshev polynomials of the second kind are used less often.

**Definition 1.2.** The Chebyshev polynomials of the second kind are given by
\[
\varphi_n(x) = \frac{\sin[(n+1)\theta] / \sin \theta}{\sin [n\theta - \theta]}, \quad x = \cos \theta, \quad -1 \leq x \leq 1,
\]
where \( n = 0, 1, 2, \ldots \) .

It may be immediately obvious from the above definitions that Chebyshev polynomials are indeed polynomials. Therefore, note that the trigonometric identity,
\[
\cos[(n+1)\theta] + \cos[(n-1)\theta] = 2\cos \theta \cos n\theta,
\]
supplies a recurrence relation for the functions, \( T_n(x) = \cos n\theta \),
\[
T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)
\]
or
\[
T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \quad (1)
\]
Since \( T_0(x) = \cos 0 = 1 \), and \( T_1(x) = \cos \theta = x \), (1) can be used to generate any number of the \( T_n(x) \) in their specific polynomial forms.

\[
T_2(x) = 2x^2 - 1 \\
T_3(x) = 4x^3 - 3x \\
T_4(x) = 8x^4 - 8x^2 - 1 \\
T_5(x) = 16x^5 - 20x^3 + 5x \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\text{(2)}
\]

The graphs of these polynomials are shown on page 8.

Similarly, the trigonometric identity,

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

supplies a recurrence relation for Chebyshev polynomials of the second kind.

\[
\begin{align*}
\psi_n(x) &= \frac{\sin[(n+1)\theta]}{\sin \theta} \text{ by definition,} \\
&= \frac{(2\sin n\theta \cos \theta - \sin[(n-1)\theta])}{\sin \theta} \text{ by (3),} \\
&= (2\cos \theta)(\sin n\theta/\sin \theta) - \sin[(n-1)\theta]/\sin \theta.
\end{align*}
\]

Hence,

\[
\psi_n(x) = 2x\psi_{n-1}(x) - \psi_{n-2}(x).
\]

Since \( \psi_0(x) = \sin \theta/\sin \theta = 1 \), and \( \psi_1(x) = \sin 2\theta/\sin \theta = 2x \),

\[
\begin{align*}
\psi_2(x) &= 4x^2 - 1 \\
\psi_3(x) &= 8x^3 - 4x^2 \\
\psi_4(x) &= 16x^4 - 8x^3 - 8x^2 + 2x + 1 \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\text{(4)}
\]
GRAPH I
We now present another way of generating the Chebyshev polynomials which will provide more insight into their nature. We will see that if the Gram-Schmidt method is used to generate the unique set of orthonormal polynomials with respect to a particular inner product, then the Chebyshev polynomials result. The inner product will be given by
\[
\langle g(x), h(x) \rangle = \int_a^b g(x)h(x)w(x)dx,
\]
where \(g(x)\) and \(h(x)\) are any two functions, \(w(x)\), the weight function is defined by
\[
w(x) \equiv (1-x^2)^{-\frac{1}{2}},
\]
and the interval \([a, b] = [-1, 1]\) is used. Following the Gram-Schmidt procedure, first observe that the \(n+1\) polynomials, \(g_j(x) = x^j, j = 0, 1, \ldots, n\), are linearly independent over any interval. Let
\[
P_0(x) = d_0(1)
\]
\[
P_1(x) = d_1[x-c_{01}P_0(x)]
\]
\[
P_2(x) = d_2[x^2-c_{02}P_0(x)-c_{12}P_1(x)]
\]
\[\vdots\]
\[
P_n(x) = d_n[x^n-c_{0n}P_0(x)-c_{1n}P_1(x)-\ldots-c_{n-1,n}P_{n-1}(x)].
\]
The next step is to determine the constants \(d_i\) and \(c_{ij}\) so that \(P_0(x), P_1(x), \ldots, P_n(x)\) are mutually orthonormal. We require that
\[
\langle P_0(x), P_0(x) \rangle = \langle d_0, d_0 \rangle = \int_{-1}^1 d_0^2(1-x^2)^{-\frac{1}{2}}dx = 1.
\]
This implies that
\[ d_0^2 \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \, dx = d_0^2 \pi = 1, \]
from which
\[ d_0 = \sqrt{1/\pi}. \]
Hence,
\[ P_0(x) = \sqrt{1/\pi} = \sqrt{1/\pi} \cdot T_0(x). \]
According to the Gram-Schmidt method, \( P_1(x) = d_1 [x - c_{01} P_0(x)] \), where \( c_{01} = \langle P_0(x), x \rangle = 0 \). So \( P_1(x) = d_1 x \). We require that
\[ \langle P_1(x), P_1(x) \rangle = d_1^2 \int_{-1}^{1} x^2 (1-x^2)^{-\frac{1}{2}} \, dx = 1, \]
which implies \( d_1 = \sqrt{2/\pi} \).
Therefore,
\[ P_1(x) = x \sqrt{2/\pi} = \sqrt{2/\pi} \cdot T_1(x). \]
Using \( P_2(x) = d_2 [x^2 - c_{02} P_0(x) - c_{12} P_1(x)] \), where \( c_{02} = \langle P_0(x), x^2 \rangle \), \( c_{12} = \langle P_1(x), x^2 \rangle \), we find that \( c_{02} = \sqrt{\pi}/2; \ c_{12} = 0; \ d_2 = 2 \sqrt{2/\pi} \).
Hence,
\[ P_2(x) = 2 \sqrt{2/\pi} x^2 - \sqrt{2/\pi} = \sqrt{2/\pi} (2x^2 - 1). \]
Recalling (2), we see that \( P_2(x) = \sqrt{2/\pi} T_2(x) \). In order to show that the general relation \( P_j(x) = \sqrt{2/\pi} T_j(x) \), \( j = 1, 2, 3, \ldots \), holds true, first we remind ourselves that we are producing a set of polynomials \( P_0(x), P_1(x), P_2(x), \ldots \), which are mutually orthonormal with respect to the inner product,
\[ \langle g(x), h(x) \rangle = \int_{-1}^{1} g(x) h(x) (1-x^2)^{-\frac{1}{2}} \, dx. \]
We will then show that \( Q_0(x) = \sqrt{1/\pi} T_0(x) \), \( Q_j(x) = \sqrt{2/\pi} T_j(x) \), \( j = 1, 2, \ldots \), define another orthonormal set, and that \( P_j(x) = Q_j(x) \), \( j = 0, 1, 2, \ldots \), because such an orthonormal set
is unique. Start by noting that

\[ \langle Q_i(x), Q_j(x) \rangle = \int_{-1}^{1} Q_i(x)Q_j(x)(1-x^2)^{-\frac{3}{2}}dx. \]

So,

\[ \langle Q_i(x), Q_j(x) \rangle = \left(\frac{2}{\pi}\right) \int_{0}^{\pi} \cos(i\theta) \cdot \cos(j\theta) d\theta. \] (5)

If \( i = j \), then by (5),

\[ \langle Q_i(x), Q_j(x) \rangle = \left(\frac{2}{\pi}\right) \int_{0}^{\pi} \cos^2(i\theta) d\theta = 1. \] (6)

If \( i \neq j \), then by (5),

\[ \langle Q_i(x), Q_j(x) \rangle = \left(\frac{2}{\pi}\right) \int_{0}^{\pi} \cos(i\theta) \cdot \cos(j\theta) d\theta = 0. \] (7)

From (6) and (7), it is known that the polynomials \( \{Q_i(x)\} \) are orthonormal.

One well-known property of orthonormal polynomials is that, given any polynomial \( R_n(x) \) of exact degree \( n \), there exist constants \( \alpha_{0n}, \alpha_{1n}, \ldots, \alpha_{nn} \) such that

\[ R_n(x) = \alpha_{0n} Q_0(x) + \alpha_{1n} Q_1(x) + \ldots + \alpha_{nn} Q_n(x), \]

where \( Q_0(x), Q_1(x), Q_2(x), \ldots \) are orthonormal. Using this property we can write \( P_0(x) = \alpha_{00} Q_0(x) \), for some constant \( \alpha_{00} \).

Since \( \langle P_0(x), P_0(x) \rangle = 1 \), \( \langle \alpha_{00} Q_0(x), \alpha_{00} Q_0(x) \rangle = \alpha_{00}^2 \langle Q_0(x), Q_0(x) \rangle = 1 \)

which means \( \alpha_{00} = 1 \), and \( P_0(x) = Q_0(x) \). Also,

\[ P_1(x) = \alpha_{01} Q_0(x) + \alpha_{11} Q_1(x). \]

Since \( \langle P_1(x), P_0(x) \rangle = 0 \),

\[ \langle \alpha_{01} Q_0(x) + \alpha_{11} Q_1(x), P_0(x) \rangle = \alpha_{01} \langle Q_0(x), P_0(x) \rangle + \alpha_{11} \langle Q_1(x), P_0(x) \rangle \]

\[ = \alpha_{01} \langle Q_0(x), P_0(x) \rangle = \alpha_{01} = 0. \]

Hence,

\[ P_1(x) = \alpha_{11} Q_1(x). \]
Since $\langle P_1(x), P_1(x) \rangle = 1$,
\[
\langle \alpha_{11} Q_1(x), \alpha_{11} Q_1(x) \rangle = \alpha_{11}^2 \langle Q_1(x), Q_1(x) \rangle = \alpha_{11}^2 = 1.
\]
This means that $\alpha_{11} = 1$ and $P_1(x) = Q_1(x)$. Suppose that $P_j(x) = Q_j(x)$, for $j = 0, 1, \ldots, k-1$, where $k \geq 2$. We know that $P_k(x) = a_0 Q_0(x) + a_1 Q_1(x) + \ldots + a_k Q_k(x)$. Since for $i = 0, 1, \ldots, k-1$, $\langle P_k(x), P_i(x) \rangle = 0$,
\[
\langle a_{0k} Q_0(x) + \ldots + a_{kk} Q_k(x), P_i(x) \rangle = \alpha_{ik} \langle Q_i(x), P_i(x) \rangle = \alpha_{ik} = 0.
\]
So, $P_k(x) = a_{kk} Q_k(x)$. Since $\langle P_k(x), P_k(x) \rangle = 1$,
\[
\langle a_{kk} Q_k(x), a_{kk} Q_k(x) \rangle = \alpha_{kk}^2 \langle Q_k(x), Q_k(x) \rangle = \alpha_{kk}^2 = 1,
\]
so that $P_k(x) = Q_k(x)$. By induction, $P_j(x) = Q_j(x)$ for $j = 0, 1, \ldots$. We have thus proved that the polynomials $P_0(x), P_1(x), \ldots$, produced by the Gram-Schmidt method are actually multiples of the Chebyshev polynomials.

It is sometimes handy to have available the expressions for powers of $x$ in terms of the $T_r$. Reversing (2), we get
\[
\begin{align*}
1 & = T_0(x) \\
x & = T_1(x) \\
x^2 & = (T_0(x) + T_2(x))/2 \\
x^3 & = (3T_1(x) + T_3(x))/4 \\
x^4 & = (3T_0(x) + 4T_2(x) + T_4(x))/8 \\
x^5 & = (10T_1(x) + 5T_3(x) + T_5(x))/16 \\
& \vdots \\
& \vdots \\
& \vdots
\end{align*}
\]
The range, -1 ≤ x ≤ 1, may be changed to the range, 0 ≤ y ≤ 1, by the linear transformation, y = (x+1)/2, x = 2y - 1.

We introduce some "new" Chebyshev polynomials $T_r^*(x)$,

$$T_r^*(x) = T_r(2x-1), \ 0 ≤ x ≤ 1.$$  

All of the properties of the $T_r^*(x)$ can be deduced from those of $T_r(2x-1)$. For example, if in the recurrence relation (1), x is replaced by 2x-1, then

$$T_{r+1}^*(x) = 2(2x-1)T_r^*(x) - T_{r-1}^*(x), \quad (9)$$

a recurrence relation for the $T_r^*(x)$. Noting that

$$T_0^*(x) = T_0(2x-1) = 1,$$

and

$$T_1^*(x) = T_1(2x-1) = 2x - 1,$$

and employing (9), yields

$$T_2^*(x) = 8x^2 - 8x + 1$$
$$T_3^*(x) = 32x^3 - 48x^2 + 18x - 1$$

... 

Reversing (10) produced

$$1 = T_0^*(x)$$
$$x = (T_0^*(x)+T_1^*(x))/2$$
$$x^2 = (3T_0^*(x)+4T_1^*(x)+T_2^*(x))/8$$
$$x^3 = (10T_0^*(x)+15T_1^*(x)+6T_2^*(x)+T_3^*(x))/32$$
$$x^4 = (34T_0^*(x)+56T_1^*(x)+28T_2^*(x)+8T_3^*(x)+T_4^*(x))/64$$
$$x^5 = (118T_0^*(x)+210T_1^*(x)+120T_2^*(x)+45T_3^*(x)+10T_4^*(x)+T_5^*(x))/512$$

... 

(11).
A general expression for $T_r(x)$ can be obtained through establishing a generating function. The generating function will be some function, $g(x)$, expressed in a closed form and involving $p$ and $x$. If $g(x)$ can be written as a series of the form $\sum_{r=0}^{\infty} p^r G_r(x)$, then each $T_r(x)$ is given by $G_r(x)$. Since

$$\sum_{r=0}^{\infty} p^r e^{i r \theta} = \sum_{r=0}^{\infty} p^r (\cos r \theta + i \sin r \theta)$$

$$= \sum_{r=0}^{\infty} p^r \cos r \theta + i \sum_{r=0}^{\infty} p^r \sin r \theta, \quad (12)$$

it is seen that the real part of (12) is exactly the series, $\sum_{r=0}^{\infty} p^r T_r(x)$.

The real part of (12) can also be written in an alternate form. This is because

$$\sum_{r=0}^{\infty} p^r e^{i r \theta} = [1-p(\cos \theta+i \sin \theta)]^{-1}$$

$$= [1-px-ip/1-x^2]^{-1}$$

$$= (1-px+ip/1-x^2)/((1-px)^2+p^2(1-x^2)).$$

Hence the real part of the expression yields the sought after generating function

$$\sum_{r=0}^{\infty} p^r T_r(x) = (1-px)/(1-2px+p^2)$$

$$= (1-px)/(1-p(2x-p))$$

$$= [1-px][1+p(2x-p)+p^2(2x-p)^2+\ldots].$$

Equating the coefficients of $p^r$ gives for $r\geq 1$,

$$T_r(x) = \{(2x)^r-\left[2B(r-1,1)-B(r-2,1)\right](2x)^{r-2}$$

$$+ \left[2B(r-2,2)-B(r-3,2)\right](2x)^{r-4}-\ldots\}/2, \quad (13)$$

where $B(i,j) = \binom{i}{j}$. Once again we have a device by which $T_r(x)$ can be written in its explicit polynomial form.
In order to obtain the general expression for \( x^r \) in terms of \( T_r(x) \), notice that

\[
x^r = (\cos \theta)^r = ((e^{i\theta} + e^{-i\theta})/2)^r/2^{r-1}
\]

\[
= \frac{1}{2^{n-1}} \{T_r(x)+B(r,1)T_{r-2}(x)+B(r,2)T_{r-4}(x)+\ldots\}, \quad (14)
\]

in which, for even \( r \), we take half the coefficient of \( T_0(x) \).

In order to obtain the corresponding results for the \( T_r^*(x) \), we note that

\[
T_s(T_r(x)) = T_r(T_s(x)) = T_r(\cos s\theta)
\]

\[
= \cos(r[\cos^{-1}(\cos s\theta)])
\]

\[
= \cos(rs\theta).
\]

So,

\[
T_s(T_r(x)) = T_{rs}(\theta). \quad (15)
\]

In particular, for \( s = 2 \), (15) yields

\[
T_r(T_2(x)) = T_{2r}(x). \quad (16)
\]

Using (16), we can write

\[
T_r(T_2(x)) = T_r(2x^2-1) = T_2(T_r(x))
\]

\[
= 2T_r^2(x)-1 = T_{2r}(x). \quad (17)
\]

If in (17), \( x^2 \) is replaced by \( x \), then

\[
T_r(2x-1) = T_r^*(x) = 2T_r^2(x^{1/2}) - 1
\]

\[
= T_{2r}(x^{1/2}). \quad (18)
\]

So, from (18) and (13),

\[
T_r^*(x) = T_{2r}(x^{1/2}) = \{2^{2r}x^r-[2B(2r-1,1)-B(2r-2,1)]2^{2r-2}x^{r-1}
\]

\[
+ [2B(2r-2,2)-B(2r-3,2)]2^{2r-4}x^{r-2} - \ldots\}/2.
\]
Similarly, from (18) and (14), an expression for $x^r$ in terms of the $T_r(x)$ can be produced,

$$x^r = \left(\frac{k^2}{2r}\right)^{2r} = \left\{ T_r(x) + \binom{2r}{1} T_{r-1}(x) + \binom{2r}{2} T_{r-2}(x) + \ldots \right\}/2^{r-1}.$$

The employment of certain properties of the Chebyshev polynomials leads to the development of some useful methods for solving ordinary differential equations and other practical problems. It is these properties which make the study of Chebyshev polynomials a profitable one.

An inspection of the recurrence relation (1) and of the first few $T_r(x)$ reveals that if $r$ is even, then $T_r(x)$ is an even function. Likewise, if $r$ is odd, then $T_r(x)$ is an odd function.

All of the $T_r(x)$ have the value of unity at $x = 1$, and at $x = -1$, the value is $+1$ for even $r$ and $-1$ for odd $r$. This is because $T_r(1) = \cos r(\cos^{-1}1) = 1$, and

$$T_r(-1) = \cos r[\cos^{-1}(-1)] = \cos r\pi = \{-1\text{ for even } r\}.$$  

The turning points of $T_r(x)$ occur at the zeros of $U_{r-1}(x)$. This is because

$$T_r'(x) = \frac{d}{dx} \cos r\theta = r(\sin r\theta/\sin \theta).$$

So the turning points of $T_r(x)$ occur at the zeros of $\sin r\theta/\sin \theta$ which is $U_{r-1}(x)$. These zeros are

$$x_k = \cos(k\pi/r), \quad k = 1, 2, \ldots, r-1.$$  

The turning points separate the $r$ zeros of $T_r(x)$,

$$y_k = \cos(k+1/2)(\pi/r), \quad k = 0, 1, \ldots, r-1.$$
Both the turning points and the zeros are symmetrically placed about the origin. Each $T_r(x)$ has $r+1$ equal maximum and minimum values in $[-1,1]$ which occur at the $r-1$ turning points and at the two endpoints $x = -1$ and $x = 1$.

We now establish various functional equations which involve the $T_r(x)$ and the $T^*_r(x)$ and which will prove useful later.

$$T_s(x) \cdot T_r(x) = (\cos s \theta)(\cos r \theta)$$
$$= \frac{1}{2} \cos(s-r)\theta + \frac{1}{2} \cos(s+r)\theta. $$

Hence,

$$T_s(x) \cdot T_r(x) = \frac{1}{2} [T_{s-r}(x) + T_{s+r}(x)].$$

$$T_s^*(x)T_r^*(x) = T_{2s}(x^{\frac{1}{2}})T_{2r}(x^{\frac{1}{2}}) = \frac{1}{2} [T_{2s-2r}(x^{\frac{1}{2}}) + T_{2s+2r}(x^{\frac{1}{2}})].$$

Thus, $T_s^*(x)T_r^*(x) = \frac{1}{2} [T_{s-r}^*(x) + T_{s+r}^*(x)].$

$$x^r T_s(x) = \frac{1}{2^{r-1}} [T_{r-2}(x)T_s(x) + B(r,1)T_{r-4}(x)T_s(x) + B(r,2)T_{r-6}(x)T_s(x) + \ldots]$$
$$= \frac{1}{2^r} [T_{r-s}(x) + T_{r+s}(x) + B(r,1)(T_{r-s-2}(x) + T_{r+s-2}(x)) + B(r,2)T_{r-s-4}(x)$$
$$+ T_{r+s-4}(x)] + \ldots.$$ 

Since $B(r,i) = B(r,r-i),$

$$x^r T_s(x) = \frac{1}{2^{r-1}} \sum_{i=0}^{\infty} B(r,i) T_{s-r+2i}(x). \quad (19)$$

Notice $x^r T^*_s(x) = (x^*)^{2r} T_{2s}(x^{\frac{1}{2}}).$ Hence, from (19),

$$x^r T^*_s(x) = \frac{1}{2^{2r-1}} \sum_{i=0}^{2r} B(2r,i) T_{2s-2r+2i}(x^{\frac{1}{2}})$$
$$= \frac{1}{2^{2r-1}} \sum_{i=0}^{2r} B(2r,i) T^*_{s-r+i}(x). \quad (20)$$
As shown earlier, the Chebyshev polynomials are orthogonal with respect to the weight function, \((1-x^2)^{-\frac{1}{2}}\). Hence, \(T_r(x)\) is orthogonal to every polynomial \(q_{r-1}(x)\) of degree \(r-1\) or less; that is,

\[
\int_{-1}^{1} \left( T_r(x)q_r(x) / \sqrt{1-x^2} \right) dx.
\]

Define

\[
V_r^{(r)}(x) = T_r(x) / \sqrt{1-x^2},
\]
so that

\[
V_r^{(r)}(x)q_{r-1}(x) = T_r(x)q_{r-1}(x) / \sqrt{1-x^2}.
\]

Now, if we perform definite integration \(r\) times over the interval \([-1,1]\), on both sides of (22), then

\[
[V_r^{(r-1)}(x)q_{r-1}(x) - V_r^{(r-2)}(x)q_{r-1}(x) + \ldots + (-1)^{r-1} V_r(x)q_{r-1}(x)]_{-1}^{1} = 0.
\]

This result, coupled with the fact that \(T_{r+1}(x) \equiv 0\), yields the differential system,

\[
\frac{d^{r+1}}{dx^{r+1}} \left( \sqrt{1-x^2} \right) V_r(x) = 0,
\]

\[
V_r(\#1) = V_r'(\#1) = \ldots = V_r^{(r-1)}(\#1) = 0.
\]

The solution of (23) is

\[
V_r(x) = C_r (1-x^2)^{r-\frac{1}{2}},
\]

where \(C_r\) is a constant. So, from (21),

\[
T_r(x) = C_r \sqrt{1-x^2} \frac{d^r}{dx^r} \left( (1-x^2)^{r-\frac{1}{2}} \right),
\]

where the constant \(C_r\) is to be determined by some normalization requirement.
The polynomial $T_r(x)$ also satisfies the differential equation,

$$(1-x^2)T''_r(x) - xT'_r(x) + r^2 T_r(x) = 0. \quad (24)$$

We establish this property by first considering $T_r(\theta)$.

Recall that $T_r(\theta) = \cos r\theta$. So, $T'_r(\theta) = -r \sin r\theta$, and $T''_r(\theta) = -r^2 \cos r\theta$, so that

$$T''_r(\theta) + r^2 T_r(\theta) = -r^2 \cos r\theta + r^2 \cos r\theta = 0. \quad (25)$$

Now we restore $x$ as the independent variable of (25) by first transforming $T''_r(\theta)$.

$$T''_r(\theta) = \frac{d}{d\theta} \left( \frac{dx}{d\theta} T'_r(x) \right)$$

$$= \frac{dx}{d\theta} \left( -\sqrt{1-x^2} T'_r(x) \right)$$

$$= -\sqrt{1-x^2} \left[ -\sqrt{1-x^2} T''_r(x) - \frac{1}{2}(-2x)(1-x^2)^{-1/2}T'_r(x) \right].$$

Hence,

$$T''_r(\theta) = (1-x^2)T''_r(x) - xT'_r(x). \quad (26)$$

Putting (26) into (25) verifies (24).

Just as the Chebyshev polynomials of the first kind are orthogonal with respect to the weight function, $(1-x^2)^{-1/2}$, the Chebyshev polynomials of the second kind are orthogonal with respect to the weight function, $(1-x^2)^{1/2}$. Also, similarly, the $U_r(x)$ satisfy

$$(1-x^2)U''_r(x) - 3xU'_r(x) + (r^2+2r)U_r(x) = 0. \quad (27)$$

The proof of this follows as before. $U_r(\theta) = \sin[(r+1)\theta]/\sin \theta$, $U'_r(\theta) = (r+1)\cos[(r+1)\theta]/\sin \theta - (\cos \theta \sin[(r+1)\theta])/\sin^2 \theta$. 
\[ U''(\theta) = \left\{ -(r+1)^2 \sin \theta \sin[(r+1)\theta] + \sin \theta \sin[(r+1)\theta] - 2(r+1) \cos \theta \cos[(r+1)\theta]/\sin^2 \theta + (2 \cos^2 \theta \sin[(r+1)\theta]/\sin^3 \theta. \right\} \] 

So,

\[ (r^2+2r)U_r(\theta) = (r^2+2r)\sin[(r+1)\theta]/\sin \theta. \] \tag{29}

\[ \frac{2 \cos \theta}{\sin \theta} U_r'(\theta) = \frac{2(r+1) \cos \theta \cos[(r+1)\theta]}{\sin^2 \theta} - \frac{2 \cos^2 \theta \sin[(r+1)\theta]}{\sin^3 \theta}. \tag{30} \]

Using (28), (29) and (30), we can write

\[ U''(\theta) + \frac{2 \cos \theta}{\sin \theta} U'_r(\theta) + (r^2+2r)U_r(\theta) = \left\{ \frac{[(r+1)^2-1] \sin[(r+1)\theta]}{\sin \theta} - \frac{2(r+1) \cos \theta \cos[(r+1)\theta]}{\sin^2 \theta} + \frac{2 \cos^2 \theta \sin[(r+1)\theta]}{\sin^3 \theta} + \frac{(r^2+2r)\sin[(r+1)\theta]}{\sin \theta} - \frac{2 \cos^2 \theta \sin[(r+1)\theta]}{\sin^3 \theta} \right\} \]

\[ = 0. \tag{31} \]

To revert back to \( x \), we write

\[ U''_r(\theta) = \frac{d}{d\theta} \left( -\sqrt{1-x^2} U'_r(x) \right) \]

\[ = -\sqrt{1-x^2} \left[ -\sqrt{1-x^2} U''_r(x) + U'_r(x) x(1-x^2)^{-1/2} \right] \]

\[ = (1-x^2) U''_r(x) - xU'_r(x). \tag{32} \]

\[ \frac{2 \cos \theta}{\sin \theta} U_r'(\theta) = -2xU'_r(x). \tag{33} \]

\[ (r^2+2r)U_r(\theta) = (r^2+2r)U_r(x). \tag{34} \]
Substituting (32), (33), and (34) back into (31) verifies the original differential equation, (27).

The hypergeometric function, \( F(\alpha, \beta, \gamma; g) \), defined by the series

\[
F = 1 + \frac{\alpha \beta}{\gamma \cdot 1} y + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1) \cdot 1 \cdot 2} y^2 \\
+ \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{\gamma(\gamma+1)(\gamma+2) \cdot 1 \cdot 2 \cdot 3} y^3 + \ldots,
\]

is a standard function in mathematical physics. It satisfies the differential equation,

\[
y(1-y)\frac{d^2F}{dy^2} + \{\gamma-(\alpha+\beta+1)y\}\frac{dF}{dy} - \alpha\beta F = 0. \tag{35}
\]

If the substitution, \( y = (x+1)/2 \), is made into (35), then

\[
(1-x^2)F'' + \{\gamma-(\alpha+\beta+1)(x+1)\}F' - \alpha\beta F = 0.
\]

Observe that this means \( F(\alpha+1, -\alpha, 1; (1+x)/2) \) satisfies

\[
(1-x^2)F''(x) - xF'(x) + r^2 F(x) = 0, \quad \text{the same as (24), the differential equation which } T_r(x) \text{ satisfies. So,}
\]

\[
T_r(x) = k_1 F(\alpha+1, -\alpha, 1; (1+x)/2), \quad \text{for some constant } k_1.
\]

Similarly, observe that \( F(\alpha, 3, -\alpha; (1+x)/2) \) satisfies

\[
(1-x^2)F''(x) + \{3-(\alpha+2-r+1)(1+x)\}F'(x) + r(r+2)F(x) = 0,
\]
or

\[
(1-x^2)F''(x) - 3xF'(x) + (r^2+2r)F(x) = 0,
\]

the same as (27), the differential equation which \( U_r(x) \) satisfies. Hence,

\[
U_r(x) = k_2 \cdot F(\alpha, 3, -\alpha; (1+x)/2),
\]

for some constant \( k_2 \).
Later on, the need to integrate $T_r(x)$ arises. So, we investigate:

$$\int T_r(x) \, dx = -\int \cos \theta \sin \theta \, d\theta + C,$$

$$= -\frac{1}{2} \int \{\sin[(r+1)\theta] - \sin[(r-1)\theta]\} \, d\theta + C,$$

$$= \frac{1}{2} \left[ \frac{1}{r+1} \cos[(r+1)\theta] - \frac{1}{r-1} \cos[(r-1)\theta] \right] + C,$$

$$= \frac{1}{2} \left[ \frac{1}{r+1} T_{r+1}(x) - \frac{1}{r-1} T_{r-1}(x) \right] + C. \quad (36)$$

This formula makes no sense for $r = 0$ or $1$. We treat these cases individually.

$$\int T_0(x) \, dx = x = T_1(x) + C,$$

$$\int T_1(x) \, dx = \int x \, dx = \frac{x^2}{4} = \frac{1}{4} T_2(x) + C. \quad (37)$$

Below, we list other properties which are of interest.

$$T'_r(x) = \frac{d^2}{dx} \frac{d}{d\theta} (\cos r\theta) = \frac{r \sin r\theta}{\sin \theta} = rU_{r-1}(x).$$

So, $\int U_{r-1}(x) = \frac{1}{r} T_{r}(x) + C.$

$$U'_r(x) = \frac{d^2}{dx} \frac{d}{d\theta} \frac{\sin[(r+1)\theta]}{\sin \theta} = \frac{1}{1-x^2} [xU_r(x) - (r+1)T_{r+1}(x)].$$

A truncated Chebyshev series is a particular linear combination of Chebyshev polynomials which we will completely define in Chapter II, but, for now, let it be represented by

$$p(x) = \sum_{r=0}^{n'} a_r x^r, \quad (38)$$

where the prime indicates that the first term is taken with a factor of $1/2$. In the techniques studied in Chapter III
for the treatment of an ordinary differential equation, it is essential to be able to evaluate \( p(x) \) at some given point, to integrate \( p(x) \), and to differentiate \( p(x) \). These are easier said than done, but, nevertheless, carrying them out is straightforward.

To evaluate \( p(x) = \sum_{r=0}^{n} a_r T_r(x) \) at a given point, \( x = x_0 \), we can convert to powers of \( x \),
\[
p(x) = \sum_{r=0}^{n} c_r x^r
\]
and then perform nested multiplication,
\[
q_r(x_0) = x_0 q_{r+1}(x_0) + c_r; \quad q_{n+1}(x_0) = 0
\]
to yield \( p(x_0) = q_0(x_0) \). Since \( |x_0| \leq 1 \), this process is relatively stable. However, it has two disadvantages. First, the coefficients \( c_r \) may be very much larger than the coefficients \( a_r \), even if \( p(x_0) \) is small, and this sometimes makes it hard to store the coefficients to the desired precision.

For example, conversion of the truncated Chebyshev series,
\[
p(x) = \left(\frac{1}{2}\right) T_4(x) + \left(\frac{1}{4}\right) T_6(x) + \left(\frac{1}{8}\right) T_8(x),
\]
to powers of \( x \) results in
\[
p(x) = -\frac{1}{4} - 5x^2 + 58x^4 - 211x^6 + 64x^8,
\]
where \( p(.1) = -0.294411 \).

Second, if (38) is truncated at an earlier term, to give a poorer but still satisfactory precision for \( p(x) \), then most of the coefficients \( c_r \) in (39) are changed, although the \( a_r \) retained in (38) have their original values. The Chebyshev series
\[
p(x) = \left(\frac{15}{8}\right) T_0(x) - \left(\frac{7}{4}\right) T_1(x) + T_2(x) - \left(\frac{1}{4}\right) T_3(x) + \left(\frac{1}{8}\right) T_4(x),
\]
is an exact representation of the polynomial,
\[
q(x) = 1 - x + x^2 - x^3 + x^4.
\]
If we now leave off the final term of $p(x)$ so as to give a lower degree approximation $p^*(x)$ to $q(x)$, we get

$$p^*(x) = \left( \frac{15}{8} \right) T_0(x) - \left( \frac{7}{4} \right) T_1(x) + T_2(x) - \left( \frac{1}{4} \right) T_3(x)$$

$$= \left( \frac{7}{8} \right) - x + 2x^2 - x^3.$$

We therefore seek a method which does not involve the polynomial form (39). Let $\phi_0(x), \phi_1(x), \ldots, \phi_r(x)$ be a series of functions which satisfy a linear recurrence relation

$$\phi_r + a_{r-1}(x)\phi_{r-1}(x) + \beta_{r-1}(x)\phi_{r-2}(x) = 0.$$ 

The sum of this series of functions is given by

$$\sum_{r=0}^{n} a_r \phi_r(x) = a_0 \phi_0(x) + a_1 \phi_1(x) + \ldots + a_n \phi_n(x)$$

$$= a_0 \phi_0(x) + \ldots + a_n (-\alpha_{n-1}(x) \phi_{n-1}(x))$$

$$- \beta_{n-1}(x) \phi_{n-2}(x))$$

$$= a_0 \phi_0(x) + \ldots + (a_{n-2} - a_n \beta_{n-1}(x)) \phi_{n-2}(x)$$

$$+ (a_{n-1} - a_n \alpha_{n-1}(x)) \phi_{n-1}(x)$$

(40)

Let $b_n(x) = a_n$ so that (40) is written

$$= a_0 \phi_0(x) + \ldots + (a_{n-2} - b_n(x) \beta_{n-1}(x)) \phi_{n-2}(x)$$

$$+ (a_{n-1} - b_n(x) \alpha_{n-1}(x)) \phi_{n-1}(x).$$

Now let $b_{n-1}(x) = -\alpha_{n-1}(x) b_n(x) + a_{n-1}$ so that the last equation becomes

$$= a_0 \phi_0(x) + \ldots + (a_{n-2} - b_n(x) \beta_{n-1}(x)) \phi_{n-2}(x)$$

$$+ b_{n-1}(x) \phi_{n-1}(x),$$

(41)
\[ a_0 \phi_0(x) + \ldots + (a_n - 2 \cdot b_n(x) \beta_{n-2}(x)) \phi_{n-2}(x) + b_{n-1}(x) (-\alpha_{n-2}(x) \phi_{n-2}(x) - \beta_{n-2}(x) \phi_{n-3}(x)) = a_0 \phi_0(x) + \ldots + (a_n - 2 \cdot \beta_{n-2}(x)b_{n-1}(x)) \phi_{n-3}(x) + (a_{n-2} - \beta_{n-1}(x)b_n(x) - \alpha_{n-2}(x)b_{n-1}(x)) \phi_{n-2}(x). \]  

Let 
\[ b_{n-2}(x) = -\alpha_{n-2}(x)b_{n-1}(x) - \beta_{n-1}(x)b_n(x) + a_{n-2} \]
and (42) becomes 
\[ = a_0 \phi_0(x) + \ldots + (a_n - 3 \cdot \beta_{n-2}(x)b_{n-1}(x)) \phi_{n-3}(x) + b_{n-2}(x) \phi_{n-2}(x). \]  

Note that (43) is of the same form as (41). Hence, additional steps may be performed, each time setting 
\[ b_r(x) = -\alpha_r(x)b_{r+1}(x) - \beta_{r+1}(x)b_{r+2}(x) + a_r. \]  

\[ \sum_{r=0}^{n} a_r \phi_r(x) = (a_0 - b_2(x) \beta_1(x)) \phi_0(x) + b_1(x) \phi_1(x), \]
\[ = (a_0 - b_2(x) \beta_1(x) - \alpha_0(x)b_1(x)) \phi_0(x) + \alpha_0(x)b_1(x) \phi_0(x) + b_1(x) \phi_1(x), \]
\[ = b_0(x) \phi_0(x) + b_1(x)(\alpha_0(x) \phi_0(x) + \phi_1(x)). \]

Hence, all of the above \( b_r(x) \) result from the recurrence relation 
\[ b_r(x) = -\alpha_r(x)b_{r+1}(x) - \beta_{r+1}(x)b_{r+2}(x) + a_r, \]
r = n, n-1, \ldots, 0, with the starting conditions 
\[ b_{n+1}(x) = b_{n+2}(x) = 0. \]
For the Chebyshev case, \( \sum_{r=0}^{n'} \alpha_r T_r(x) \), we have

\[
\alpha_r(x) = -2x, \quad \beta_r(x) = 1
\]

so that

\[
\sum_{r=0}^{n'} \alpha_r T_r(x) = (b_0(x) - a_0/2)T_0(x) + b_1(x)(T_1(x) - 2xT_0(x)). \tag{44}
\]

Since \( T_0(x) = 1, T_1(x) = xT_0(x), b_0(x) = a_0 + 2xb_1(x) - b_2(x) \), then from (44)

\[
\sum_{r=0}^{n'} \alpha_r T_r(x) = (b_0(x) - b_2(x))/2, \tag{45}
\]

where \( b_{n+1}(x) = b_{n+2}(x) = 0 \),

\[
b_r(x) = 2xb_{r+1}(x) - b_{r+2}(x) + a_r, \quad r = n, n-1, \ldots, 0. \tag{46}
\]

Three special point evaluations of \( p(x) \) are

\[
p(1) = \frac{a_0}{2} + a_1 + a_2 + \ldots + a_n,
\]

\[
p(-1) = \frac{a_0}{2} - a_1 + a_2 - a_3 + \ldots + (-1)^n a_n,
\]

\[
p(0) = \frac{a_0}{2} - a_2 + a_4 - a_6 + \ldots + (-1)^n a_{2n}. \tag{47}
\]

The corresponding results for the \( T^*(x) \) follow easily.

From (45) and (46),

\[
\sum_{r=0}^{n} a_r T^*_r(x_0) = \sum_{r=0}^{n} a_r T_r(2x_0 - 1)
\]

\[
= \frac{1}{2} [b_0(x_0) - b_2(x_0)],
\]

where

\[
b_r(x_0) = a_r + 2(2x_0 - 1)b_{r+1}(x_0) - b_{r+2}(x_0),
\]

\[
b_{r+1}(x_0) = b_{r+2}(x_0) = 0.
\]
If \( q(x) = \sum_{r=0}^{n} a_r T_r(x) \), then
\[
q(x) = \sum_{r=0}^{n} a_r T_r(2x-1).
\]

So, from (47),
\[
\begin{align*}
q(1) &= \sum_{r=0}^{n} a_r T_r(1) = \frac{a_0}{2} + a_1 + a_2 + \ldots + a_n, \\
q(0) &= \sum_{r=0}^{n} a_r T_r(-1) = \frac{a_0}{2} - a_1 + a_2 - \ldots + (-1)^n a_n, \\
q(\frac{1}{2}) &= \sum_{r=0}^{n} a_r T_r(0) = \frac{a_0}{2} - a_2 + a_4 - \ldots + (-1)^n a_{2n}.
\end{align*}
\]

Now, we apply ourselves to performing indefinite integration on \( p(x) \).
\[
\begin{align*}
\int_{-1}^{x} p(x) \, dx &= \int p(x) \, dx - \left[ \int p(x) \, dx \right]_{x=-1} \\
&= \sum_{r=0}^{n} a_r \int T_r(x) \, dx - \left[ \sum_{r=0}^{n} a_r \int T_r(x) \, dx \right]_{x=-1}.
\end{align*}
\]
Let \( P(x) \equiv \int p(x) \, dx \). Continuing from (48), we have
\[
\begin{align*}
\int_{-1}^{x} p(x) \, dx &= \frac{a_0}{2} T_0(x) + \frac{a_1}{4} (T_2(x)+1) \\
&\quad + \sum_{r=2}^{n} a_r \left[ \frac{1}{r+1} T_{r+1}(x) - \frac{1}{r-1} T_{r-1}(x) \right] \\
&\quad - P(-1) \\
&= \frac{a_1}{4} T_0(x) + \left( \frac{a_0}{2} - \frac{a_2}{2} \right) T_1(x) \\
&\quad + \left( \frac{a_1}{4} - \frac{a_3}{4} \right) T_2(x) \\
&\quad + \sum_{r=3}^{n-1} \left( \frac{r-1}{2} - \frac{r+1}{2} \right) T_r(x) \\
&\quad + \left( \frac{a_{n-1}}{2n} \right) T_n(x) + \frac{a_n}{2(n+1)} - P(-1) T_0(x).
\end{align*}
\]
Writing (49) in a neater form, we get

\[ \int_{-1}^{x} p(x)dx = \sum_{r=0}^{n+1} b_r T_r(x), \quad (50) \]

where

\[
\begin{align*}
    b_{n+1} & = \frac{a_n}{2(n+1)}, \\
    b_n & = \frac{a_{n-1}}{2n}, \\
    b_r & = \frac{a_{r-1}}{2r} - \frac{a_{r+1}}{2r} \quad \text{for } r = 1, 2, \ldots, n-1,
\end{align*}
\]

and

\[ b_0 = 2 \left[ \frac{a_1}{2} - P(-1) \right]. \]

Therefore, what we essentially have in (50) is an expression for \( P(x) \) which involves \( P(-1) \), an unknown quantity.

We remedy this difficulty by realizing that

\[ \int_{-1}^{n+1} p(x)dx = \sum_{r=0}^{n+1} b_r T_r(-1) = 0. \]

This forces

\[ \frac{1}{2} b_0 = b_1 - b_2 + \ldots + (-1)^{n-1} b_n + (-1)^n b_{n+1}, \]

so that now we possess an expression for \( b_0 \) (and thus \( P(x) \)) which does not involve any unknown or non-computable quantities.

Also, the value of the definite integral,

\[ \int_{-1}^{1} \sum_{r=0}^{\infty} a_r T_r(x)dx, \]

where the \( a_r \) are constants, is often valuable. Using the integration formulas (36) and (37),
\[ \int_{-\infty}^{\infty} a_r T_r(x) \, dx = \sum_{r=0}^{\infty} a_r \int_{-1}^{1} T_r(x) \, dx \]

\[ = \frac{a_0}{2} T_1(x) + \frac{a_1}{4} (T_2(x)+1) + \sum_{r=2}^{\infty} a_r \left( \frac{1}{r+1} T_{r+1}(x) - \frac{1}{r-1} T_{r-1}(x) \right) \]

\[ = a_0 + \sum_{r=2}^{\infty} a_r \left( \frac{1}{r+1} - \frac{1}{r-1} \right) \]

\[ = a_0 - \sum_{r=2}^{\infty} \frac{(1+(-1)^r)}{r} a_r. \quad (51) \]

To integrate the two corresponding series, \( \sum_{r=0}^{\infty} a_r T_r^*(x) \) and \( \sum_{r=0}^{\infty} a_r T_r^*(x) \), in a similar fashion, we first need to note that by (36),

\[ \int_{-\infty}^{\infty} T_r^*(x) \, dx = \int_{-1}^{1} (2x-1) \, dx = \left[ \frac{1}{4} (r+1) T_r^*(x) - \frac{1}{r-1} T_{r-1}^*(x) \right] + \text{constant} \]

for \( r = 2, 3, \ldots \). Also, by (37)

\[ \int_{-\infty}^{\infty} T_0^*(x) \, dx = x = \frac{1}{2} [T_0^*(x)+T_0^*(x)] = \frac{1}{2} T_1^*(x) + \text{constant}, \]

and

\[ \int_{-\infty}^{\infty} T_1^*(x) \, dx = x^2 - x = \frac{1}{8} T_2^*(x) + \text{constant}. \]

We are now equipped to perform indefinite integration.

\[ \int_{0}^{x} \sum_{r=0}^{n} a_r T_r^*(x) \, dx = \sum_{r=0}^{n} a_r \int_{0}^{x} T_r^*(x) \, dx, \]

\[ = a_0 \left( T_0^*(x) + T_1^*(x) \right) + a_1 \left( \frac{1}{8} T_2^*(x) - T_0^*(x) \right) + \sum_{r=2}^{n} \frac{a_r}{4} \left( \frac{1}{r+1} T_{r+1}^*(x) - \frac{1}{r-1} T_{r-1}^*(x) \right) \]

\[ = a_0 - \frac{a_1}{8} T_0^*(x) + \sum_{r=1}^{n-1} \frac{1}{4r} (a_{r-1} - a_{r+1}) T_r^*(x) + \frac{a_{n-1}}{4n} T_1^*(x) + \frac{a_n}{4(n+1)} T_n^*(x) \]

\[ = \sum_{r=0}^{n+1} b_r T_r^*(x), \quad (52) \]
where \( b_n = \frac{a_{n-1}}{4n}, \ b_{n-1} = \frac{a_n}{4(n+1)}, \ b_r = \frac{1}{4r} (a_{r-1} - a_{r+1}) \) for \( r = 1,2,\ldots,n-1. \) As for \( b_0, \) notice that

\[
\int_0^1 \sum_{r=0}^{n} a_r T_r(x) dx = \sum_{r=0}^{n} b_r T_r(0) = 0.
\]

This means that \( b_0 = 2[b_1 - b_2^2 + \ldots + (-1)^n b_{n+1}]. \) Now we compute the definite integral

\[
\int_0^1 \sum_{r=0}^{n} a_r T_r(x) dx.
\]

By (52) above,

\[
\int_0^1 \sum_{r=0}^{n} a_r T_r(x) dx = \sum_{r=0}^{n} b_r T_r(1) - \sum_{r=0}^{n} b_r T_r(0).
\]

From (47),

\[
\int_0^1 \sum_{r=0}^{n} a_r T_r(x) dx = \left[ \frac{b_0}{2} + b_1 + b_2^2 + \ldots \right] - \left[ \frac{b_0}{2} - b_1 + b_2 - b_3^2 + \ldots \right]
\]

\[
= 2(b_1 + b_3^2 + b_5^2 + \ldots)
\]

\[
= 2 \left[ \frac{1}{4} (a_0 - a_2) + \frac{1}{4 \cdot 3} (a_2 - a_4) + \frac{1}{4 \cdot 5} (a_4 - a_6) + \ldots \right]
\]

\[
= \frac{a_0}{2} - \frac{1}{2} \sum_{r=0}^{\infty} \left( \frac{1 + (-1)^r}{r^2 - 1} \right) a_r,
\]

which is exactly half of (51), the value of \( \int_{-1}^{1} \sum_{r=0}^{\infty} c_r T_r(x), \)

as would be expected.

Turning to the term-by-term differentiation of the finite Chebyshev series,

\[
p(x) = \sum_{r=0}^{n} a_r T_r(x), \quad (53)
\]
we seek to compute the coefficients $c_r$ in

$$p'(x) = \sum_{r=0}^{n-1} c_r T_r(x)$$

(54)

which is a polynomial of degree $n-1$. Making use of (36) and (37), equation (54) can be integrated to yield

$$p(x) = \frac{c_0}{2} T_1(x) + \frac{c_1}{4} T_2(x) + \sum_{r=2}^{n-1} \frac{c_r}{2} \left[ \frac{1}{r+1} T_{r+1}(x) - \frac{1}{r-1} T_{r-1}(x) \right] + \text{constant}.$$  

(55)

The constant term in (55) is, of course, the constant term of $p(x)$, namely $\frac{a_0}{2} T_0(x)$. Hence, considering the definition of $p(x)$ in (53), (55) becomes

$$\sum_{r=0}^{n} a_r T_r(x) = \frac{a_0}{2} T_0(x) + \sum_{r=1}^{n-2} \frac{1}{2r} [c_{r-1} - c_{r+1}] T_r(x) + \frac{c_{n-2}}{2(n-1)} T_{n-1}(x) + \frac{c_{n-1}}{2n} T_n(x).$$  

(56)

If the coefficients of the $T_r(x)$ on both sides of equation (56) are set equal, then $c_{r-1} = 2r a_r + c_{r+1}$ for $r = 1, 2, \ldots, n-2,$

and $c_{n-1} = 2 a_n$, $c_{n-2} = 2(n-1) a_{n-1}$.

Thus, from this recurrence relation the unknown $c_r$ can be made known, and the differentiation is complete.

The range $[0,1]$ can be treated similarly using the polynomials $T_r^*(x)$. Let

$$p(x) = \sum_{r=0}^{n} a_r T_r^*(x).$$  

(57)
Integrate \( \sum_{r=0}^{n-1} c_r T_r^*(x) \) to get

\[
p(x) = \frac{c_0}{4} T_1^*(x) + \frac{c_1}{8} T_2^*(x) + \frac{1}{4} \sum_{r=2}^{n-1} c_r \left[ \frac{1}{r+1} T_{r+1}^*(x) - \frac{1}{r-1} T_{r-1}^*(x) \right],
\]

\[
= \sum_{r=1}^{n-2} \left[ \frac{c_{r-1}}{4r} - \frac{c_{r+1}}{4r} \right] T_r^*(x) + \frac{c_{n-2}}{4(n-1)} T_{n-1}^*(x) + \frac{c_{n-1}}{4n} T_n^*(x),
\]

\[
= \sum_{r=0}^{n} a_r T_r^*(x).
\]

Equating the coefficients of the \( T_r^*(x) \) yields this recurrence relation for the \( c_r \),

\[
c_{n-1} = 4n a_n, \quad c_{n-2} = 4(n-1) a_{n-1},
\]

\[
c_{r-1} = 4r a_r + c_{r+1}, \text{ for } r = 1, 2, \ldots, n-2.
\]

Note that the differentiation above was performed only on the finite series (53) and (57). If we were to try to apply the same method of differentiation to an infinite Chebyshev series, then some caution must be used because the resulting expressions for the \( c_r \) are infinite series. Of course, these infinite series must converge quickly enough to give an acceptable error in the \( c_r \).
REFERENCES


CHAPTER III

APPROXIMATION AND REPRESENTATION OF
FUNCTIONS BY CHEBYSHEV POLYNOMIALS

Suppose that \( f(x) \) is a function defined over the interval \([a, b]\) and that \( P_n(x) \) is a polynomial of degree \( n \). A polynomial which minimizes the norm \( || \cdot || \) defined by

\[
||f-P_n||_\infty \max_{a \leq x \leq b} |f(x) - P_n(x)| = D(f, P_n)
\]

is conventionally called a polynomial of "best" approximation. As we pointed out in Chapter I, the Chebyshev polynomials are related to the norm above in some way, and the following sequence of theorems will show how they are related.

We state the first theorem without its proof, one version of which can be found in Isaacson and Keller's *Analysis of Numerical Methods*.

**Theorem 3.1:** Let the semi-norm \( || \cdot || \) be defined in \([a, b]\), the linear space of functions defined on the closed bounded interval \([a, b]\), and let there exist positive numbers \( m \) and \( M \) which satisfy \( 0 < m \leq \sum_{j=0}^{n} b_j x^j \leq M, n = 0, 1, \ldots \), for all \( \{b_j\} \) such that \( \sum_{j=0}^{n} |b_j|^2 = 1 \). Then for any integer \( n \) and \( f(x) \) in \([a, b]\), there exists a polynomial of degree at most \( n \) for which

\[
||f-P_n|| \text{ attains its minimum over all such polynomials.}
\]
Equipped with the preceding theorem, we now proceed to show the existence of a polynomial of best approximation for the general case.

**Theorem 3.2.** Let \( f(x) \) be a given function continuous in \([a,b]\). Then for any integer \( n \), there exists a polynomial \( \hat{P}_n(x) \) of degree at most \( n \) which minimizes \( \| f - P_n \|_\infty \) over all polynomials of degree at most \( n \).

**Proof.** The method of proof shall be to verify the hypotheses of Theorem 3.1 to obtain the existence of \( \hat{P}_n(x) \). Notice that

\[
\| \sum_{j=0}^{n} b_j x^j \|_\infty = \max_{a \leq x \leq b} \left| \sum_{j=0}^{n} b_j x^j \right|
\]

\[
\leq \max_{a \leq x \leq b} \sum_{j=0}^{n} |b_j x^j|
\]

\[
\leq \max_{a \leq x \leq b} \sum_{j=0}^{n} |x^j|
\]

\[
= \sum_{j=0}^{n} |b_j| = M.
\]

Hence, we have obtained an upper bound \( M \) for \( \| \sum_{j=0}^{n} b_j x^j \|_\infty \).

Let \( \phi \), a function of the variables \( \{b_i\} \), be defined by

\[
\phi(b_0, b_1, \ldots, b_n) = \| \sum_{j=0}^{n} b_j x^j \|_\infty.
\]

We plan to show that \( \phi \) is a continuous function. Choose some \( \varepsilon > 0 \). Let \( \delta = (\delta_0, \delta_1, \ldots, \delta_n) \) where \( \delta_i = \varepsilon / ((n+1)M) \), for all \( i \).
If \( \bar{x} = (x_0, \ldots, x_n) \) such that \( |a_i - x_i| < \delta_i \), for all \( i \),
then
\[
|\phi(b_0, \ldots, b_n) - \phi(x_0, \ldots, x_n)| \leq \| \sum_{j=0}^{n} b_j x_j - \sum_{j=0}^{n} x_j x_j \|_\infty
\]
\[
= \| \sum_{j=0}^{n} (b_j - x_j) x_j \|_\infty = \max_{a \leq x \leq b} \| \sum_{j=0}^{n} (b_j - x_j) x_j \| \\
= \max_{a \leq x \leq b} \sum_{j=0}^{n} |b_j - x_j| \cdot |x_j| \leq M \sum_{j=0}^{n} |b_j - x_j| \\
< M \sum_{j=0}^{n} \epsilon / ((n+1)M) = \epsilon.
\]
This proves that \( \phi \) is a continuous function of the variables \{b_j\}.

Let \( A \equiv \{ a \mid \sum_{j=0}^{n} a_j^2 = 1 \} \). Since \( A \) is a closed and bounded
set and since \( \phi \) is continuous, then by a certain Weierstrauss
theorem, we know that \( \phi \) attains its minimum on \( A \). That is,
there exists \( \hat{a} = (\hat{a}_0, \ldots, \hat{a}_n) \in A \) such that
\[
\min_{\hat{a} \in A} \| \sum_{j=0}^{n} \hat{a}_j x_j \|_\infty = \| \sum_{j=0}^{n} \hat{a}_j x_j \|_\infty = m.
\]
Since \( \sum_{j=0}^{n} \hat{a}_j^2 = 1 \), then \( \sum_{j=0}^{n} \hat{a}_j x_j \) is not the zero polynomial
which prevents \( \| \sum_{j=0}^{n} \hat{a}_j x_j \|_\infty \) from vanishing, hence \( m \neq 0 \), Q.E.D.

Theorem 3 will lead us up to a characterization of the
best polynomial and to the establishment of its uniqueness.

**Theorem 3.3** (De La Vallée-Poussin). Let an nth degree
polynomial \( P_n(x) \) have the deviations from \( f(x) \) given by
f(x_j) - P_n(x_j) = (-1)^j e_j, \ j = 0, \ldots, n+1, \text{ where}

0 \leq x_0 < x_1 < \ldots < x_{n+1} \leq b \text{ and all } e_j > 0 \text{ or else all } e_j < 0. \text{ Then}

\min_j |e_j| \leq d_n(f) = \min\{D(f,R):R(x) \text{ is a polynomial of degree } n \text{ or less}\}.

Proof. Assume that for some polynomial Q_n(x) of degree n, D(f,Q_n) < \min_j |e_j|. Note that Q_n(x) - P_n(x) is a polynomial of degree at most n. Also,

Q_n(x) - P_n(x) = (f(x) - P_n(x)) - (f(x) - Q_n(x))

has the same sign at the points x_j as does f(x) - P_n(x).

This is because

|f(x_j) - Q_n(x_j)| \leq \max_{a \leq x \leq b} |f - Q_n|

= D(f,Q_n) < \min_j |e_j| \leq |e_j|

= |f(x_j) - P_n(x_j)|, \ j = 0, \ldots, n+1.

Thus, Q_n(x) - P_n(x) has at least n+2 sign changes which cause at least n+1 zeros of Q_n(x) - P_n(x), a polynomial of degree at most n. This implies that Q_n(x) - P_n(x) = 0 and Q_n(x) = P_n(x), an impossibility, since

|f(x_j) - Q_n(x_j)| < |f(x_j) - P_n(x_j)|.

Hence, for every polynomial n(x) of degree at most n,

\min_j |e_j| \leq D(f, n), \ Q.E.D.

Theorem 3.4. (Chebyshev) A polynomial of degree at most P_n(x) is a best approximation of degree at most n to f(x) in [a,b] if and only if f(x) - P_n(x) assumes the values

\pm D(f,P_n) = \max_{a \leq x \leq b} |f(x) - P_n(x)|.
with alternate changes of sign, at least \( n+2 \) times in \([a,b]\).

This best approximation polynomial \( P_n(x) \) is unique.

**Proof.** For the sufficiency, suppose that some polynomial \( P_n(x) \) oscillates as in the statement of the theorem. Let \( x_j, j = 0,1,\ldots,n+1 \), be \( n+2 \) points at which the maximum deviation \( |D(f,P_n)| \) is attained with alternate sign changes so that \( a\leq x_0 < x_1 < \ldots < x_{n+1} \leq b \). We have now satisfied the hypotheses of Theorem 3.3 so that \( \min \{|e_j| = D(f,P_n)\leq \min\{D(f,Q_n) : Q_n(x) \text{ is a polynomial of degree } n \text{ or less}\} \leq D(f,P_n) \). This implies that \( D(f,P_n) = \min\{D(f,Q_n) : Q_n(x) \text{ is a polynomial of degree } n \text{ or less}\} \) and hence that \( P_n(x) \) is a best approximation of degree at most \( n \) to \( f(x) \) in \([a,b]\).

For the necessity, let \( P_n(x) \) be a polynomial of best approximation. We will show that if \( f(x) - P_n(x) \) does not assume the values \( \pm D(f,P_n) \), with alternate changes in sign, at least \( n+2 \) times in \([a,b]\), then \( P_n(x) \) is not a best polynomial approximation, a contradiction.

Let \( f(x) - P_n(x) \) attain the values \( \pm D(f,P_n) \), with alternate sign changes, at most \( k \) times where \( 1 \leq k \leq n+1 \). We assume without loss of generality, that
\[
f(x_j) - P_n(x_j) = (-1)^j D(f,P_n),
\]
j = 0,1,\ldots,k, where \( a\leq x_1 < x_2 < \ldots < x_k \leq b \). Let \( \xi_1 = (\eta_1 + \zeta_1)/2 \)
where \( \eta_1 = g.l.b.\{\eta \in \{a \leq \eta \leq x_2 \text{ and } f(\eta) - P_n(\eta) = D(f,P_n)\} \} \) and \( \zeta_1 = u.b. \{\zeta | a \leq \zeta \leq x_2 \text{ and } f(\zeta) - P_n(\zeta) = -D(f,P_n)\} \). Note that \( x_1 \leq \xi_1 \) and \( \eta_1 \leq x_2 \). For such a pair, \( \eta_1 \) and \( \zeta_1 \), there are three possibilities, \( \xi_1 > \eta_1, \xi_1 = \eta_1, \) and \( \xi_1 < \eta_1 \). If
\( \zeta_1 > \eta_1 \), then \( x_1 \leq \eta_1 < \zeta_1 \leq x_2 \), which yields \( k+3 \) alternations of sign of \( f(x) - P_n(x) \) a violation of our assumption of only \( k \) changes. Suppose that \( \zeta_1 = \eta_1 \). Then at \( \xi_1 = (\eta_1 + \zeta_1)/2 = \eta_1 \), \( f(x) - P_n(x) \) has vertical slope, that is, \( f(x) - P_n(x) \) is not continuous at \( \xi_1 \), another violation. Hence,

\( x_1 < \xi_1 < \eta_1 < x_2 \). Note that on the interval \([a, \xi_1]\) the deviation \( f(x) - P_n(x) \) takes on the extreme value \(-D(f, P_n)\) at least once and doesn't take on the value \( D(f, P_n) \) at all.

Now, let \( \xi_2 = (\eta_2 + \zeta_2)/2 \), where \( \eta_2 = \text{g.l.b.} \{ \eta \mid x_2 \leq \eta \leq x_3 \} \) and \( f(\eta) - P_n(\eta) = -D(f, P_n) \), and \( \zeta_2 = \text{l.u.b.} \{ \zeta \mid x_2 \leq \zeta \leq x_3 \} \) and \( f(\zeta) - P_n(\zeta) = +D(f, P_n) \). We could now show, in a manner similar to that above for \( \xi_1 \), that \( x_2 < \eta_2 < \zeta_2 \leq x_3 \) and that on the interval \([\xi_1, \xi_2]\), \( f(x) - P_n(x) \) takes on the value \( D(f, P_n) \) at least once and never takes on the value \(-D(f, P_n)\).

It should now be clear that there exist points \( \xi_1, \xi_2, \ldots, \xi_{k-1} \) such that

\[
x_1 < \xi_1 < x_2 < \xi_2 < \ldots < x_{k-1} < \xi_{k-1} < x_k
\]

and such that, for each of the \( k \) intervals \([a, \xi_1], [\xi_1, \xi_2], \ldots, [\xi_{k-1}, b]\), the deviation \( f(x) - P_n(x) \) takes on only one of the two extreme deviations \( \pm D(f, P_n) \), the negative of the extreme taken on by the preceding interval, and is bounded away from the extreme of opposite sign \( \pm D(f, P_n) \). In other words, there exists some \( \varepsilon > 0 \) such that, for \( x \) in the "odd" intervals, \([a, \xi_1], [\xi_2, \xi_3], [\xi_4, \xi_5], \ldots\),

\[
-D(f, P_n) + \varepsilon < f(x) - P_n(x) \leq D(f, P_n),
\]

(1)
and such that, for \( x \) in the "even" intervals, \([\xi_1, \xi_2],
\[\xi_3, \xi_4], [\xi_5, \xi_6], \ldots\),
\[-D(f, P_n) + \varepsilon < f(x) - P_n(x) \leq D(f, P_n). \tag{2}\]

The polynomial \( r(x) = (x-\xi_1)(x-\xi_2)\cdots(x-\xi_{k-1}) \) is of
degree \( k-1 \) and has only one sign throughout each of the \( k \)
intervals. This sign alternates from one interval to the
next. Let the maximum value of \( |r(x)| \) in \([a, b]\) be given
by \( M \). Now define \( q(x) = (-1)^k r(x)/2M \) and look at the poly-
nomial \( Q_n(x) = P_n(x) + e q(x) \). \( Q_n(x) \) is of degree \( n \) since
deg\( q(x) = deg(r(x)) = k-1 \leq n \), and yet we will soon show
that \( D(f, Q_n) < D(f, P_n) \), exposing \( P_n(x) \) as an imposter of
the best approximation polynomial. In the interior of any
of the odd intervals, \( q(x) < 0 \), since \( (-1)^k \) and \( r(x) \) will be
of opposite sign. Also, we can say that \( -1/2 \leq q(x) < 0 \).
Similarly, in the interior of the even intervals, we get
that \( 0 < q(x) \leq 1/2 \).

Recalling the above inequalities (1) and (2) and that
\( P_n(x) = Q_n(x) - e q(x) \), we can say that, for \( x \) in odd
intervals,
\[-D(f, P_n) - e q(x) \leq f(x) - Q_n(x) < D(f, P_n) - e(l+q(x)), \]
and, for \( x \) in even intervals,
\[-D(f, P_n) + e(1-q(x)) < f(x) - Q_n(x) \leq D(f, P_n) - e q(x).\]
Since, for \( x \) in odd intervals, \(-1/2 \leq q(x) < 0\), then
\[-D(f, P_n) + e/2 \leq f(x) - Q_n(x) < D(f, P_n) - e, \]
which leads to \( |f(x) - Q_n(x)| < D(f, P_n) \). Since, for \( x \) in even
intervals, \( 0 < q(x) \leq 1/2 \), then,
\[-D(f, P_n) + e < f(x) - Q_n(x) \leq D(f, P_n) - e/2,\]
which leads to $|f(x) - Q_n(x)| < D(f, P_n)$. Therefore, $D(f, Q_n) < D(f, P_n)$, a contradiction, which forces $f(x) - P_n(x)$ to assume the values $\pm D(f, P_n)$, with alternate changes in sign, at least $n+2$ times.

To show uniqueness, suppose that there are two best approximation polynomials, $P_n(x)$ and $Q_n(x)$, that is, $D(f, P_n) = D(f, Q_n) = d_n(f)$.

Notice that

$$|f(x) - 1/2(P_n(x) + Q_n(x))|$$

$$= |1/2(f(x) - P_n(x)) + 1/2(f(x) - Q_n(x))|$$

$$\leq 1/2 |f(x) - P_n(x)| + 1/2 |f(x) - Q_n(x)| \leq d_n(f).$$

This means that $1/2(P_n(x) + Q_n(x))$ is another best polynomial approximation, and hence there exist points $x_j$ in $[a, b]$, $j = 0, \ldots, n+1$, such that $f(x_j) - 1/2(P_n(x_j) + Q_n(x_j)) = d_n(f)$, or, according to the inequalities, now turned equalities, in (3), $1/2 |f(x_j) - P_n(x_j)| + 1/2 |f(x_j) - Q_n(x_j)| = d_n(f)$, $j = 0, \ldots, n+1$. Suppose that $1/2(f(x_j) - P_n(x_j)) \neq 1/2(f(x_j) - Q_n(x_j))$ for some $j = 0, \ldots, n+1$. We say, without a loss of generality, that

$$1/2(f(x_j) - P_n(x_j)) > 1/2(f(x_j) - Q_n(x_j)),$$

which forces $f(x_j) - P_n(x_j) > d_n(f)$, an impossibility. Thus, $f(x_j) - P_n(x_j) = f(x_j) - Q_n(x_j)$, $j = 0, \ldots, n+1$. Stated another way,

$$(f(x_j) - P_n(x_j)) - (f(x_j) - Q_n(x_j)) = Q_n(x_j) - P_n(x_j) = 0,$$

$j = 0, \ldots, n+1$. This indicates that the polynomial $Q_n(x) - P_n(x)$,
of degree at most $n$, has $n+2$ distinct zeros. The only way this can happen is for $Q_n(x) = P_n(x)$. Q.E.D.

The above characterization enables us to easily produce the best approximation of degree at most $n-1$ to the polynomial, $p(x) = a_n x^n + q(x)$, where $q(x)$ is a polynomial of degree $n-1$ or less. We propose that $p^*(x) = p(x) - a_n 2^{1-n} T_n(x)$ is this approximation. By a property of orthogonal polynomials, $p(x)$ is uniquely expressible as

$$p(x) = d_0 T_0(x) + \ldots + d_n T_n(x), \quad (4)$$

for certain coefficients $d_0, \ldots, d_n$. Using the recursion formula (2.1) for the $T_r(x)$, one induces that $T_r(x) = 2^{r-1} x^r$ a polynomial of degree at most $r-1$, $r = 1, 2, \ldots$. Since $x^n$ appears only in $T_n(x)$, we must have $a_n = d_n 2^{n-1}$ or $d_n = a_n 2^{1-n}$. Hence, $p^*(x) = d_0 T_0(x) + \ldots + d_n T_n(x) - d_n T_n(x)$, showing that $p^*(x)$ is indeed a polynomial of degree at most $n-1$. Notice

$$\max_{-1 \leq x \leq 1} |p(x) - p^*(x)| = \max_{-1 \leq x \leq 1} |a_n 2^{1-n} T_n(x)| = |a_n 2^{1-n}|$$

since $\max_{-1 \leq x \leq 1} |T_n(x)| = 1$. Also notice that

$$a_n 2^{1-n} T_n(\xi_k) = a_n 2^{1-n} (-1)^k$$

for $\xi_k = \cos(k\pi/n)$, $k = 0, \ldots, n$.

Thus, the difference $p(x) - p^*(x)$ assumes the values $D(p, p^*)$, with alternate changes of sign, $n+1$ times in $[-1,1]$, and by Theorem 3.4, $p^*(x)$ is the best approximation of degree $n-1$ or less for $p(x)$.
In summary, once we have expressed $p(x)$ in the form (4), then simply dropping the last term produces the best approximation of degree $n-1$ or less. The maximum error in $[-1,1]$ is then $d_n$. This type of truncation is sometimes called Chebyshev economization or telescoping.

The most practical and, hence, the most valuable use of Chebyshev polynomials is in producing approximations to functions which are defined explicitly or which are defined discretely, as in a mathematical table. We first look into computing the continuous least-squares fit using Chebyshev polynomials which results in the already mentioned Chebyshev series. The Chebyshev series generally converges rapidly which saves in computational time. Also, it is easy to estimate an accurate upper bound for the error of this series. Such advantages have led to the common use of Chebyshev series subroutines for evaluating the elementary functions by digital computer, (2, p. 65). Also studied in this chapter are the discreet least-squares fit using Chebyshev polynomials and the Chebyshev approximation methods for rational functions.

Let $f(x)$ be a function defined on the interval $[-1,1]$, and suppose that we wish to find the polynomial $p(x)$, of degree $n$ or less such that

$$
\int_{-1}^{1} w(x) [f(x) - p(x)]^2 \, dx \leq \int_{-1}^{1} w(x) [f(x) - q(x)]^2 \, dx
$$
where \( w(x) \) is a given function defined on \([-1,1]\) and positive there, and where \( q(x) \) is any polynomial of degree \( n \) or less. If a set of \( n+1 \) polynomials, \( \{\phi_r(x)\}^n_{r=0} \), where \( \phi_r(x) \) is of degree \( r \), is available, then \( p(x) \) can be represented as

\[
p(x) = c_0\phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x).
\]

The object is to find those \( c_r \) for which

\[
E(c_0, c_1, \ldots, c_n) = \int_{-1}^{1} w(x) \left[ f(x) - \sum_{r=0}^{n} c_r\phi_r(x) \right]^2 dx
\]

is a minimum. The problem can be solved by recalling that at a minimum of \( E \),

\[
\frac{\partial E}{\partial c_i} = 0, \quad i = 0, 1, \ldots, n. \tag{5}
\]

Now, we use the definition of \( E \) to see what (5) actually requires. We have, for \( i = 0, 1, \ldots, n \)

\[
\frac{\partial E}{\partial c_i} = \frac{\partial}{\partial c_i} \left\{ \int_{-1}^{1} w(x) \left[ f(x) - \sum_{r=0}^{n} c_r\phi_r(x) \right]^2 dx \right\} = 0.
\]

This leads to the following sequence of equations:

\[
\frac{\partial E}{\partial c_i} = \int_{-1}^{1} \frac{\partial}{\partial c_i} \left[ w(x) \left[ f(x) - \sum_{r=0}^{n} c_r\phi_r(x) \right]^2 \right] dx
\]

\[
= 2 \int_{-1}^{1} w(x) \phi_i(x) [f(x) - \sum_{r=0}^{n} c_r\phi_r(x)] dx
\]

\[
= 2 \int_{-1}^{1} w(x) \phi_i(x) f(x) dx - 2 \sum_{r=0}^{n} c_r \int_{-1}^{1} w(x) \phi_i(x) \phi_r(x) dx
\]

\[
= 0. \tag{6}
\]
If it is true that the $\phi_i(x)$ are orthogonal, i.e.,

$$
\int_{-1}^{1} w(x)\phi_i(x)\phi_r(x)dx \begin{cases} = 0, & \text{for } i \neq r, \\ \neq 0, & \text{for } i = r,
\end{cases}
$$

then from (6),

$$
\int_{-1}^{1} w(x)\phi_i(x)f(x)dx - c_i \int_{-1}^{1} w(x)\phi_i^2(x)dx = 0,
$$

$$
i = 0,1,\ldots,n. \tag{7}
$$

The coefficient matrix of the linear system (7) is invertible, and therefore the following solution of this system is unique:

$$
c_i = \frac{\int_{-1}^{1} w(x)\phi_i(x)f(x)dx}{\int_{-1}^{1} w(x)\phi_i^2(x)dx}, \text{ for } i = 0,1,\ldots,n.
$$

The problem is solved if the above "ifs" can be deleted by our actually exhibiting a set of such polynomials, $\phi_i(x)$. Toward this end, we assign $\phi_i(x) = T_i(x)$ and $w(x) = (1-x^2)^{-1/2}$ and study

$$
\int_{-1}^{1} (1-x^2)^{-1/2}T_i(x)T_r(x)dx
$$

for case 1: $i \neq r$, and for case 2: $i = r$.

Case 1: $i \neq r$.

$$
\int_{-1}^{1} (1-x^2)^{-1/2}T_i(x)T_r(x)dx
$$

$$
= \int_{0}^{\pi} (1-\cos^2 \theta)^{-1/2} \cos i \theta \cos r \theta (-\sin \theta) d\theta,
$$

where $x = \cos \theta$

$$
= \int_{0}^{\pi} \cos i \theta \cos r \theta d\theta = \frac{\sin[(i-r)\theta]}{2(i-r)} \bigg|_{0}^{\pi} = 0.
$$
Case 2: \( i = r \).
\[
\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_i(x) T_r(x) \, dx
\]
\[
= \int_{0}^{\pi} \cos^2 r \theta \, d\theta
\]
\[
= \left\{ \begin{array}{ll}
\frac{\theta}{2} + \frac{1}{4r} \sin 2r \theta \bigg|_{0}^{\pi} &= \frac{\pi}{2}, \text{ for } r \neq 0, \\
\int_{0}^{\pi} d\theta &= \pi, \text{ for } r = 0.
\end{array} \right.
\]
Thus, the shoe fits for the \( T_i(x) \) and we can now say that the polynomial,
\[
p(x) = \sum_{r=0}^{n} c_r T_r(x),
\]
where
\[
c_r = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_r(x) f(x) \, dx, \text{ for } r = 0,1,\ldots,n,
\]
is the polynomial such that
\[
\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} [f(x) - p(x)]^2 \, dx
\]
is a minimum over all polynomials of degree \( n \) or less.

The finite series \( \sum_{r=0}^{n} c_r T_r(x) \), we call the continuous least-squares fit for \( f(x) \), and the infinite series \( \sum_{r=0}^{\infty} c_r T_r(x) \), we call the Chebyshev series for \( f(x) \).

**Example 1.** We shall now compute the continuous least-squares fit of degree 4 over the interval \([-1,1]\) for the function \( f(x) = x + \cos^{-1} x \). By (8), the fit is
\[ p_4(x) = \frac{c_0}{2} T_0(x) + c_1 T_1(x) + c_2 T_2(x) \]
\[ + c_3 T_3(x) + c_4 T_4(x), \]
where
\[ c_0 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_0(x)(x+\cos^{-1} x)dx = \pi \]
\[ c_1 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_1(x)(x+\cos^{-1} x)dx = 1 - \frac{4}{\pi} \]
\[ c_2 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_2(x)(x+\cos^{-1} x)dx = 0 \]
\[ c_3 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_3(x)(x+\cos^{-1} x)dx = -\frac{4}{9\pi} \]
\[ c_4 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_4(x)(x+\cos^{-1} x)dx = 0. \]
Therefore,
\[ p_4(x) = \frac{\pi}{2} + (1-\frac{4}{\pi}) T_1(x) - \frac{4}{9\pi} T_3(x). \]

By (2.2),
\[ p_4(x) = \frac{\pi}{2} + (1-\frac{8}{3\pi}) x - \frac{16}{9\pi} x^3 \]
\[ = 1.5707963 + 0.1511737x - 0.5658842x^3. \]

If Chebyshev series are to be used as representations of functions, then it is helpful to know something about the convergence properties of these series. For reasons which we will see later, we first study the convergence of the Fourier series
\[ a_0/2 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \]
in which
\[ a_n = \frac{1}{\pi} \int_{0}^{2\pi} f(x) \cos nx \, dx \text{ and } b_n = \frac{1}{\pi} \int_{0}^{2\pi} f(x) \sin nx \, dx. \]

First, a few definitions below will be helpful in proving the sequence of theorems which will follow thereafter.
Definition. If $h \to 0$ purely through positive values of $h$, then we say that $\Psi(x+h) \to \Psi(x)$ if such a limit exists.

Definition. A function $\Psi(x)$ is said to be piecewise continuous in a finite interval $[a,b]$ if

(i) the interval $[a,b]$ can be subdivided into a finite number, $m$ say, of intervals, $[(a,a_1),(a_1,a_2),\ldots,(a_r,a_{r+1}),\ldots,(a_{m-1},b)]$, in each of which $\Psi(x)$ is continuous, and if
(ii) $\Psi(a^+), \Psi(b^-), \Psi(a_r^-), \Psi(a_r^+)$ for $r = 1,2,\ldots,n-1$ are all finite.

Because of condition (ii) and the continuity of $\Psi(x)$, we know that there exist constants $M_{r+1}$, $r = 1,\ldots,m$, such that $|\Psi(x)| < M_{r+1}$ for $a_r \leq x \leq a_{r+1}$, $r = 0,\ldots,m-1$, where $a_0 = a$ and $a_m = b$. So, if we define $M = \text{u.b.}\{M_1,M_2,\ldots,M_m\}$, then $|\Psi(x)| < M$ for $a \leq x \leq b$. It follows that if $\Psi(x)$ is piecewise continuous, then $\int_a^b \Psi(x)\,dx$ is finite.

Definition. A function $f(x)$ has a right-hand derivative at the point $x = \xi$ if and only if the limit,

$$\lim_{h \to 0} [f(\xi+h)-f(\xi^+)]/h$$

where $h \to 0$ only through positive values of $h$, is finite.

Definition. A function has a left-hand derivative at the point $x = \xi$ if and only if the limit,

$$\lim_{h \to 0} [f(\xi+h)-f(\xi^-)]/h$$

where $h \to 0$ only through negative values of $h$, is finite.
Theorem 3.5. If \( \psi(u) \) is piecewise continuous in the interval \( a \leq u \leq b \), then \( \lim_{N \to \infty} \int_{a}^{b} \psi(u) \sin(Nu) du = 0 \).

Theorem 3.5 is a valid one, and we shall employ it in its above form later in the paper. However, its proof is difficult; so, we shall instead prove a less general case by assuming in addition that \( \psi'(u) \) is piecewise continuous in \( a \leq x \leq b \).

Proof. We again divide the interval \([a,b]\) into the intervals \([a = a_0,a_1),(a_1,a_2),\ldots,(a_{m-1},a_m = b]\). It follows that
\[
\int_{a}^{b} \psi(u) \sin Nu du = \sum_{r=0}^{m-1} \int_{a_r}^{a_{r+1}} \psi(u) \sin Nu du.
\]
For \( r = 0,\ldots,m-1 \), we use integration by parts to get
\[
\int_{a_r}^{a_{r+1}} \psi(u) \sin Nu du = \left[ -\psi(u) \cdot (\cos Nu)/N \right]_{a_r}^{a_{r+1}} + \frac{1}{N} \int_{a_r}^{a_{r+1}} \psi'(u) \cos Nu du.
\]
(9)
The last integral is bounded so that the whole right-hand side of (9) is less than \( M_{r+1}/N \) where \( M_{r+1} \) is a constant. If \( M = \text{l.u.b.}\{M_1,M_2,\ldots,M_m\} \), then
\[
\left| \int_{a}^{b} \psi(u) \sin Nu du \right| < Mm/N.
\]
Since \( M \) and \( m \) are finite, then as \( N \to \infty \), \( Mm/N \to 0 \), forcing
\[
\int_{a}^{b} \psi(u) \sin Nu du \to 0. \quad \text{Q.E.D.}
\]

Theorem 3.6. If \( \psi(u) \) is piecewise continuous in any closed interval \( \gamma \leq u \leq a \) where \( 0 < \gamma \leq a \) and if \( \psi(u) \) has a right-
hand derivative at \( u = 0 \), then

\[
\lim_{N \to \infty} \int_0^a \psi(u) \left( \frac{\sin Nu}{u} \right) du = \frac{\pi \psi(0^+)}{2}.
\]

**Proof.** Note that \( \psi(u) = \psi(0^+) + [\psi(u) - \psi(0^+)] \). So,

\[
\int_0^a \psi(u) \left( \frac{\sin Nu}{u} \right) du = \psi(0^+) \int_0^a \left( \frac{\sin Nu}{u} \right) du + \int_0^a \phi(u) \sin Nu \ du \tag{10}
\]

where \( \phi(u) = \frac{[\psi(u) - \psi(0^+)]}{u} \). Let \( \xi \equiv Nu \) so that

\[
\int_0^a \left( \frac{\sin Nu}{u} \right) du = \int_0^a \left( \frac{\sin \xi}{\xi} \right) d\xi.
\]

As \( N \to \infty \),

\[
\int_0^a \left( \frac{\sin \xi}{\xi} \right) d\xi + \int_0^\infty \left( \frac{\sin \xi}{\xi} \right) d\xi = \frac{\pi}{2}. \tag{11}
\]

Since \( \psi(u) \) is piecewise continuous in any interval \( \gamma \leq u \leq a \) where \( 0 < \gamma \leq a \), then \( \phi(u) \) has the same property. We were given that \( \lim_{h \to 0} \frac{[\psi(0^+ + h) - \psi(0^+)]}{h} \) is finite where \( h \to 0 \) only through positive values of \( h \), but this statement is precisely the same as saying that \( \phi(0^+) \) is finite. This enables us to say that \( \phi(u) \) is piecewise continuous in \( 0 \leq u \leq a \). An application of Theorem 3.5 gives

\[
\lim_{N \to \infty} \int_0^a \phi(u) \sin Nu \ du = 0.
\]

In light of this, looking back at (10) and (11) reveals that

\[
\lim_{N \to \infty} \int_0^a \psi(u) \left( \frac{\sin Nu}{u} \right) du = \frac{\pi \psi(0^+)}{2}. \quad \text{Q.E.D.}
\]
Theorem 3.7. If $\psi(u)$ is piecewise continuous in the open interval $a<u<b$ and if $\psi(u)$ has right and left derivatives at a point $u = x$, where $a<x<b$, then

$$\lim_{N \to \infty} \int_a^b \psi(u) \frac{[\sin N(u-x)]}{(u-x)} du = \pi \frac{[\psi(x^+) + \psi(x^-)]}{2}.$$  

Proof. We begin by dividing up the integral in question as

$$\int_a^b \psi(u) \frac{[\sin N(u-x)]}{(u-x)} du = \int_a^x \psi(u) \frac{[\sin N(u-x)]}{(u-x)} du + \int_x^b \psi(u) \frac{[\sin N(u-x)]}{(u-x)} du. \quad (12)$$

The change of variable, $u = x - \xi$, precipitates

$$\int_a^x \psi(u) \frac{[\sin N(u-x)]}{(u-x)} du = \int_0^{x-a} \phi(\xi) \frac{[\sin N\xi]}{\xi} \ d\xi$$

where $\phi(\xi) = \psi(x-\xi)$. We now assert that $\phi(\xi)$ which is equal to $\psi(x-\xi)$ has a right-hand derivative at $\xi = 0$, namely, the left-hand derivative of $\psi(u)$ at $u = x$. This fact coupled with the fact that $\phi(\xi)$ is piecewise continuous in $[0,x-a]$ allows us to employ Theorem 3.6 to conclude that

$$\lim_{N \to \infty} \int_0^{x-a} \phi(\xi) \frac{[\sin N\xi]}{\xi} \ d\xi = \pi \phi(0+)/2 = \pi \psi(x^-)/2.$$  

For the treatment of the second integral of (12), we proceed similarly, this time letting $u = x + \eta$ be the change
of variable, to get
\[ \int_{0}^{b} x \psi(x+\eta)(\sin N\eta)/\eta \, d\eta = \pi \psi(x+)/2. \quad \text{Q.E.D.} \]

In order to state the next theorem in an economical way, we need the following definition.

**Definition.** A function \( f(x) \) is said to be **piecewise smooth** in an interval \((a,b)\) if and only if \( f(x) \) is piecewise continuous in \((a,b)\) and \( f(x) \) has right-hand and left-hand derivatives at every point in \((a,b)\).

**Theorem 3.8 (Fourier's Theorem).** If \( f(x+2\pi) = f(x) \) for all \( x \) and if \( f(x) \) is piecewise-smooth in the interval \((0,2\pi)\), then the Fourier series

\[ S(x) = a_0/2 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \quad (13) \]

where
\[
\begin{align*}
a_n &= \frac{1}{\pi} \int_{0}^{2\pi} f(u) \cos nu \, du, \\
b_n &= \frac{1}{\pi} \int_{0}^{2\pi} f(u) \sin nu \, du,
\end{align*}
\]

converges to the sum \( [f(x+)+f(x-)]/2 \) at each point \( x \) in \((0,2\pi)\).

**Proof.** We presently consider the sum of only the first \( 2N+1 \) terms of (13),
\[ S_N(x) = a_0/2 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx). \quad (14) \]

After substitution into (14) for the coefficients \( a_r \) and \( b_r \),
\[ S_N(x) \text{ reduces to} \]
\[
S_N(x) = \left\{ \int_0^{2\pi} f(u) \left[ \frac{1}{2} + \sum_{n=1}^{N} \cos nu \cos nx + \sin nu \sin nx \right] du \right\} / \pi. \tag{15}\]

A trigonometric substitution for the factor in brackets in (15) results in
\[
S_N(x) = \left\{ \int_0^{2\pi} f(u) \sin \left[ (N+1/2)(u-x) \right] / (2\sin((u-x)/2)) du \right\} / \pi,
\]
\[
= \int_0^{2\pi} \psi(u) \sin \left[ (N+1/2)(u-x) \right] / (u-x) du,
\]
where
\[
\psi(u) = f(u) [(u-x)/2] / \pi \sin[(u-x)/2]. \tag{16}\]

It is given that \( f(u) \) is piecewise smooth in \((0, 2\pi)\). So, by definition, \( f(u) \) is piecewise continuous in \((0, 2\pi)\) and has right-hand and left-hand derivatives at the point \( u = x \). By its form given in (16), it follows that \( \psi(u) \) has these same properties. This allows application of Theorem 3.7 to conclude that
\[
S_N(x) = \lim_{N \to \infty} \int_0^{2\pi} \psi(u) \sin \left[ (N+1/2)(u-x) \right] / (u-x) du
\]
\[
= \left( \pi/2 \right) \left[ \psi(x^+) + \psi(x^-) \right]. \tag{17}\]

Use of (16) in (17) yields
\[
S_N(x) = \left( \pi/2 \right) \left\{ \left( f(x^+)/\pi \right) \lim_{u \to x^+} \left[ (u-x)/(2\sin((u-x)/2)) \right] \right\}
\]
\[
+ \left( f(x^-)/\pi \right) \lim_{u \to x^-} \left[ (u-x)/(2\sin((u-x)/2)) \right] \right\}. \]
By L'Hospital's rule,

\[ S_N(x) = (\pi/2)\left( (f(x+)/\pi)\lim_{u \to x+} \cos[(u-x)/2] \right)^{-1} + (f(x-)/\pi)\lim_{u \to x-} \cos[(u-x)/2]^{-1} \]

\[ = (\pi/2)\{f(x+)/\pi + f(x-)/\pi\} \]

\[ = [f(x+)+f(x-)]/2. \quad \text{Q.E.D.} \]

Hence, we have established sufficient conditions for the convergence of a Fourier series. Now let us see what this information tells us about the convergence of a Chebyshev series.

Let \( f(x) \) be some function defined in \(-1 \leq x \leq 1\). If we make the change of variable \( x = \cos \theta \), then

\[ f(x) = f(\cos \theta) = g(\theta), \]

\(-\pi \leq \theta \leq \pi\). The function \( \cos \theta \) is even and periodic with period \( 2\pi \). Hence, \( g(\theta) \) is also even and periodic with period \( 2\pi \).

The Fourier series for \( g(\theta) \) is then given by

\[ g(\theta) = a_0/2 + \sum_{r=1}^{\infty} (a_r \cos r\theta + b_r \sin r\theta), \]

where

\[ a_r = \pi^{-1}\int_{-\pi}^{\pi} g(\theta)\cos r\theta \, d\theta, \]

\[ b_r = \pi^{-1}\int_{-\pi}^{\pi} g(\theta)\sin r\theta \, d\theta. \]

Since the product \( g(\theta)\sin r\theta \) is an odd function, we have that \( b_r = 0 \) for \( r = 1, 2, \ldots \). So,

\[ g(\theta) = a_0/2 + \sum_{r=1}^{\infty} a_r \cos r\theta, \quad (18) \]

where \( a_r = (2/\pi)\int_{0}^{\pi} g(\theta)\cos r\theta \, d\theta. \)
Conversion of (18) back to the independent variable \( x \) reveals that

\[
g(\theta) = f(\cos \theta) = f(x) = a_0/2 + \sum_{r=1}^{\infty} a_r T_r(x),
\]

where

\[
a_r = (2/\pi) \int_{-1}^{1} f(x) T_r(x) (1-x^2)^{-1/2} dx,
\]

which is precisely the Chebyshev series for \( f(x) \).

In conclusion, the convergence or divergence of the Chebyshev series for a function \( f(x) \) is determined by the convergence or divergence of the Fourier series for \( g(\theta) = f(\cos \theta) \), and the conditions under which the latter series converges are given in Theorem 3.8.

Let \( f(x) \) be a function defined on the interval \([-1,1]\], and, this time, suppose that we desire the polynomial \( p(x) \) of degree \( n \) or less such that

\[
S \equiv \sum_{k=0}^{N} w(x_k)[f(x_k) - p(x_k)]^2
\]

\[
\leq \sum_{k=0}^{N} w(x_k)[f(x_k) - q(x_k)]^2,
\]

(19)

where \( q(x) \) is any polynomial of degree \( n \) or less, \( w(x_k) > 0 \) for all \( k \), and \( x_0, \ldots, x_N \) are \( N+1 \) given points. If we can produce a set of polynomials \( \{\phi_r(x)\}_{r=0}^{n} \) such that \( \phi_r(x) \) is of degree \( r \), then we can write

\[
p(x) = c_0 \phi_0(x) + c_1 \phi_1(x) + \ldots + c_n \phi_n(x).
\]

We wish to find those \( c_i \) for which

\[
S(c_0, \ldots, c_n) = \sum_{k=0}^{N} w(x_k)[f(x_k) - \sum_{r=0}^{n} c_r \phi_r(x_k)]^2
\]

is a minimum. For minimum \( S \), \( \partial S/\partial c_i = 0 \), for \( i = 0, \ldots, n \).
This yields

\[ \frac{\partial S}{\partial c_i} = 2 \sum_{k=0}^{N} w(x_k)f(x_k)\phi_i(x_k) \]

\[ - 2 \sum_{r=0}^{n} c_r \sum_{k=0}^{N} w(x_k)\phi_i(x_k)\phi_r(x_k) = 0, \quad i = 0, 1, \ldots, n. \]  

(20)

Suppose that it is true that

\[ \sum_{k=0}^{N} w(x_k)\phi_i(x_k)\phi_r(x_k) \]

\[ = 0, \text{ for } i \neq r \]

\[ \neq 0, \text{ for } i = r. \]

Since the coefficient matrix of the linear system (20) is invertible, the unique solution of this system is given by

\[ c_i = \frac{\sum_{k=0}^{N} w(x_k)f(x_k)\phi_i(x_k)}{\sum_{k=0}^{N} w(x_k)\phi_i^2(x_k)}. \]  

(21)

Hence, the problem is solved if we can find some polynomials \( \phi_r(x) \) and the points \( x_0, \ldots, x_N \) for which the assumed conditions hold. Toward this end, we assign

\[ \phi_r(x) = T_r(x), \text{ for } r = 0, \ldots, n, \]

\[ x_k = \cos(k\pi/N), \text{ for } k = 0, \ldots, N, \]

\[ w(x_0) = w(x_N) = 1/2, \]

\[ w(x_k) = 1, \text{ for } k = 1, \ldots, N-1, \]

and now test these assignments.

For \( i \neq r \), we find

\[ \sum_{k=0}^{N} T_i(x_k)T_r(x_k) = 0, \]
and for \( i = r \), it turns out that

\[
\sum_{k=0}^{N} T_i(x_k) T_r(x_k) = \begin{cases} 
\frac{N}{2} & \text{if } r \neq 0, N \\
N & \text{if } r = 0, N,
\end{cases}
\]

(where the double prime indicates that both the first and the last terms of the summation are taken with a factor of \( 1/2 \).)

So, the polynomials \( T_r(x) \) have the desired properties and we know that the polynomial

\[
p(x) = \sum_{r=0}^{n} c_r T_r(x),
\]

where

\[
c_r = \frac{2}{N} \sum_{k=0}^{N} f(x_k) T_r(x_k)
\]

and \( x_k = \cos(k\pi/N) \) satisfies

\[
\sum_{k=0}^{N} [f(x_k) - p(x_k)]^2 \leq \sum_{k=0}^{N} [f(x_k) - q(x_k)]^2
\]

for any polynomial \( q(x) \) of degree \( n \) or less.

Example 2. We will produce the polynomial \( p(x) \) of degree 4 or less which is the Chebyshev discrete least-squares approximation of degree 4 to the function \( f(x) = x + \cos^{-1}x \) using five nodal points, that is, \( N = 4 \). Using (22), we obtain

\[
p(x) = \frac{\pi}{2} + \left(1 - \frac{\pi}{(4\sqrt{2})}\right) T_1(x) + \left(\frac{\pi}{(4\sqrt{2})} - \frac{\pi}{4}\right) T_3(x),
\]

\[
= 1.5707963 + 0.3493548x - 0.9201511x^3.
\]

Notice the special case \( n = N \). We are attempting to minimize \( S \) as defined in (19). Of course, \( S = 0 \) for the unique polynomial of degree \( N \) or less which passes through
the N+1 nodal points. Hence, the minimizing polynomial \( p(x) \) given by (22) is also the interpolating polynomial for the N+1 points.

It is also possible to formulate a similar, but different, discrete least-squares approach. We have found that if we have a sequence of orthogonal polynomials, \( \phi_0(x), \phi_1(x), \ldots \), at our disposal, then least-squares approximations can be readily computed. By [2, p. 11], these orthogonal polynomials satisfy the following three-term recurrence relation:

\[
\phi_{r+1}(x) = (\alpha_r x + \beta_r) \phi_r(x) + \gamma_{r-1} \phi_{r-1}(x)
\]

(23)

with the coefficients given by

\[
\alpha_r = A_{r+1}/A_r,
\]

\[
\gamma_{r-1} = -(A_{r+1} A_{r-1} k_r)/(A_r^2 k_{r-1}),
\]

where \( A_r \) is the coefficient of \( x^r \) in \( \phi_r(x) \) and where

\[
k_r = \int_{-1}^{1} w(x) \phi_r^2(x) dx.
\]

If we choose \( k_r = 1 \) for all \( r \) and if we define \( p_r = A_r/A_{r+1} \), \( q_r = -\beta_r p_r \), \( \phi_{-1}(x) = 0 \) and choose the N+1 data points \( x_k \) so that \( \phi_{N+1}(x_k) = 0, k = 0, \ldots, N \), then from (23) above, we have

\[
A \Phi(x_k) = x_k \Phi(x_k), \quad k = 0, \ldots, N,
\]

where
\[
A = \begin{pmatrix}
q_0 & p_0 & 0 & 0 & 0 & 0 & \ldots \\
p_0 & q_1 & p_1 & 0 & 0 & 0 & \ldots \\
0 & p_1 & q_2 & p_2 & 0 & 0 & \ldots \\
0 & 0 & p_2 & q_3 & p_3 & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\
\vdots & \vdots & \vdots & \vdots & \ddots & p_{N-2} & q_{N-1} & p_{N-1} \\
0 & \cdots & \cdots & \cdots & \cdots & p_{N-2} & q_{N-1} & p_N
\end{pmatrix}
\]

and \( \phi(x_k) = (\phi_0(x_k), \phi_1(x_k), \ldots, \phi_N(x_k)) \). Hence, the numbers \( x_0, x_1, \ldots, x_N \) are the eigenvalues of the symmetric tridiagonal matrix \( A \). The eigenvector corresponding to the eigenvalue \( x_k \) is \( \phi(x_k) \). Since \( A \) is a symmetric matrix, its set of eigenvectors form an independent orthogonal system.

Define \( \lambda_k = (\sum_{r=0}^{2} \phi_r^2(x_k))^{-1}, k = 0, 1, \ldots, N. \) Let a matrix \( X \) be defined by

\[
X = \begin{pmatrix}
\lambda_0^0 \phi_0(x_0) & \lambda_1^0 \phi_0(x_1) & \cdots & \lambda_N^0 \phi_0(x_N) \\
\lambda_0^1 \phi_1(x_0) & \lambda_1^1 \phi_1(x_1) & \cdots & \lambda_N^1 \phi_1(x_N) \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_0^N \phi_N(x_0) & \lambda_1^N \phi_N(x_1) & \cdots & \lambda_N^N \phi_N(x_N)
\end{pmatrix}
\]

Notice that \( X \) is orthogonal since its columns form an independent orthogonal system. Therefore \( XX^{-1} = XX^T = I \) which gives us

\[
\sum_{k=0}^{N} \lambda_k^r \phi_r^2(x_k) = 1, \ r = 0, 1, \ldots, N, \\
\sum_{k=0}^{N} \lambda_k^r \phi_r(x_k) \phi_s(x_k) = 0, \ r \neq s.
\]
Therefore, these orthogonal polynomials \( \phi_r(x) \) provide, by (21), the following least-squares approximation \( p(x) \) for the weight function \( w(x_k) = \lambda_k \), \( k = 0, \ldots, N \):

\[
p(x) = \sum_{r=0}^{n} c_r \phi_r(x),
\]

\[
c_r = \sum_{k=0}^{N} \lambda_k f(x_k) \phi_r(x_k).
\]  

(24)

For the Chebyshev polynomials \( T_r(x) \), we find, for \( x_0, \ldots, x_N \), the \( N+1 \) zeros of \( T_{N+1}(x) \),

\[
\lambda_k = \left[ \frac{1}{N} \sum_{r=0}^{n} T_r^2(x_k) \right]^{-1} = 2/(N+1), \quad k = 0, \ldots, N.
\]

By (24), the Chebyshev approximation,

\[
p(x) = \sum_{r=0}^{n} c_r T_r(x),
\]

\[
c_r = \left[ \frac{2}{(N+1)} \right] \sum_{k=0}^{N} f(x_k) T_r(x_k),
\]  

(25)

where

\[
x_k = \cos\left(\frac{(2k+1)/(N+1) \cdot (\pi/2)}{N}\right),
\]

satisfies

\[
N \sum_{k=0}^{N} [f(x_k) - p(x_k)]^2 \leq \sum_{k=0}^{N} [f(x_k) - q(x_k)]^2
\]

for any polynomial \( q(x) \) of degree \( n \) or less.

Observe that for \( n = N \) the polynomial \( p(x) \) defined in (25) must be the interpolation polynomial \( p^*(x) \) of degree \( N \) which agrees with \( f(x) \) at the \( N+1 \) nodes \( x_0, \ldots, x_N \). This is true since

\[
N \sum_{k=0}^{N} [f(x_k) - p^*(x_k)] = 0.
\]
Example 3. Corresponding to the above treatment, let us now compute a discrete least-squares approximation of degree 4 to our standby function \( f(x) = x + \cos^{-1}x \). Specifically, we will use (25) to find \( p(x) \) such that
\[
\sum_{k=0}^{4} [f(x_k) - p(x_k)]^2 \leq \sum_{k=0}^{4} [f(x_k) - q(x_k)]^2
\]
where \( q(x) \) is any polynomial of degree 4 or less and where
\[
x_k = \cos[(2k+1)\pi/10], \quad k = 0, \ldots, 4.
\]
Let \( p(x) = \sum_{r=0}^{4} c_r T_r(x) \). It turns out that \( c_0 = \pi \), \( c_1 = -0.2515581 \), \( c_2 = 0 \), \( c_3 = -0.1128528 \), and \( c_4 = 0 \).
Upon conversion to powers of \( x \) we find the result
\[
p(x) = 1.5707963 + 0.0870003x - 0.4514112x^3.
\]
As the first method of obtaining a polynomial approximation to a rational function, we seek to compute the Chebyshev series for that function. Let \( f(x) = P(x)/Q(x) \) where \( P(x) \) and \( Q(x) \) are polynomials. We try to determine the coefficients \( a_r \) in the equation
\[
P(x)/Q(x) - \sum_{r=0}^{\infty} a_r T_r(x) = 0.
\]
Multiplying by \( Q(x) \) yields
\[
P(x) = Q(x) \sum_{r=0}^{\infty} a_r T_r(x). \quad (26)
\]
If application of the identities (1.19) are made to the right-hand side of (26) and also if \( P(x) \) is expressed in terms of \( T_r(x) \), then it is possible to equate the coefficients of the \( T_r(x) \) and to come up with a set of solvable difference
equations, the unknown variables of which are the coefficients $a_r$. Hence $\sum_{r=0}^{\infty} a_r T_r(x)$ can be produced by this procedure which is more clear if one studies the following example.

Example 4. We proceed to find the Chebyshev series, in $0 \leq x \leq 1$, for the rational function $f(x) = x/(1+x)$. Set

$$x/(1+x) - \sum_{r=0}^{\infty} a_r T_r(x) = 0.$$ 

Multiply by $1+x$ to get

$$x = (1+x) \sum_{r=0}^{\infty} a_r T_r(x)$$

$$= \sum_{r=0}^{\infty} a_r T_r^*(x) + \sum_{r=0}^{\infty} a_r x T_r^*(x). \quad (27)$$

Since

$$x T^*_r(x) = (1/4)[T^*_{r-1}(x) + 2T^*_r(x) + T^*_{r+1}(x)]$$

and since

$$x = [T^*_0(x) + T^*_1(x)]/2$$

we have from (27)

$$[T^*_0(x) + T^*_1(x)]/2 = [(3a_0 + a_1)/4]T^*_0(x)$$

$$+ \sum_{r=1}^{\infty} [(a_{r-1} + 2a_r + a_{r+1})/4]T^*_r(x).$$

Equating the coefficients of the $T^*_r(x)$ on both sides of this equation produces the system

$$3a_0/4 + a_1/4 = 1/2, \quad (28)$$

$$a_0/4 + 3a_1/2 + a_2/4 = 1/2, \quad (29)$$

$$(a_{r-1} - 6a_r + a_{r+1})/4 = 0, \text{ for } r = 2, 3, \ldots. \quad (30)$$
Equation (30) is a difference equation with constant coefficients and has the general solution

\[ a_r = A(-3+2\sqrt{2})^r + B(-3-2\sqrt{2})^r, \quad r = 1, 2, \ldots \]

In order for the Chebyshev series for \( f(x) \) to be convergent, it is necessary that \( B = 0 \). Thus,

\[ a_r = A(-3+2\sqrt{2})^r, \quad r = 1, 2, \ldots \quad (31) \]

Using (31) to substitute for \( a_r \) back into (28) and (29), we solve to get \( A = -\sqrt{2} \) and, finally, the Chebyshev series

\[ \frac{x}{1+x} = (1-1/\sqrt{2}) - \sqrt{2} \sum_{r=1}^{\infty} (-3+2\sqrt{2})^r T_r^*(x). \quad (32) \]

Note that

\[ \sum_{r=1}^{\infty} (-3+2\sqrt{2})^r T_r^*(x) \leq \sum_{r=1}^{\infty} (-3+2\sqrt{2})^r |T_r^*(x)| \leq \sum_{r=1}^{\infty} |3+2\sqrt{2}|^r. \]

The sum \( \sum_{r=1}^{\infty} |3+2\sqrt{2}|^r \) is a geometric series, and hence the Chebyshev series (32) is seen to converge rapidly. Note that (32) was produced without the burdensome calculation by integrals of the coefficients \( a_r \).

A second method of finding an approximation for the rational function \( f(x) = P(x)/Q(x) \) involves finding the coefficients \( a_r \) such that

\[ P(x) = Q(x) \sum_{r=0}^{n} a_r T_r(x), \quad (33) \]

where \( n \) is an integer. It may well be impossible to find such coefficients so that (33) is satisfied exactly.
However, we may be able to exactly satisfy the perturbed variation of (33),
\[
P(x) + E(x) = Q(x) \sum_{r=0}^{n} a_r T_r(x),
\]
where \(E(x) = \sum_{m=p}^{q} \tau_m T_m(x)\) is a perturbing polynomial and where the integers \(p\) and \(q\) depend on the relative degrees of both sides of (33). If the constants \(\tau_m\) and the coefficients \(a_r\) are determined in some manner, then since from (34) we have
\[
P(x)/Q(x) + E(x)/Q(x) = \sum_{r=0}^{n} a_r T_r(x),
\]
the summation \(\sum_{r=0}^{n} a_r T_r(x)\) is taken as the desired approximation with the error \(\epsilon(x)\) given by \(\epsilon(x) = E(x)/Q(x)\).

Taking a closer look at this error we see
\[
|\epsilon(x)| \leq \max_{-1 \leq x \leq 1} \frac{|E(x)|}{\min_{-1 \leq x \leq 1} |Q(x)|} \leq \sum_{m=p}^{q} |\tau_m| / \min_{-1 \leq x \leq 1} |Q(x)|. \tag{35}
\]

**Example 5.** Getting more specific, we now compute a Chebyshev approximation of degree 4 to the rational function \(f(x) = x/(1+x)\) over the interval \([0,1]\). We are seeking \(a_0, a_1, a_2, a_3,\) and \(a_4\) such that
\[
x \approx (1+x) \sum_{r=0}^{4} a_r T_r^*(x). \tag{36}
\]
As in the first method of approximating rational functions, we equate the coefficients of the \(T_r^*(x)\) on opposite sides of equation (36) to obtain the system
\[3a_0/4 + a_1/4 = 1/2,\]
\[a_0/4 + 3a_1/2 + a_2/4 = 1/2\]
\[a_{r-1}/4 + 3a_r/2 + a_{r+1}/4 = 0 \text{ for } r = 2, 3,\]
\[a_3/4 + 3a_4/2 = 0,\]
\[a_4/4 = 0.\]

The sole solution \((0,0,0,0,0)\) of this system is a useless one. If, instead, we seek the \(a_r\) such that
\[x + \tau_5 T_5^*(x) = (1+x) \sum_{r=0}^{4} a_r T_r^*(x),\]
then the resulting system of equations is unchanged from (18) above with the exception of the last equation which becomes \(a_4/4 = \tau_5\). With this change, the system (37) becomes six equations in six variables. The solution of this new system yields \(\tau_5 = -1/3363\) and the approximation
\[f(x) = x/(1+x) \sim 0.2928933 + 0.2426404T_1^*(x) - 0.0415241T_2^*(x) - 0.0011894T_4^*(x).\]

This result compares favorably with the first five terms of the Chebyshev series for \(f(x)\). From (32),
\[x/(1+x) = 0.2928932 + 0.2426408T_1^*(x) - 0.0416302T_2^*(x) + 0.0071364T_3^*(x) - 0.0012261T_4^*(x) + \sqrt{2} \sum_{r=5}^{\infty} (-3+2\sqrt{2})^r T_r^*(x).\]
Using (35), an upper bound for the error $\varepsilon(x)$ of the approximation is given by

$$|\varepsilon(x)| \leq \frac{|-1/3633|}{\min_{0 \leq x \leq 1} |1+x|}$$

$$= \frac{1}{3363} = 0.0002973.$$
REFERENCES


CHAPTER IV

ORDINARY DIFFERENTIAL EQUATIONS
AND A QUADRATURE FORMULA

In this chapter, we present two methods which employ Chebyshev polynomials to solve certain types of first order differential systems, and then we develop a Chebyshev quadrature formula. The methods for the solution of differential systems bear a strong resemblance to the methods in Chapter III for obtaining Chebyshev approximations to functions.

Consider the first-order differential equation

\[ p_1(x)y'(x) + p_2(x)y(x) = p_3(x), \tag{1} \]

where \( p_1(x) \), \( p_2(x) \), and \( p_3(x) \) are polynomials and where an associated starting condition such as \( y(0) = 1 \) is given.

In search for a solution of the form

\[ y(x) = \sum_{r=0}^{\infty} a_r T_r(x) \tag{2} \]

in \(-1 \leq x \leq 1\), we begin by integrating (1) and substituting for \( y(x) \) from (2) to get

\[ p_1(x) \sum_{r=0}^{\infty} a_r T_r(x) + \int (p_2(x) - p_1(x)) \sum_{r=0}^{\infty} a_r T_r(x) = \int p_3(x)dx + C. \tag{3} \]
Using formulas (2.19) and (2.36), the left-hand side of (3) can be converted into an infinite Chebyshev series in which the coefficient of $T_r(x)$, $r = 0,1,\ldots$, is a finite linear combination of coefficients $a_r$. The right-hand side of (3) can easily be written as a finite Chebyshev series so that the transformed (3) takes on the form

$$
\sum_{r=0}^{\infty} b_r T_r(x) = \sum_{r=p}^{q} c_r T_r(x),
$$

where the $b_r, c_r, p$ and $q$ are constants. Equating the coefficients of corresponding $T_r(x)$ yields an infinite set of linear equations involving the unknowns $a_0, a_1, a_2, a_3, \ldots$. An additional linear equation of this type can be produced from the given starting condition. The resulting linear system can be used to compute successively improving approximations for the $a_r$ until we are satisfied with their accuracy. This technique we call the infinite series method to distinguish it from the $\tau$-method to be introduced later in this chapter.

Second-order differential equations can be solved in a similar manner but with increased difficulty. To solve

$$q_1(x)y''(x) + q_2(x)y'(x) + q_3(x) = q_4(x),$$

where $q_r(x)$, $r = 1,\ldots, 4$ are polynomials and where two starting conditions are given, we proceed as before except that this time two integrations of the left-hand side of (5)
are required to produce the desired infinite Chebyshev series there. The increased difficulty comes mostly in the solving of a slightly more complex linear system.

Let us apply the infinite series method to a particular example problem.

**Example 6.** Take a look at the details of the infinite series method by applying it to the first-order differential system,

\[(1+x)y'(x) + (1+x+x^2)y(x) = 1-x, \quad \text{(6)}\]
\[y(0) - (3/4)y(1) = 1, \text{ where } 0 \leq x \leq 1. \quad \text{(7)}\]

From (6) we have that

\[(1+x)dy + (1+x+x^2)y(x)dx = (1-x)dx.\]

Integration of both sides of this equation brings us to

\[(1+x)y(x) + \int(xy(x)+x^2y(x))dx = x - x^2/2 + C. \quad \text{(8)}\]

Setting \(y(x) = \sum a_r T^*(x)\) in (8) yields

\[\sum a_r T^*(x) + \sum a_r xT^*(x) + \int(\sum a_r xT^*(x))dx + \int(\sum a_r x^2T^*(x))dx = x - x^2/2 + C. \quad \text{(9)}\]

If we now use a formula derived in Chapter II,

\[x^r T^*_s(x) = 2^{-2r} \sum_{i=0}^{2r} T^*_{s-r+i}(x),\]

and if we also put the right-hand side of (9) into terms of \(T^*_1(x)\) and \(T^*_2(x)\), then after a considerable amount of
rearranging we get
\[29a_0/64 + 13a_1/8 + a_2/16 - a_3/8 - a_4/64)T_1^*(x)\]
\[+ (a_0/16 + 23a_1/64 + 3a_2/2 + 19a_3/128 - a_4/16 - a_5/128)T_2^*(x)\]
\[+ \sum_{r=3}^{\infty}[a_{r-1}/4 + 3a_r/2 + a_{r+1}/4 + r^{-1}(a_{r-5}/64 + a_{r-2}/8\]
\[+ 13a_{r-1}/64 - 13a_{r+1}/64 - a_{r+2}/8 - a_{r+3}/64]T_r^*(x)\]
\[= T_1^*(x)/4 - T_2^*(x)/16 + C_1. \quad (10)\]

According to the initial condition (7),
\[\sum_{r=0}^{\infty}a_rT_r^*(0) - (3/4)\sum_{r=0}^{\infty}a_rT_r^*(1) = 1.\]

This implies, in view of the point evaluations (2.47),
that \(a_0/8 - 7a_1/4 + a_2/4 - 7a_3/4 + a_4/4 - \ldots = 1.\) This linear equation together with the infinite set of linear equations which result from equating the coefficients of corresponding \(T_r^*(x)\) on each side of equation (10) form an infinite linear system. From this system, we can produce the approximations
\[a_0 \approx 2.11641, \quad a_1 \approx -0.43448, \quad a_2 \approx -0.02679, \quad a_3 \approx 0.01039\]
which yield
\[y(x) \approx 1.05821 - 0.43448T_1^*(x) - 0.02679T_2^*(x) + 0.01039T_3^*(x),\]
an approximate solution of the original differential system.

Consider again the differential equation (1). The \(r\)-method involves finding a finite series approximation to the solution \(y(x),\)
\[y(x) = \sum_{r=0}^{n}a_rT_r(x),\]
where \( n \) is some integer. If (1) is again integrated and \( y(x) \) is replaced by \( \sum_{r=0}^{n} a_r T_r(x) \), then the result is

\[
p_1(x) \sum_{r=0}^{n} a_r T_r(x) + \int_{0}^{\omega} \left( \int (p_2(x) - p_1(x)) \sum a_r T_r(x) \right) = \int p_3(x) dx + C_2. \tag{11}
\]

It is possible that there may not exist coefficients \( a_r, r = 0, \ldots, n \), such that (11) is satisfied. However, the addition of some perturbing terms on the right-hand side of (11) produces the variant

\[
p_1(x)y(x) + \int (p_2(x) - p_1(x))y(x) dx = \int p_3(x) dx + \sum_{r=a}^{b} \tau_r T_r(x), \tag{12}
\]

where \( a = \text{deg}(p_3(x)) + 2 \), \( b = n + \max[\text{deg}(p_1(x)), \text{deg}(p_2(x)) + 1] \), and \( \tau_r, r = a, \ldots, b \), are constants to be determined. The exact solution of (12) is now a polynomial, say \( Q(x) \), so that coefficients exist such that

\[
Q(x) = \sum_{r=0}^{n} a_r T_r(x).
\]

To finish up, proceed just as before in the infinite series method by equating coefficients of the corresponding \( T_r(x) \) and by using the given initial condition to produce a linear system. Of course, this time the linear system is finite, and in addition to the unknowns \( a_r \) it contains the unknown perturbing term coefficients \( \tau_r \).
The treatment of a second-order differential system follows the same lines. Key differences are that two initial conditions are supplied, and two integrations are required for the necessary transformation to a finite Chebyshev series.

Let us now study the error which is involved in taking the finite series

\[ \bar{y}(x) = \sum_{r=0}^{n} a_r T_r(x) \]

to represent the exact solution \( y(x) \) of (1). Since \( \bar{y}(x) \) satisfies (12), we have

\[ p_1(x)\bar{y}(x) + \int (p_2(x) - p_1'(x))y(x)dx = p_3(x)dx + \sum_{r=a}^{b} \tau_r T_r(x). \]  

(13)

By integrating (1), we get a similar equation which \( y \) satisfies,

\[ p_1(x)y(x) + \int (p_2(x) - p_1'(x))y(x)dx = \int p_3(x)dx + C. \]  

(14)

Subtracting (13) from (14) yields

\[ p_1(x)(y(x) - \bar{y}(x)) + \int (p_2(x) - p_1'(x))(y(x) - \bar{y}(x))dx = C - \sum_{r=a}^{b} \tau_r T_r(x), \]  

(15)

an expression which directly involves the error

\[ e_n(x) = y(x) - \bar{y}(x) \]

and which may lend itself to providing an approximation for this error as is the case in the following example.
Example 7. The differential system of (6) and (7) serves as a good sample case. By the $\tau$-method, we can obtain

$$y(x) = 1.04212T_0^*(x) - 0.42632T_1^*(x) - 0.02637T_2^*(x)$$

as the exact solution of a perturbed form of (8). This perturbed equation is given by

$$(1+x)y(x) + \int (x+x^2)y(x)dx = x - x^2/2 + E(x) + A,$$  \hspace{1cm} (16)

where

$$E(x) = \tau_3 T_3^*(x) + \tau_4 T_4^*(x) + \tau_5 T_5^*(x),$$

and where the constant $A$ is chosen so that

$$\overline{y}(0) - (3/4)\overline{y}(1) = 1.$$

Subtraction of (8) from (16) shows that

$$(1+x)e_2(x) + \int (x+x^2)e_2(x)dx = E(x) + \alpha,$$  \hspace{1cm} (17)

where $e_2(x) = \overline{y}(x) - y(x)$ and $\alpha$ is a constant such that

$$e_2(0) - (3/4)e_2(1) = 0.$$  \hspace{1cm} (18)

By (17), we see that

$$(1+x)e_2(x) = E(x) + \alpha - \int (x+x^2)e_2(x)dx.$$

The solution of equations like (17) and (18) is not particularly easy, but the nature of $E(x)$ often enables us to find a good estimate for $e_2(x)$ from the iterative scheme defined in this case by

$$(1+x)e_2^{(r+1)}(x) = E(x) + \alpha_r - \int (x+x^2)e_2^{(r)}(x)dx,$$  \hspace{1cm} (19)
with $a_r$ adjusted so that (18) is satisfied at every step and with $e_2^{(0)}(x) = 0$. The integral in (19) can be expressed in closed analytical form in this example, but it is just as easy, and sometimes necessary, to use a simple quadrature rule such as the trapezoid rule.

Example 8. Consider the differential equation

$$\begin{align*}
(1+x)y'(x) + (1+x+x^2)y(x) &= x^4 + x^3 + 3x^2 + 2x, \\
\end{align*}$$

and the associated condition $y(1) = 1$. It is easily verified that $y(x) = x^2$ is the exact solution of this differential system. Since in the $\tau$-method we try to approximate the solution of a system by assuming it to take the form of a polynomial, then using the $\tau$-method on (20) to compute a cubic approximation in $0 \leq x \leq 1$ should result in $\overline{y}(x) = x^2$. We execute to find out if this is so.

Integrating (20) yields

$$\begin{align*}
(1+x)y(x) + \int xy(x)\,dx + \int x^2y(x)\,dx &= x^5/5 + x^4/4 + x^3 + x^2 + C. \\
\end{align*}$$

(21)

When we let $y(x) = \sum_{r=0}^{3} a_r T_r^*(x)$, notice that the left-hand side of (21) becomes a polynomial of degree 6 while the right-hand side is a polynomial of degree 5. So, we add the perturbing term $\tau_6 T^*_6(x)$ to the right-hand side so that the equation which we are trying to satisfy will then become
\[
\sum_{r=0}^{3} a_r T_r^*(x) + \sum_{r=0}^{3} a_r T_r^*(x)
\]
\[
+ \int (\sum_{r=0}^{3} a_r x T_r^*(x)) \, dx + \int (\sum_{r=0}^{3} a_r T_r^*(x)) \, dx
\]
\[
= (2970T_1^*(x) + 1060T_2^*(x) + 165T_3^*(x)
\]
\[
+ 15T_4^*(x) + T_5^*(x))/2560 + \tau_6 T_6^*(x) + K,
\]

where the powers of \(x\) on the right-hand side have been converted to their Chebyshev equivalents. Now, taking advantage of the fact that the left-hand side of (21) is the same as that of (8) in Example 6, we produce the following linear system by equating coefficients of corresponding \(T_r(x)\), \(r = 1, \ldots, 6\) and by using the condition \(y(1) = 1\):

\[
\begin{align*}
\frac{a_0}{2} + a_1 + a_2 + a_3 &= 1 \\
29a_0/64 + 13a_1/8 + a_2/16 - a_3/8 &= 2970/2560 \\
a_0/16 + 23a_1/64 + 3a_2/2 + 19a_3/128 &= 1060/2560 \\
a_0/192 + a_1/24 + 61a_2/192 + 3a_3/2 &= 165/2560 \\
a_2/256 + a_2/32 + 77a_3/256 &= 15/2560 \\
a_2/320 + a_3/40 &= 1/2560 \\
a_3/384 &= \tau_6.
\end{align*}
\]

This linear system has the unique solution, \(a_0 = 3/8\), \(a_1 = 1/2, a_2 = 1/8, a_3 = 0, \) and \(\tau_6 = 0\). Hence, the finite "approximation" \(\overline{y}(x)\) is given by

\[
\overline{y}(x) = 3/4 + T_1^*(x)/2 + T_2^*(x)/8 = x^2,
\]

as expected.
Chebyshev polynomials can also be employed in the solution of more general linear differential equations, non-linear differential equations, and partial differential equations. For a survey of these methods see (2, pp. 145-175).

Let us now pursue the study of another useful role played by the Chebyshev polynomials in numerical mathematics.

Since the Chebyshev polynomials are orthogonal functions with respect to a certain weight function, we can choose them to be used in conjunction with so-called Gaussian quadrature formula to produce what we call a Chebyshev-Gauss rule. We lead up to this production through a series of theorems which will define the nature of a Gaussian quadrature formula.

First, explanation of a few helpful terms is needed. The term "quadrature formula" refers to a sum of the type
\[ \sum_{j=1}^{n} \alpha_j f(x_j) \]  
which is employed as an approximation to the integral
\[ \int_{a}^{b} w(x)f(x)dx, \]  
where \( f(x) \) is a function defined on \([a,b]\), and \( w(x) \) is a positive function defined on the same interval.

The term "nodes" used in this context refers to the \( n \) points \( x_1, x_2, \ldots, x_n \). A quadrature formula \( I_{n+1}(f) \) is said to be interpolatory if
\[ I_{n+1}(f) = \int_{a}^{b} P_n(x)w(x)dx, \]  
where \( P_n(x) \) is a polynomial of degree at most \( n \) and where
\[ f(x_j) = P_n(x_j), \]  
for \( j = 0, \ldots, n \).
Lastly, if we let
\[ E_n(f) \equiv \int_a^b w(x)f(x)dx - \sum_{j=1}^{n} \alpha_j f(x_j) \]
denote the error of a quadrature formula, then we define the "degree of precision" of this quadrature formula as the largest \( m \) such that \( E_n(x^k) = 0 \), for \( k = 0, \ldots, m \), but \( E_n(x^{m+1}) \neq 0 \).

The definitions of the above terms make possible the concise statement of the next theorem which we state without proof. For a proof, see (4, p. 316).

**Theorem 4.1.** A quadrature formula which uses \( n \) distinct nodes is an interpolatory formula if and only if it has degree of precision at least \( n-1 \).

**Theorem 4.2.** The quadrature formula (23) has degree of precision at most \( 2n-1 \). This maximum degree of precision is attained if and only if the \( n \) nodes, \( x_j \), are the zeros of \( Q(x) \), the \( n \)th orthogonal polynomial with respect to the weight \( w(x) \) over \([a,b]\).

**Proof.** Suppose that the quadrature formula
\[ I(f) = \sum_{j=1}^{n} \alpha_j f(x_j) \]
has degree of precision \( 2n-1 \). Let \( g(x) \) be any polynomial of degree \( 2n-1 \). By Theorem 4.1, \( I(f) \) is an interpolating formula. Since \( g(x) = P(x) + E(g) \) where \( P(x) \) is a polynomial of degree at most \( n-1 \) such that \( g(x_j) = P(x_j), j = 1, \ldots, n \) and where
\[ E(g) = Q(x)g[x_1, \ldots, x_n, x] \]

in which

\[ Q(x) \equiv \prod_{j=1}^{n} (x-x_j), \]

then \( E(g) = 0 \) which implies that

\[ \int_a^b w(x)E(g) = \int_a^b w(x)Q(x)g[x_1, \ldots, x_n, x] = 0. \] (24)

The nth divided difference of any polynomial of degree \( n+r \) is a polynomial of degree at most \( r \). So, \( g[x_1, \ldots, x_n, x] \) is a polynomial of degree at most \( n-1 \). Therefore, by (24), the polynomial \( Q(x) \) is the nth orthogonal polynomial with respect to the weight \( w(x) \) over \([a, b]\).

For the necessity, suppose that the \( n \) nodes, \( x_j \), are the zeros of \( Q(x) \), the nth orthogonal polynomial with respect to the weight \( w(x) \) over \([a, b]\). Then a quadrature formula of the type \( \sum_{j=1}^{n} \alpha_j f(x_j) \) can have maximum degree of precision of at least \( n-1 \) since this is the case for the Lagrangian interpolation formula. The error involved when such a quadrature formula is used to approximate

\[ \int_a^b w(x)f(x)dx \]

is given by

\[ E(f) = \int_a^b w(x)Q(x)f[x_1, \ldots, x_n, x]dx. \] (25)

We know if \( f(x) \) is a polynomial of degree \( n+s \), then

\( f[x_1, \ldots, x_n, x] \) is a polynomial of degree at most \( s \). Hence, for such an \( f(x) \), if \( s = 1, \ldots, n-1 \), then (25) vanishes. Furthermore, if \( s = n \), then

\[ \int_a^b w(x)Q(x)f[x_1, \ldots, x_n, x] \neq 0. \] Q.E.D.
Since the quadrature formula of maximum degree of precision given by $I(f)$ is interpolatory, then

$$I(f) = \int_a^b w(x) P_{n-1}(x) \, dx,$$

where

$$P_{n-1}(x) = \sum_{j=1}^n \ell_j(x) f(x_j)$$

in which

$$\ell_j(x) = \left[ \prod_{k=1}^n (x-x_k) \right] / (x-x_j) \frac{d}{dx} \left[ \prod_{k=1}^n (x-x_k) \right]_{x=x_j}.$$

So,

$$I(f) = \sum_{j=1}^n f(x_j) \int_a^b w(x) \ell_j(x) \, dx.$$

In other words, the $\alpha_j$ of (23) are given by

$$\alpha_j = \int_a^b w(x) \ell_j(x) \, dx = \left\{ \frac{d}{dx} \left[ \prod_{k=1}^n (x-x_k) \right]_{x=x_j} \right\}^{-1} \cdot \left( \int_a^b \left[ w(x) \prod_{k=1}^n (x-x_k) \right] / (x-x_j) \, dx \right)$$

for $j = 1, \ldots, n$. 

A simpler expression than (26) can be found for the $\alpha_j$. Toward this end, let $Q_n(x)$ denote the $n$th orthonormal polynomial over $[a,b]$ with respect to the given weight function $w(x)$. If the leading coefficient of $Q_n(x)$ is $a_n$, then since $I(f)$ the nodes $x_1, \ldots, x_n$ are the zeros of $Q_n(x)$, we have

$$Q_n(x) = \alpha_n \prod_{j=1}^n (x-x_j)$$

and

$$Q_n'(x) = \alpha_n \frac{d}{dx} \left[ \prod_{j=1}^n (x-x_j) \right]$$
so that (26) may be written as
\[
\alpha_j = [Q'_n(x_j)]^{-1} \int_a^b Q_n(x)w(x)/(x-x_j)\,dx.
\]  \hspace{1cm} (27)

According to the Christoffel-Darboux relation (see (4, p. 205)),
\[
\frac{a_n}{a_{n+1}} [Q_{n+1}(x)Q_n(x_j) - Q_n(x_j)Q_{n+1}(x)]
\]
\[
= (x-x_j) \sum_{k=0}^n Q_k(x)Q_k(x_j).
\]  \hspace{1cm} (28)

Multiplying both sides of (28) by \(w(x)/(x-x_j)\) yields
\[
\left[\frac{a_n w(x)}{a_{n+1}(x-x_j)}\right][Q_{n+1}(x)Q_n(x_j) - Q_n(x_j)Q_{n+1}(x)]
\]
\[
= \sum_{k=0}^n w(x)Q_k(x)Q_k(x_j).
\]  \hspace{1cm} (29)

Since \(Q_n(x_k) = 0\) for \(k = 1, \ldots, n\) and
\[
Q_0(x) = \left[\int_a^b w(x)\,dx\right]^{-\frac{1}{2}},
\]
equation (29) reduces to
\[
\frac{a_n w(x)}{a_{n+1}(x-x_j)} [-Q_{n+1}(x_j)Q_n(x)]
\]
\[
= \sum_{k=0}^{n-1} w(x)Q_k(x)Q_k(x_j).
\]  \hspace{1cm} (30)

Integration of both sides of (30) over \([a,b]\) brings us to
\[
(-\frac{a_n}{a_{n+1}})Q_{n+1}(x_j) \int_a^b Q_n(x)w(x)/(x-x_j)\,dx
\]
\[
= Q_0^2(x) \int_a^b w(x)\,dx = 1.
\]  \hspace{1cm} (31)

From (31), it follows that
\[ \alpha_j = \left[ Q_n'(x_j) \right]^{-1} \int_a^b Q_n(x)w(x)/(x-x_j) \, dx \]

\[ = -a_{n+1} \frac{[a_n Q_n'(x_j) Q_{n+1}(x_j)]^{-1}}{a_n} , \quad (32) \]

\( j = 1, \ldots, n \). This last expression for \( \alpha_j \) is generally much easier to evaluate than the one in (27).

Now that it seems simple enough to compute the Gaussian approximation \( I(f) = \sum_{j=1}^n \alpha_j f(x_j) \) which has degree of precision \( 2n-1 \), let us delve into an analysis of the error

\[ E(f) = \int_a^b f(x)w(x) \, dx - I(f) . \]

**Theorem 4.3.** Let \( f'(x) \) be continuous in the closed interval \([a,b]\). Let \( \xi_1, \xi_2, \ldots, \xi_n \) be any \( n \) distinct points in \([a,b]\) which do not coincide with the zeros, \( x_1, x_2, \ldots, x_n \), of the \( n \)th orthogonal polynomial, \( Q_n(x) \), over \([a,b]\). Then the error in \( n \) point Gaussian quadrature applied to \( \int_a^b f(x)w(x) \, dx \) is

\[ E(f) = \int_a^b w(x)Q_n(x)(x-\xi_1)(x-\xi_2) \ldots (x-\xi_n) \]

\[ \cdot (f[x_1, \ldots, x_n, \xi_1, \ldots, \xi_n, x]) \, dx . \quad (33) \]

**Proof.** It is true that \( f(x) = P(x) + R(x) \) where \( P(x) \) is the polynomial of degree \( 2n - 1 \) or less which interpolates \( f(x) \) at the \( 2n \) points, \( x_1, \ldots, x_n, \xi_1, \ldots, \xi_n \), and where \( R(x) \) is the interpolation error given by

\[ R(x) = \prod_{j=1}^n \frac{[(x-x_j)(x-\xi_j)]f[x_1, \ldots, x_n, \xi_1, \ldots, \xi_n, x]}{[x-x_j](x-\xi_j)} . \quad (34) \]
Since
\[ \int_{a}^{b} f(x)w(x)dx = \sum_{j=1}^{n} \alpha_j f(x_j) + E(f), \]
then
\[ \int_{a}^{b} w(x)P(x)dx + \int_{a}^{b} w(x)R(x)dx = \sum_{j=1}^{n} \alpha_j P(x_j) + \sum_{j=1}^{n} \alpha_j R(x_j) + E(f). \] (35)

We know that
\[ \int_{a}^{b} w(x)P(x)dx = \sum_{j=1}^{n} \alpha_j P(x_j) \] (36)
since the Gaussian quadrature formula has degree of precision 2n-1 and since the coefficients \( \alpha_j \) are independent of the function whose integral is being approximated. By (36) then, (35) becomes
\[ \int_{a}^{b} w(x)R(x)dx = \sum_{j=1}^{n} \alpha_j R(x_j) + E(f). \] (37)

In view of the expression for \( R(x) \) in (35), we see that \( R(x_i) = 0 \) for \( i = 1, \ldots, n \). This is so because \( f'(x) \) being continuous forces \( f[x_1, \ldots, x_n, \xi_1, \ldots, \xi_n, x_i] \) to be finite for \( i = 1, \ldots, n \). Hence, equation (37) yields
\[ E(f) = \int_{a}^{b} w(x)R(x)dx. \] Q.E.D.

The following corollary packages up the error in an even tidier form.

**Corollary 1.** Let \( f(x) \) have a continuous derivative of order 2n in \([a,b]\). Then the error \( E(f) \) in Theorem 4.3 is given by
\[ E\{f\} = \int_a^b w(x)Q_n^2(x)dx, \quad (38) \]

where \( \xi \) is some point in \([a,b]\).

**Proof.** Under the assumed continuity conditions on \( f(x) \), the integrand in (33) is a continuous function of the \( n \) points \( \xi_1, \ldots, \xi_n \). Thus, we can let \( \xi_j \to x_j \) for \( j = 1, \ldots, n \) so that

\[ E\{f\} = \int_a^b w(x)Q_n^2(x)f[x_1, \ldots, x_n, x_1, \ldots, x_n, x]dx. \]

Since \( w(x)Q_n^2(x) \) is a nonnegative integrable function on \([a,b]\) and since \( f[x_1, \ldots, x_n, x_1, \ldots, x_n, x] \) is continuous on \([a,b]\), then by the Mean-value Theorem for integrals we have

\[ E\{f\} = f[x_1, \ldots, x_n, x_1, \ldots, x_n, \eta]\int_a^b w(x)Q_n^2(x)dx \]

where \( a < \eta < b \). Thus, since \( f'(2n)(x) \) is continuous and since \( x_1, \ldots, x_n, \eta \) all are in \([a,b]\), then

\[ E_n\{f\} = [f'(2n)(\xi)/(2n)!] \int_a^b w(x)Q_n^2(x)dx. \quad Q.E.D. \]

Taking the Chebyshev polynomials to be the set of orthogonal polynomials involved in the just discussed general case of Gaussian quadrature results in what we refer to as Chebyshev-Gauss quadrature. This quadrature rule is therefore useful in approximating integrals of the form

\[ \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x)dx \]

where \( f(x) \) is a function defined on the interval \([-1,1]\).

For \( n \) point Chebyshev-Gauss quadrature, this approximation
is given by
\[ \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x) \, dx = \frac{\alpha_j f(x_j)}{\sum_{j=1}^{n} \alpha_j f(x_j)}, \]
where according to (32) the coefficients \( \alpha_j \) are given by
\[ \alpha_j = -2^n \left[ 2^{n-1} \cdot (2/\pi) T'_n(x_j) T_{n+1}(x_j) \right]^{-1} \]
\[ = -\pi \left[ T'_n(x_j) T_{n+1}(x_j) \right]^{-1} \text{ for } j = 1, \ldots, n, \quad (39) \]
and where \( x_j = \cos \left[ (2j-1)\pi/2n \right], j = 1, \ldots, n. \) By direct calculation, one can find that
\[ T'_n(x_j) = (-1)^{j+1} n/\sin \beta_j \text{ and } \]
\[ T_{n+1}(x_j) = (-1)^j \sin \beta_j \text{ where } \beta_j = (2j-1)\pi/2n. \quad (40) \]
Hence, equation (39) reduces to \( \alpha_j = \pi/n \) for \( j = 1, \ldots, n, \) and the Chebyshev-Gauss approximation is given by
\[ \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x) \, dx \approx \frac{\pi}{n} \sum_{j=1}^{n} f[\cos((2j-1)\pi/2n)]. \quad (41) \]
According to (38) the error caused by the approximation is given by
\[ E[f] = f(2n)(\xi) \left[ (2^{n-1})^2 (2n)! \right]^{-1} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_n^2(x) \, dx, \]
where \( \xi \) is some point in the interval \([-1,1]. \) In the light of equations (40), this reduces to
\[ E[f] = 2\pi f(2n)(\xi) \left[ 2^{2n} (2n)! \right]^{-1}. \quad (42) \]
\textbf{Example 9.} For the integral \( \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \sin^2 x \, dx, \) let us compute the 5 point Chebyshev-Gauss estimate. First assign \( f(x) = \sin^2 x. \) From (41), we get
\[
\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \sin^2 x \, dx \sim (\pi/5) \sum_{j=1}^{n} \sin^2 x_j
\]

where

\[
x_j = \cos[(2j-1)\pi/10] \quad \text{for} \quad j = 1, \ldots, n.
\]

It turns out that

\[
\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \sin^2 x \, dx \sim 1.2191070
\]

where the error \( E \) is given by

\[
E = \frac{2\pi}{(2^{10}(10!))} \cdot |\xi(10)(\xi)| \cdot \frac{\pi}{(10!)} |\cos 2\xi|
\]

for some \( \xi \) in \([-1,1]\). So, \( E \leq \pi/(10!) \leq 0.0000008 \) which allows us to write

\[
1.2191062 \leq \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \sin^2 x \, dx \leq 1.2191078. \quad (43)
\]

Approximations to the integral above were also computed by using the composite forms of the rectangle rule, the trapezoid rule, and Simpson's rule in order to compare accuracy. A change of variable was employed so as to get around the improper integral. Note that

\[
\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \sin^2 x \, dx = \int_{-\pi/2}^{\pi/2} \sin^2 (\sin \theta) \, d\theta.
\]

With a step size of \( \pi/16 \) all three approximations were close to the range (43):

\[
\begin{align*}
\text{rectangle rule estimate} &= 1.2191060 \\
\text{trapezoid rule estimate} &= 1.2191060 \\
\text{Simpson's rule estimate} &= 1.2191050.
\end{align*}
\]

(44)

Due to the finite arithmetic of the computer used, decreased step size only resulted in poorer estimates.
To give an idea of the superiority of the Chebyshev-Gauss estimate for this example, the slightly poorer estimates of (44) were obtained by 16, 17, and 33 evaluations of \( f(x) \), respectively, while the Chebyshev-Gauss estimate required only 5 such evaluations.

Many more practical applications for Chebyshev polynomials exist other than those included in this paper. Among the most important of these is in the acceleration of the rate of convergence of iterative processes for solving certain classes of linear algebraic equations. Another use is in a method for finding eigenvalues of matrices. There also exists some theory for finding the representation in Chebyshev series of functions of two independent variables. However, many of the methods of this theory are speculative, seldom of practical value, and generally lacking a rigorous background, (2, p. 155).
REFERENCES


BIBLIOGRAPHY


