3

Chebyshev and Descartes Systems

Overview

A *Chebyshev space* is a finite-dimensional subspace of $C(A)$ of dimension $n+1$ that has the property that any element that vanishes at $n+1$ points vanishes identically. Such spaces, whose prototype is the space $\mathcal{P}_n$ of real algebraic polynomials of degree at most n, share with the polynomials many basic properties. The first section is an introduction to these Chebyshev spaces. A basis for a Chebyshev space is called a *Chebyshev system*. Two special families of Chebyshev systems, namely, *Markov systems* and *Descartes systems*, are examined in the second section. The third section examines the *Chebyshev "polynomials"* associated with Chebyshev spaces. These associated Chebyshev polynomials, which equioscillate like the usual Chebyshev polynomials, are extremal for various problems in the supremum norm. The fourth section studies particular Descartes systems

$$\left(x^{\lambda_0}, x^{\lambda_1}, \dots\right), \qquad \lambda_0 < \lambda_1 < \dots$$

on $(0,\infty)$ in detail. These systems, which we call *Müntz systems*, can be very explicitly orthonormalized on $[0,1]$, and this orthogonalization is also examined. The final section constructs Chebyshev "polynomials" associated with the Chebyshev spaces

$$\operatorname{span}\left\{1, \frac{1}{x-a_1}, \dots, \frac{1}{x-a_n}\right\}, \qquad a_i \in \mathbb{R}\setminus[-1,1]$$

on $[-1,1]$ and explores their various properties.

3.1 Chebyshev Systems

From an approximation theoretic point of view an essential property that polynomials of degree at most n have is that they can uniquely interpolate at $n+1$ points. This is equivalent to the fact that a polynomial of degree at most n that vanishes at $n+1$ points vanishes identically. Any $(n+1)$-dimensional vector space of continuous functions with this property is called a *Chebyshev space* or sometimes a *Haar space*. Many basic approximation properties extend to these spaces. The precise definition is the following.

Definition 3.1.1 (Chebyshev System). Let A be a Hausdorff space. The sequence $(f_0, \dots, f_n)$ is called a *real* (or *complex*) *Chebyshev system* or *Haar system* of dimension $n+1$ on A if $f_0, \dots, f_n$ are real- (or complex-) valued continuous functions on A, $\operatorname{span}\{f_0, \dots, f_n\}$ over $\mathbb{R}$ (or $\mathbb{C}$) is an $(n+1)$-dimensional subspace of $C(A)$, and any element of $\operatorname{span}\{f_0, \dots, f_n\}$ that has $n+1$ distinct zeros in A is identically zero.

If $(f_0, \dots, f_n)$ is a Chebyshev system on A, then $\operatorname{span}\{f_0, \dots, f_n\}$ is called a *Chebyshev space* or *Haar space* on A.

Chebyshev systems and spaces will be assumed to be real, unless we explicitly specify otherwise. If $A \subset \mathbb{R}$, then the topology on A is always meant to be the usual metric topology.

Implicit in the definition is that A contains at least $n+1$ points. Being a Chebyshev system is a property of the space spanned by the elements of the system, so every basis of a Chebyshev space is a Chebyshev system.

A point $x_0 \in (a,b)$ is called a *double zero* of an $f \in C[a,b]$ if $f(x_0) = 0$ and $f(x_0 - \epsilon) f(x_0 + \epsilon) > 0$ for all sufficiently small $\epsilon > 0$ (in other words, if f vanishes without changing sign at x_0). It is easy to see that if $(f_0, \dots, f_n)$ is a Chebyshev system on $[a,b] \subset \mathbb{R}$, then every $0 \neq f \in \operatorname{span}\{f_0, \dots, f_n\}$ has at most n zeros even if each double zero is counted twice; see E.10. Chebyshev spaces are defined via zero counting, and many of the theorems in the theory of Chebyshev spaces are proved by zero counting arguments. So it is important to make the agreement that, unless it is stated explicitly otherwise, we count the zeros of an element f from a Chebyshev space on $[a,b]$ so that each double zero of f is counted twice.

The following simple equivalences hold:

Proposition 3.1.2 (Equivalences). *Let $f_0, \dots, f_n$ be real- (or complex-) valued continuous functions on a Hausdorff space A (containing at least $n+1$ points). Then the following are equivalent:*

a] *Every $0 \neq p \in \operatorname{span}\{f_0, \dots, f_n\}$ has at most n distinct zeros in A.*

b] *If $x_0, \dots, x_n$ are distinct elements of A and $y_0, \dots, y_n$ are real (or complex) numbers, then there exists a unique $p \in \operatorname{span}\{f_0, \dots, f_n\}$ such that*

$$p(x_i) = y_i\,, \qquad i = 1, 2, \dots, n\,.$$

c] *If $x_0, \dots, x_n$ are distinct points of A, then*

$$D(x_0, \dots, x_n) := \begin{vmatrix} f_0(x_0) & \dots & f_n(x_0) \\ \vdots & \ddots & \vdots \\ f_0(x_n) & \dots & f_n(x_n) \end{vmatrix} \neq 0\,.$$

Proof. These equivalences are all elementary facts in linear algebra. □

On an interval there is a sign regularity to the determinants in c].

Proposition 3.1.3. *Suppose $(f_0, \dots, f_n)$ is a (real) Chebyshev system on $[a, b] \subset \mathbb{R}$. Then there exists a $\delta := -1$ or $\delta := 1$ such that*

$$\delta \begin{vmatrix} f_0(x_0) & \dots & f_n(x_0) \\ \vdots & \ddots & \vdots \\ f_0(x_n) & \dots & f_n(x_n) \end{vmatrix} > 0$$

for any $a \leq x_0 < x_1 < \cdots < x_n \leq b$.

Proof. This follows immediately from part c] of the previous proposition and continuity considerations. That is, if $D(x_0, \dots, x_n) < 0$ while $D(y_0, \dots, y_n) > 0$, then for some $\lambda \in (0, 1)$

$$D(\lambda x_0 + (1-\lambda) y_0, \dots, \lambda x_n + (1-\lambda) y_n) = 0\,,$$

which is impossible. □

The intimate relationship between Chebyshev systems and best approximation in the uniform norm is indicated by the next result. In order to state it we need to introduce the notion of an alternation sequence.

Definition 3.1.4 (Alternation Sequence). Let $A \subset \mathbb{R}$ and let

$$x_0 < x_1 < \cdots < x_n$$

be $n+1$ points of A. Then $(x_0, x_1, \dots, x_n)$ is said to be an *alternation sequence* of *length* $n+1$ for a real valued $f \in C(A)$ if

$$|f(x_i)| = \|f\|_A\,, \qquad i = 0, 1, \dots, n$$

and

$$\operatorname{sign}(f(x_{i+1})) = -\operatorname{sign}(f(x_i))\,, \qquad i = 0, 1, \dots, n-1\,.$$

Definition 3.1.5 (Best Approximation). Suppose that U is a (finite-dimensional) subspace of a normed space $(V, \|\cdot\|)$. If $g \in V$ and $p \in U$ satisfy

$$\|g - p\| = \inf_{h \in U} \|g - h\|,$$

then p is said to be a *best approximation* to g from U.

As a result of the finite dimensionality of the subspace U, at least one best approximation to any $g \in V$ from U exists. This is straightforward since

$$T := \{p \in U : \|p\| \leq \|g\| + 1\}$$

is a compact subset of U, so any sequence (p_j) of approximations to g from U satisfying

$$\|g - p_j\| \leq j^{-1} + \inf_{h \in U} \|g - h\|$$

has a convergent subsequence with limit in U. This limit is then a best approximation to g from U.

Theorem 3.1.6 (Alternation of Best Approximations). *Suppose $(f_0, \ldots, f_n)$ is a Chebyshev system on $[a, b] \subset \mathbb{R}$. Let A be a closed subset of $[a, b]$ containing at least $n + 2$ distinct points. Then $p \in H_n := \operatorname{span}\{f_0, \ldots, f_n\}$ is a best approximation to $g \in C(A)$ from H_n in the uniform norm on A if and only if there exists an alternation sequence of length $n + 2$ for $g - p$ on A.*

Proof. The proof of the *only if* part of the theorem is mostly an example of a standard type of perturbation argument that will recur later.

The perturbation argument goes as follows. Suppose p is a best approximation of required type and suppose a alternation sequence of maximal length for $g - p$ is

$$(x_0 < x_1, < \cdots < x_m)$$

where $x_i \in A$ and where m is strictly less than $n + 1$. Suppose, without loss of generality, that

$$g(x_0) - p(x_0) > 0$$

(otherwise multiply by -1). Now let

$$Y := \{x \in A : |g(x) - p(x)| = \|g - p\|_A\}.$$

Note that Y is compact. Since $(x_0 < x_1 < \ldots < x_m)$ is an alternation sequence of maximal length, we can divide Y into $m + 1$ disjoint compact subsets $Y_0, Y_1, \ldots, Y_m$ with

$$x_0 \in Y_0, \quad x_1 \in Y_1, \ldots, \quad x_m \in Y_m$$

so that

$$\operatorname{sign}(g(x)-p(x)) = -\operatorname{sign}(g(y)-p(y)) \neq 0\,, \qquad x \in Y_i\,,\ \ y \in Y_{i+1}\,.$$

Now choose a $p^* \in \operatorname{span}\{f_0,\dots,f_n\}$ satisfying

$$\operatorname{sign}_{x\in Y_i}(p^*(x)) = (-1)^i\,, \qquad i = 0,1,\dots,m\,.$$

This can be done by choosing points z_i with

$$\max Y_{i-1} < z_i < \min Y_i\,, \qquad i = 1,2,\dots,m$$

and then applying E.11. We now claim that, for $\delta > 0$ sufficiently small,

$$\|g-(p+\delta p^*)\|_A < \|g-p\|_A\,, \tag{3.1.1}$$

which contradicts the fact that p is a best approximation, and so there must exist an alternation set of length $n+2$ for $g-p$ on A. To verify (3.1.1) we proceed as follows:

For each $i = 0,1,\dots,m$ choose an open set $O_i \subset [a,b]$ (in the usual metric topology relative to $[a,b]$) containing Y_i so that for every $x \in \overline{O}_i$,

$$\operatorname{sign}(g(x)-p(x)) = \operatorname{sign}(p^*(x)) \tag{3.1.2}$$

and

$$|g(x)-p(x)| \geq \tfrac{1}{2}\|g-p\|_A\,. \tag{3.1.3}$$

Now pick a $\delta_1 > 0$ such that for every $x \in B := A \setminus \bigcup_{i=0}^m O_i$ and $\delta \in (0,\delta_1)$,

$$|g(x)-(p(x)+\delta p^*(x))| < \|g-p\|_A\,,$$

which can be done since B is compact and by construction

$$\|g-p\|_B < \|g-p\|_A\,.$$

Note that (3.1.2) and (3.1.3) allow us to pick a $\delta_2 > 0$ such that for $x \in \bigcup_{i=0}^m \overline{O}_i$ and $\delta \in (0,\delta_2)$,

$$|g(x)-(p(x)+\delta p^*(x))| < |g(x)-p(x)|\,.$$

This verifies (3.1.1) and finishes the direct half of the theorem.

The proof of the converse is simple. Suppose there is an alternation sequence of length $n+2$ for $g-p$ on A, and suppose there exists a p^* with

$$\|g-p^*\|_A < \|g-p\|_A.$$

Then p^*-p has at least $n+1$ zeros on $[a,b]$, one between any two consecutive alternation points for $g-p$ on A, and hence it vanishes identically. This contradiction finishes the proof. □

In the setting of Theorem 3.1.6 the best approximation is unique; see E.5.

Comments, Exercises, and Examples.

The terminology is not entirely standard in the literature with Chebyshev systems often referred to as Haar systems on intervals of $\mathbb{R}$. There are various proofs of Theorem 3.1.6, of which ours is by no means the most elegant (see Cheney [66], for example). The point of this proof is that it easily modifies to deal with characterizations of extremal functions for various other extremal problems. Many good books cover this standard material. See, for example, Cheney [66], Lorentz [86a], or Pinkus [89]; see also Appendix 3. An extensive treatment of Chebyshev systems is available in Karlin and Studden [66] or Nürnberger [89], where E.3 can be found. E.4 shows that real Chebyshev systems are intrinsically one-dimensional.

E.1 Examples of Chebyshev systems.

a] Suppose $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that

$$(x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n})$$

is a Chebyshev system on $[0, \infty)$.

b] Suppose $\lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that

$$(x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n})$$

is a Chebyshev system on $(0, \infty)$.

c] Suppose $\lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that

$$(x^{\lambda_0},\, x^{\lambda_0} \log x\,,\, x^{\lambda_1},\, x^{\lambda_1} \log x\,, \ldots,\, x^{\lambda_n},\, x^{\lambda_n} \log x)$$

is a Chebyshev system on $(0, \infty)$.

d] Suppose $\lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that

$$\left(\frac{1}{x - \lambda_0}\,,\, \frac{1}{x - \lambda_1}\,, \cdots,\, \frac{1}{x - \lambda_n}\right)$$

is a Chebyshev system on $(-\infty, \infty) \setminus \{\lambda_0, \lambda_1, \ldots, \lambda_n\}$.

e] Suppose $\lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that

$$(e^{\lambda_0 x},\, e^{\lambda_1 x},\, \ldots,\, e^{\lambda_n x})$$

is a Chebyshev system on $(-\infty, \infty)$.

f] Show that

$$(1,\, \cos\theta,\, \sin\theta,\, \cos 2\theta,\, \sin 2\theta,\, \ldots,\, \cos n\theta,\, \sin n\theta)$$

is a Chebyshev system on $[0, 2\pi)$.

g] Show that

$$(1,\, \cos\theta,\, \cos 2\theta,\, \ldots,\, \cos n\theta)$$

is a Chebyshev system on $[0, \pi)$.

E.2 More Examples.

a] If $(f_0, \dots, f_n)$ is a Chebyshev system on A, then it is also a Chebyshev system on any subset B of A containing at least $n+1$ points.

b] If $(f_0, \dots, f_n)$ is a Chebyshev system on A and $g \in C(A)$ is strictly positive on A, then $(gf_0, \dots, gf_n)$ is also a Chebyshev system on A.

c] If $(f_0, \dots, f_n)$ is a Chebyshev system on $[0,1]$, then

$$\left(1, \int_0^x f_0(t)\,dt, \dots, \int_0^x f_n(t)\,dt\right)$$

is also a Chebyshev system on $[0,1]$.

See E.8 of Section 3.2, which treats the effect of differentiation on a Chebyshev system.

E.3 Extended Complete Chebyshev Systems. Let $(g_0, \dots, g_n)$ be a sequence of functions in $C^n[a,b]$. Define the *Wronskian determinant*

$$W(g_0, \dots, g_m)(t) := \begin{vmatrix} g_0(t) & g_1(t) & \dots & g_m(t) \\ g_0'(t) & g_1'(t) & \dots & g_m'(t) \\ \vdots & \vdots & \ddots & \vdots \\ g_0^{(m)}(t) & g_1^{(m)}(t) & \dots & g_m^{(m)}(t) \end{vmatrix}.$$

We say that $(g_0, \dots, g_n)$ is an *extended complete Chebyshev system (ECT system)* on $[a,b]$ if

$$W(g_0, \dots, g_m)(t) > 0, \quad m = 0, 1, \dots, n, \quad t \in [a,b].$$

a] Let $\operatorname{span}\{g_0, \dots, g_n\}$ be an $(n+1)$-dimensional subspace of $C^n[a,b]$. Show that the following statements are equivalent:

(i) For every $m = 0, 1, \dots, n$, $0 \neq f \in \operatorname{span}\{g_0, \dots, g_m\}$ has at most m zeros in $[a,b]$ counting multiplicities ($x_0 \in [a,b]$ is a zero of f with multiplicity k if $f(x_0) = f'(x_0) = \cdots f^{(k-1)}(x_0) = 0$ and $f^{(k)}(x_0) \neq 0$).

(ii) For each $i = 0, 1, \dots, n$, there exists a choice of $\delta_i := 1$ or $\delta_i := -1$ such that

$$(\delta_0 g_0,\ \delta_1 g_1, \dots,\ \delta_n g_n)$$

is an ECT system on $[a,b]$.

In particular, every ECT system on $[a,b]$ is a Chebyshev system on $[a,b]$.

Proof. For details see Karlin and Studden [66]. □

b] Characterization Theorem. *The following statements are equivalent:*

(i) $(g_0, \dots, g_n)$ *is an ECT system on* $[a,b]$.

(ii) There exist $w_i \in C^{n-i}[a,b]$, $i = 0, 1, \dots, n$, *with* w_i *strictly positive such that*

$$\begin{aligned} g_0(t) &= w_0(t)\,, \\ g_1(t) &= w_0(t) \int_a^t w_1(t_1)\, dt_1\,, \\ &\vdots \\ g_n(t) &= w_0(t) \int_a^t w_1(t_1) \int_a^{t_1} w_2(t_2) \cdots \int_a^{t_{n-1}} w_n(t_n)\, dt_n \cdots dt_2\, dt_1\,. \end{aligned}$$

Proof. This is proved by induction on n. See Karlin and Studden [66]. □

c] Suppose $\lambda_0 < \lambda_1 < \cdots < \lambda_n$. Show that $(x^{\lambda_0}, \dots, x^{\lambda_n})$ is an ECT system on $[a,b]$ provided $a > 0$.

E.4 Railway Track Theorem. Real Chebyshev systems exist only on very special subsets of $\mathbb{R}^m$. Indeed, real Chebyshev systems intrinsically live on one-dimensional subsets.

a] Suppose $A \subset \mathbb{R}^m$ contains three distinct arcs that join at a point x_0. Then, for $n \geq 2$, there exists no real Chebyshev system $(f_0, \dots, f_n)$ on A.

Proof. Suppose there exists a real Chebyshev system $(f_0, \dots, f_n)$ on such a set A. Let

$$V(x,y) := D(x, y, x_2, x_3, \dots, x_n)$$

(D is defined in Proposition 3.1.2) which is never zero for distinct points $x, y, x_2, x_3, \dots, x_n$. Choose distinct points $x_0, x_1, \dots, x_n$ on one of the three distinct arcs so that x_0 is adjacent to x_1. Pick the points $z_1 \neq x_0$ and $z_2 \neq x_0$ so that z_1, z_2, and x_1 are on different arcs. Now consider interchanging $x := x_0$ and $y := x_1$ by moving x from x_0 to z_1, y from x_1 to z_2, x from z_1 to x_1, and y from z_2 to x_0. Since $x, y, x_2, x_3, \dots, x_n$ remain distinct, $V(x,y)$ does not vanish in this process. This contradicts the fact that $V(x,y)$ is continuous and $V(x_0, x_1) = -V(x_1, x_0)$. □

The following more general result of Mairhuber [56] also holds:

b] Mairhuber's Theorem. *If* $(f_0, \dots, f_n)$ *is a real Chebyshev space on* A, *then* A *is homeomorphic to a subset of the unit circle.*

E.5 Uniqueness of Best Approximations. Prove that a best approximation from a Chebyshev space satisfying the conditions of Theorem 3.1.6 is unique.

Hint: Suppose f has two best approximations $p_1 \in H_n$ and $p_2 \in H_n$. Then, by the alternation characterization, $p_1 - p_2 \in H_n$ has at least $n+1$ zeros on $[a,b]$ (we count each internal zero without sign change twice). Now E.10 implies that $p_1 - p_2 = 0$. □

E.6 De la Vallée Poussin Theorem. *Suppose H_n is a Chebyshev space of dimension $(n+1)$ on $[a,b]$. If $p \in H_n$ and there exist $n+2$ points*

$$a \le x_0 < x_1 < \cdots < x_{n+1} \le b$$

so that

$$\operatorname{sign}(f(x_i) - p(x_i)) = -\operatorname{sign}(f(x_{i+1}) - p(x_{i+1})), \quad i = 0, 1, \ldots, n,$$

then

$$\inf_{p \in H_n} \|f - p\|_{[a,b]} \ge \min_{i=0,\ldots,n+1} |f(x_i) - p(x_i)|.$$

E.7 Haar's Characterization of Chebyshev Spaces. The following pretty theorem is due to Haar (for a proof, see E.3 of Appendix 3):

Theorem. *Let $f_0, \ldots, f_n \in C(A)$ where A is a compact Hausdorff space containing at least $n+1$ points. Then $(f_0, \ldots, f_n)$ is a Chebyshev system on A if and only if every $g \in C(A)$ has a unique best approximation from* $\operatorname{span}\{f_0, \ldots, f_n\}$ *in the uniform norm on A.*

E.8 Best Approximation to x^n. Reprove Theorem 2.1.1 by using the alternation characterization of best approximations.

E.9 Best Rational Approximations. Let $f \in C[a,b]$. Then $p/q \in \mathcal{R}_{n,m}$ is a best approximation to f from $\mathcal{R}_{n,m}$ in $C[a,b]$ if and only if $f - p/q$ has an alternation set of length at least

$$2 + \max\{n + \deg(q), m + \deg(p)\}$$

$[a,b]$. (Here we must assume p/q is written in a reduced form.)

The proof of this is a fairly complicated variant of the proof of Theorem 3.1.6 (see, for example, Cheney [66]).

E.10 Zeros of Functions in Chebyshev Spaces. As before, we call the point $x_0 \in (a,b)$ a double zero of $f \in C[a,b]$ if $f(x_0) = 0$ and

$$f(x_0 - \epsilon) f(x_0 + \epsilon) > 0$$

for all sufficiently small $\epsilon > 0$ (in other words, if f vanishes without changing sign at x_0). Let $(f_0, \ldots, f_n)$ be a Chebyshev system on $[a,b] \subset \mathbb{R}$. Show that every $0 \ne f \in \operatorname{span}\{f_0, \ldots, f_n\}$ has at most n zeros even if each double zero is counted twice.

Hint: Use Proposition 3.1.2 b]. □

E.11 Functions in a Chebyshev Space with Prescribed Sign Changes. Let $(f_0, \dots, f_n)$ be a Chebyshev system on $[a, b]$, and let

$$a < z_1 < z_2 < \cdots < z_m < b\,, \qquad 0 \le m \le n\,.$$

Show that there is a function $p^* \in \operatorname{span}\{f_0, \dots, f_n\}$ such that

(i) $p^*(x) = 0$ if and only if $x = z_i$ for some $i = 1, 2, \dots, m$,

(ii) $p^*(x)$ changes sign at each z_i, $i = 1, 2, \dots, m$.

Hint: If $m = n$, then use Proposition 3.1.3 and a continuity argument to show that

$$p^*(x) = \begin{vmatrix} f_0(x) & f_1(x) & \dots & f_n(x) \\ f_0(z_1) & f_1(z_1) & \dots & f_n(z_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(z_n) & f_1(z_n) & \dots & f_n(z_n) \end{vmatrix}$$

satisfies the requirements.

If $m < n$, then use the already proved case, a limiting argument, and E.10 to show that there are $p_j \in \operatorname{span}\{f_0, \dots, f_n\}$, $j = 1, 2$, such that

(1) $p_j(x)$ changes sign at x if and only if $x = z_i$, $i = 1, 2, \dots, m$,

(2) $p_1(x) \neq 0$ for every $x \in [a, z_m] \setminus \{z_1, z_2, \dots, z_m\}$,

(3) $p_2(x) \neq 0$ for every $x \in [z_1, b] \setminus \{z_1, z_2, \dots, z_m\}$.

Now show that either $p^* := p_1 + p_2$ or $p^* := p_1 - p_2$ satisfies the requirements. □

E.12 The Dimension of a Chebyshev Space on a Circle. Let $(f_0, \dots, f_n)$ be a Chebyshev system on a circle C. Show that n must be even. Observe that such Chebyshev systems exist.

Hint: Show that for every set of n distinct points $x_1, x_2, \dots, x_n$ on the circle there is a $p \in \operatorname{span}\{f_0, \dots, f_n\}$ such that $p(x) = 0$ if and only if $x \in \{x_1, x_2, \dots, x_n\}$ and $p(x)$ changes sign at each x_i. □

3.2 Descartes Systems

Chebyshev systems capture some of the essential properties of polynomials. There are two additional types of systems that capture some additional properties.

Definition 3.2.1 (Markov System). We say that $(f_0, \dots, f_n)$ is a *Markov system* on a Hausdorff space A if each $f_i \in C(A)$, and $\{f_0, \dots, f_m\}$ is a Chebyshev system for each $m = 0, 1, \dots, n$. (We allow n to tend to $+\infty$, in which case we call the system an *infinite Markov system* on A.)

A Markov system is just a Chebyshev system with each initial segment also a Chebyshev system. Being a Chebyshev system is a property of the space not of the basis. However, the Markov system depends on the basis. For example,

$$(x^{\lambda_0}, x^{\lambda_1}, \dots), \qquad \lambda_0 < \lambda_1 < \cdots$$

is a Markov system on any $A \subset (0, \infty)$ containing infinitely many points, but not every basis of $\operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \dots\}$ is a Markov system on A (see E.1).

Proposition 3.2.2. *$(f_0, \dots, f_n)$ is a Markov system on an interval $[a,b]$ if and only if for each $i = 0, 1, \dots, n$, there exists a choice of $\delta_i := 1$ or $\delta_i := -1$ such that with $g_i := \delta_i f_i$,*

$$D\begin{pmatrix} g_0 & g_1 & \dots & g_m \\ x_0 & x_1 & \dots & x_m \end{pmatrix} := \begin{vmatrix} g_0(x_0) & \dots & g_m(x_0) \\ \vdots & \ddots & \vdots \\ g_0(x_m) & \dots & g_m(x_m) \end{vmatrix} > 0$$

for every $a \le x_0 < x_1 < \cdots < x_m \le b$ and $m = 0, 1, \dots, n$.

Proof. This is an easy consequence of Proposition 3.1.3 by induction on n. □

A stronger property that a system on an interval can have is the following:

Definition 3.2.3 (Descartes System). The system $(f_0, \dots, f_n)$ is said to be a *Descartes system* (or *order complete Chebyshev system*) on an interval I if each $f_i \in C(I)$ and

$$D\begin{pmatrix} f_{i_0} & f_{i_1} & \dots & f_{i_m} \\ x_0 & x_1 & \dots & x_m \end{pmatrix} > 0$$

for any $0 \le i_0 < i_1 < \cdots < i_m \le n$ and $x_0 < x_1 < \cdots < x_m$ from I. (Once again we allow n to tend to ∞.)

This again is a property of the basis. It implies that any finite-dimensional subspace generated by some basis elements is a Chebyshev space on I. The canonical example of a Descartes system on $[a,b]$, $a > 0$, is

$$(x^{\lambda_0}, x^{\lambda_1}, \dots), \qquad \lambda_0 < \lambda_1 < \cdots$$

(see E.2). A Descartes system on I is obviously a Descartes system on any subinterval of I.

The following version of Descartes' rule of signs holds for Descartes systems.

Theorem 3.2.4 (Descartes' Rule of Signs). *If $(f_0, \dots, f_n)$ is a Descartes system on $[a,b]$, then the number of distinct zeros of any*

$$0 \neq f = \sum_{i=0}^{n} a_i f_i\,, \qquad a_i \in \mathbb{R}$$

is not greater than the number of sign changes in $(a_0, \dots, a_n)$.

A sign change occurs between a_i and a_{i+k} exactly when $a_i a_{i+k} < 0$ and $a_{i+1} = a_{i+2} = \cdots = a_{i+k-1} = 0$.

Proof. Suppose $(a_0, \dots, a_n)$ has p sign changes. Then we can partition $\{a_0, \dots, a_n\}$ into exactly $p+1$ blocks so that each block is of the form

$$a_{n_m+1}, a_{n_m+2}, \dots, a_{n_{m+1}}\,, \qquad m = 0, 1, \dots, p$$

$(n_0 := -1, n_{p+1} := n)$, where all of the coefficients in each of the blocks are of the same sign and not all the coefficients in a block vanish. Now let

$$g_m := \sum_{i=n_m+1}^{n_{m+1}} |a_i| f_i\,, \qquad m = 0, 1, \dots, p\,.$$

Then, for $a \leq x_0 < x_1 < \cdots < x_p \leq b$,

$$D\begin{pmatrix} g_0 & g_1 & \dots & g_p \\ x_0 & x_1 & \dots & x_p \end{pmatrix} = \sum_{i_0=n_0+1}^{n_1} \cdots \sum_{i_p=n_p+1}^{n_{p+1}} |a_{i_0}| \cdots |a_{i_p}| D\begin{pmatrix} f_{i_0} & f_{i_1} & \dots & f_{i_p} \\ x_0 & x_1 & \dots & x_p \end{pmatrix} > 0$$

since each of the determinants in the sum is positive. Thus $\{g_0, \dots . g_p\}$ is a $(p+1)$-dimensional Chebyshev system on $[a,b]$, and hence

$$f := \sum_{i=0}^{p} \delta_i g_i\,, \qquad \delta_i = \pm 1$$

has at most p zeros. This finishes the proof. □

A refined version of Descartes' rule of signs for ordinary polynomials is presented in the exercises. The following comparison theorem due to Pinkus and, independently, Smith [78] will be of use later.

Theorem 3.2.5. *Suppose* $(f_0, \dots, f_n)$ *is a Descartes system on* $[a,b]$. *Suppose*

$$p = f_\alpha + \sum_{i=1}^{k} a_i f_{\lambda_i}\,, \qquad \text{and} \qquad q = f_\alpha + \sum_{i=1}^{k} b_i f_{\gamma_i}\,,$$

where $0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_k \le n$, $0 \le \gamma_1 < \gamma_2 < \cdots < \gamma_k \le n$,

$$0 \le \gamma_i \le \lambda_i < \alpha\,, \qquad i = 1, 2, \dots, m\,,$$

and

$$\alpha < \lambda_i \le \gamma_i \le n\,, \qquad i = m+1, m+2, \dots, k$$

with strict inequality for at least one index $i = 1, 2, \dots, k$. *If*

$$p(x_i) = q(x_i) = 0\,, \qquad i = 1, 2, \dots, k\,,$$

where $x_i \in [a,b]$ *are distinct, then*

$$|p(x)| \le |q(x)|$$

for all $x \in [a,b]$ *with strict inequality for* $x \ne x_i$.

The proof is left as a guided exercise (see E.4) with some interesting consequences presented in E.5.

Comments, Exercises, and Examples.

Theorem 3.2.4 characterizes Descartes systems; see Karlin and Studden [66, p. 25]. Some caution must be exercised since, as in the previous section, definitions are not entirely standard. We will explore two particular Descartes systems in greater detail later; see E.2 and E.3. For further material, the reader is referred to Karlin and Studden [66], Karlin [68], and Nürnberger [89].

E.1 Distinctions.

a] Given $\lambda_0 < \lambda_1 < \cdots$ and $A \subset (0, \infty)$ with infinitely many points, show that $(x^{\lambda_0}, x^{\lambda_1}, \dots)$ is a Markov system on A, but there is a basis for $\operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \dots\}$, which is not a Markov system on A.

b] Find a Markov system that is not a Descartes system.

E.2 Examples of Descartes Systems.

a] Suppose $\lambda_0 < \lambda_1 < \cdots$. Show that the *Müntz system*

$$(x^{\lambda_0}, x^{\lambda_1}, \dots)$$

is a Descartes system on $(0, \infty)$.

Hint: For every $0 \le i_0 < i_1 < \cdots < i_m$, the determinant

$$D\begin{pmatrix} x^{\lambda_{i_0}} & x^{\lambda_{i_1}} & \dots & x^{\lambda_{i_m}} \\ x_0 & x_1 & \dots & x_m \end{pmatrix} = \begin{vmatrix} x_0^{\lambda_{i_0}} & \dots & x_0^{\lambda_{i_m}} \\ \vdots & \ddots & \vdots \\ x_m^{\lambda_{i_0}} & \dots & x_m^{\lambda_{i_m}} \end{vmatrix}$$

is nonzero for any $0 < x_0 < x_1 < \cdots < x_m < \infty$ by Proposition 3.1.2 and E.1 a] of Section 3.1. It only remains to prove that it is positive whenever $0 < x_0 < x_1 < \cdots < x_m < \infty$. Observe that the exponents λ_{i_j} can be varied continuously (for fixed x_i) without changing the sign of the determinant provided no two ever become equal. Now perturb $(\lambda_{i_0}, \lambda_{i_1}, \dots, \lambda_{i_m})$ into $(0, 1, \dots, m)$ and observe that the determinant becomes a Vandermonde determinant, as in part b], which in this case is positive. □

b] Vandermonde Determinant. Show that

$$\begin{vmatrix} 1 & x_0 & \dots & x_0^m \\ 1 & x_1 & \dots & x_1^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & \dots & x_m^m \end{vmatrix} = \prod_{0 \le i < j \le m} (x_j - x_i)\,.$$

Hint: The determinant is a polynomial in $x_0, x_1, \dots, x_m$ of degree m in each variable that vanishes whenever $x_i = x_j$. □

c] Suppose $\lambda_0 < \lambda_1 < \cdots$. Show that the *exponential system*

$$(e^{\lambda_0 t},\, e^{\lambda_1 t},\, \dots)$$

is a Descartes system on $(-\infty, \infty)$.

Hint: Use part a]. □

d] Suppose $0 < \lambda_0 < \lambda_1 < \cdots$. Show that

$$(\sinh \lambda_0 t,\, \sinh \lambda_1 t,\, \dots)$$

is a Descartes system on $(0, \infty)$.

Outline. Let $0 \le i_0 < i_1 < \cdots < i_m$ be fixed integers. First we show that

$$(\sinh \lambda_{i_0} t,\, \sinh \lambda_{i_1} t,\, \dots,\, \sinh \lambda_{i_m} t)$$

is a Chebyshev system on $(0, \infty)$. Indeed, let

$$0 \not\equiv f \in \operatorname{span}\{\sinh \lambda_{i_0} t,\, \sinh \lambda_{i_1} t,\, \dots,\, \sinh \lambda_{i_m} t\}\,.$$

Then

$$0 \not\equiv f \in \operatorname{span}\{e^{\pm\lambda_{i_0} t},\, e^{\pm\lambda_{i_1} t},\, \dots,\, e^{\pm\lambda_{i_m} t}\}$$

and by E.1 e] of Section 3.1, f has at most $2m$ zeros in $(-\infty, \infty)$. Since f is odd, it has at most m zeros in $(0, \infty)$.

Since for every $0 \le i_0 < i_1 < \cdots < i_m$, $(\sinh \lambda_{i_0} t, \ldots, \sinh \lambda_{i_m} t)$ is a Chebyshev system on $(0,\infty)$, the determinant

$$D \begin{pmatrix} \sinh \lambda_{i_0} t & \sinh \lambda_{i_1} t & \ldots & \sinh \lambda_{i_m} t \\ x_0 & x_1 & \ldots & x_m \end{pmatrix}$$

$$= \begin{vmatrix} \sinh \lambda_{i_0} x_0 & \sinh \lambda_{i_1} x_0 & \ldots & \sinh \lambda_{i_m} x_0 \\ \sinh \lambda_{i_0} x_1 & \sinh \lambda_{i_1} x_1 & \ldots & \sinh \lambda_{i_m} x_1 \\ \vdots & \vdots & \ddots & \vdots \\ \sinh \lambda_{i_0} x_m & \sinh \lambda_{i_1} x_m & \ldots & \sinh \lambda_{i_m} x_m \end{vmatrix}$$

is nonzero for any $0 < x_0 < x_1 < \cdots < x_m < \infty$ Proposition 3.1.2. So it only remains to prove that it is positive whenever $0 < x_0 < x_1 < \cdots < x_m < \infty$. Now let

$$D(\alpha) := D \begin{pmatrix} \sinh \lambda_{i_0} t & \sinh \lambda_{i_1} t & \ldots & \sinh \lambda_{i_m} t \\ \alpha x_0 & \alpha x_1 & \ldots & \alpha x_m \end{pmatrix}$$

$$= \begin{vmatrix} \sinh \lambda_{i_0} \alpha x_0 & \sinh \lambda_{i_1} \alpha x_0 & \ldots & \sinh \lambda_{i_m} \alpha x_0 \\ \sinh \lambda_{i_0} \alpha x_1 & \sinh \lambda_{i_1} \alpha x_1 & \ldots & \sinh \lambda_{i_m} \alpha x_1 \\ \vdots & \vdots & \ddots & \vdots \\ \sinh \lambda_{i_0} \alpha x_m & \sinh \lambda_{i_1} \alpha x_m & \ldots & \sinh \lambda_{i_m} \alpha x_m \end{vmatrix}$$

and

$$D^*(\alpha) := D \begin{pmatrix} \frac{1}{2} e^{\lambda_{i_0} t} & \frac{1}{2} e^{\lambda_{i_1} t} & \ldots & \frac{1}{2} e^{\lambda_{i_m} t} \\ \alpha x_0 & \alpha x_1 & \ldots & \alpha x_m \end{pmatrix}$$

$$= \begin{vmatrix} \frac{1}{2} e^{\lambda_{i_0} \alpha x_0} & \frac{1}{2} e^{\lambda_{i_1} \alpha x_0} & \ldots & \frac{1}{2} e^{\lambda_{i_m} \alpha x_0} \\ \frac{1}{2} e^{\lambda_{i_0} \alpha x_1} & \frac{1}{2} e^{\lambda_{i_1} \alpha x_1} & \ldots & \frac{1}{2} e^{\lambda_{i_m} \alpha x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} e^{\lambda_{i_0} \alpha x_m} & \frac{1}{2} e^{\lambda_{i_1} \alpha x_m} & \ldots & \frac{1}{2} e^{\lambda_{i_m} \alpha x_m} \end{vmatrix}$$

where $0 < x_0 < x_1 < \cdots < x_m < \infty$ are fixed. Since

$$(\sinh \lambda_{i_0} t, \sinh \lambda_{i_1} t, \ldots, \sinh \lambda_{i_m} t)$$

and

$$(e^{\lambda_{i_0} t}, e^{\lambda_{i_1} t}, \ldots, e^{\lambda_{i_m} t})$$

are Chebyshev systems on $(0,\infty)$, $D(\alpha)$ and $D^*(\alpha)$ are continuous nonvanishing functions of α on $(0,\infty)$. Now, observe that

$$\lim_{\alpha \to \infty} |D(\alpha)| = \lim_{\alpha \to \infty} |D^*(\alpha)| = \infty \qquad \text{and} \qquad \lim_{\alpha \to \infty} \frac{D(\alpha)}{D^*(\alpha)} = 1 .$$

By part c],

$$(e^{\lambda_{i_0} t}, e^{\lambda_{i_1} t}, \ldots, e^{\lambda_{i_m} t})$$

is a Descartes system on $(-\infty,\infty)$, hence $D^*(\alpha) > 0$ for every $\alpha > 0$. So the above limit relations imply that $D(\alpha) > 0$ for every large enough α, hence for every $\alpha > 0$. In particular,

$$D(1) = D\begin{pmatrix} \sinh\lambda_{i_0}t & \sinh\lambda_{i_1}t & \dots & \sinh\lambda_{i_m}t \\ x_0 & x_1 & \dots & x_m \end{pmatrix} > 0\,,$$

which finishes the proof. □

e] Suppose $0 < \lambda_0 < \lambda_1 < \cdots$. Show that

$$(\cosh\lambda_0 t,\, \cosh\lambda_1 t,\, \dots)$$

is a Descartes system on $(0,\infty)$.

Hint: Proceed as in the outline for part d]. □

E.3 Rational Systems.

a] **Cauchy Determinants.** Show that

$$\begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_m} \\ \vdots & \ddots & \vdots \\ \frac{1}{\alpha_m+\beta_1} & \cdots & \frac{1}{\alpha_m+\beta_m} \end{vmatrix} = \frac{\prod\limits_{1\le i<j\le m}(\alpha_j-\alpha_i)(\beta_j-\beta_i)}{\prod\limits_{1\le i,j\le m}(\alpha_i+\beta_j)}\,.$$

Hint: Multiply both sides above by $\prod\limits_{1\le i,j\le m}(\alpha_i+\beta_j)$ and observe that both sides are polynomials of the same degree, $m-1$, in each variable α_i, β_i. Also both sides vanish exactly when $\alpha_i = \alpha_j$ or $\beta_i = \beta_j$. So up to a constant both sides are the same. Now show that the constant is 1. □

b] Let $\alpha_1 > \alpha_2 > \cdots > b$. Show that

$$\left(\frac{1}{\alpha_1 - x}\,,\, \frac{1}{\alpha_2 - x}\,, \dots\right)$$

is a Descartes system on $[a,b]$.

Let $\alpha_1 < \alpha_2 < \cdots < a$. Show that

$$\left(\frac{1}{x - \alpha_1}\,,\, \frac{1}{x - \alpha_2}\,, \dots\right)$$

is a Descartes system on $[a,b]$ (see also E.6 e]).

E.4 Proof of Theorem 3.2.5. Assume the notation of this Theorem 3.2.5.

a] Let $0 \le \delta_0 < \delta_1 < \cdots < \delta_\mu \le n$ and $a < x_1 < x_2 < \cdots < x_\mu < b$. Show that there exists a unique $p = f_{\delta_\mu} + \sum_{i=0}^{\mu-1} a_i f_{\delta_i}$ such that

(1) $p(x_i) = 0\,, \quad i = 1, 2, \ldots, \mu\,.$

Show also that the above p has the following properties:

(2) $p(x)$ changes sign at each x_i,

(3) $p(x) \neq 0$ if $x \notin \{x_1, x_2, \ldots, x_\mu\}$,

(4) $a_i a_{i+1} < 0\,, \quad i = 0, 1, \ldots, \mu - 1\,, \quad a_\mu := 1\,,$

(5) $p(x) > 0\,, \quad x \in (x_\mu, b]\,.$

Hint: Since $(f_{\delta_0}, \ldots, f_{\delta_\mu})$ and $(f_{\delta_0}, \ldots, f_{\delta_{\mu-1}})$ are Chebyshev systems, E.11 of Section 3.1 shows that there exists a p of the desired form satisfying (1). Since $(f_{\delta_0}, \ldots, f_{\delta_{\mu-1}})$ is a Chebyshev system, this p is unique. Now E.11 of Section 3.1 yields that p satisfies (2) and (3). By Theorem 3.2.4, p satisfies (4). The fact that (5) holds for p follows from expanding the determinant

$$D \begin{pmatrix} f_{\delta_0} & f_{\delta_1} & \ldots & f_{\delta_{\mu-1}} & f_{\delta_\mu} \\ x_1 & x_2 & \ldots & x_\mu & x \end{pmatrix}$$

by Cramer's rule. This determinant is just $cp(x)$ with some $c > 0$ since it vanishes at each x_i, and the coefficient of f_{δ_μ} is positive; see Definition 3.2.3. Also, the above determinant is positive for all $x \in (x_\mu, b]$; see Definition 3.2.3 again. □

b] Prove Theorem 3.2.5.

Outline. For notational simplicity assume that $\alpha = n$ (hence $m = k$); the general case is analogous. Further, we may assume that there is an index j such that

$$\gamma_j < \lambda_j \quad \text{and} \quad \gamma_i = \lambda_i \quad \text{whenever} \quad i \neq j$$

since the result follows from this by a finite number of pairwise comparisons. So we assume

$$p = f_n + a_j f_{\lambda_j} + \sum_{\substack{i=1 \\ i \neq j}}^{k} a_i f_{\lambda_i}$$

and

$$q = f_n + b_j f_{\gamma_j} + \sum_{\substack{i=1 \\ i \neq j}}^{k} b_i f_{\lambda_i}\,,$$

where $0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_k < n$ and $0 \leq \lambda_{j-1} < \gamma_j < \lambda_j$ for some $1 \leq j \leq k$ (of course, the inequality $\lambda_{j-1} < \gamma_j$ holds only if λ_{j-1} is defined, that is, only if $j \geq 2$). Then

$$p - q = a_j f_{\lambda_j} - b_j f_{\gamma_j} + \sum_{\substack{i=1 \\ i \neq j}}^{k} (a_i - b_i) f_{\lambda_i}$$

has exactly k zeros on $[a, b]$ at $x_1, x_2, \ldots, x_k$ because $p - q$ is in a $(k+1)$-dimensional Chebyshev subspace.

By property (5) in part a] applied to p and q, respectively,

$$p(x) > 0 \text{ and } q(x) > 0\,, \qquad x \in (x_k, b]\,.$$

Now by property (4) in part a] applied to $c(p-q)$, where c is chosen so that the lead coefficient of $c(p-q)$ is 1, and by the fact that p and $p-q$ have the same coefficient for f_{λ_j}, the lead coefficient of $p-q$ ($a_k - b_k$ provided $\lambda_k > \lambda_j$) is negative. So property (5) in part a] implies that

$$p(x) - q(x) < 0, \quad x \in (x_k, b]\,.$$

Hence $0 < p(x) < q(x)$, $x \in (x_k, b]$.

Now use property (3) in part a] and the fact that all of p, q, and $p-q$ change sign only at x_i to finish the proof. □

The following extension of part a] will be used later:

c] Suppose

$$0 \le \delta_0 < \delta_1 < \cdots < \delta_\mu \le n\,, \quad a \le x_1 \le x_2 \le \cdots \le x_\mu \le b\,,$$

$a < x_2, x_{\mu-1} < b$, and $x_i < x_{i+2}$, $i = 1, 2, \ldots, \mu - 2$. Show that there exists a unique $p = f_{\delta_\mu} + \sum_{i=0}^{\mu-1} a_i f_{\delta_i}$ (with $a_i \in \mathbb{R}$) such that

(1) $p(x_i) = 0$, $i = 1, 2, \ldots, \mu$,

(2) $p(x)$ changes sign at x_i if and only if $x_i \notin \{a, b, x_{i-1}, x_{i+1}\}$.

Show also that

(3) $p(x) \neq 0$ if $x \notin \{x_1, x_2, \ldots, x_\mu\}$,

(4) $a_i a_{i+1} \le 0$, $i = 0, 1, \ldots, \mu - 1$, $a_\mu := 1$,

(5) $p(x) > 0$, $x \in (x_\mu, b)$,

(6) $(-1)^\mu p(x) > 0$, $x \in (a, x_1)$,

(7) $(-1)^{\mu-i} p(x) > 0$, $x \in (x_i, x_{i+1})$, $i = 1, 2, \ldots, \mu - 1$.

Hint: Use part a] and a limiting argument. The uniqueness follows from E.10 of Section 3.1. □

The next exercise provides a solution to a problem of Lorentz, which is settled in Borosh, Chui, and Smith [77].

E.5 A Problem of Lorentz on Best Approximation to x^λ. Suppose that $[a, b] \subset [0, \infty)$, $n \in \mathbb{N}$, and $p \in (0, \infty]$ are fixed. Let μ be a finite Borel measure on $[a, b]$.

a] Suppose $\lambda_1, \lambda_2, \ldots, \lambda_n$ are arbitrary fixed real numbers if $a > 0$, or fixed real numbers greater than $-1/p$ if $a = 0$. Let $f \in L_p(\mu)$ be fixed. Show that

$$E_{p,\mu}(\lambda_1, \lambda_2, \ldots, \lambda_n; f) := \min_{a_i \in \mathbb{R}} \left\| f(x) - \sum_{i=1}^{n} a_i x^{\lambda_i} \right\|_{L_p(\mu)}$$

exists and is finite.

Outline. Use a standard compactness argument. □

b] Suppose $1 < p < \infty$, the support of μ contains at least $n+1$ points, $\lambda_1, \lambda_2, \dots, \lambda_n, \lambda$ are arbitrary fixed distinct real numbers if $a > 0$, or fixed real numbers greater than $-1/p$ if $a = 0$. Show that if $(\widetilde{a}_i)_{i=1}^n \subset \mathbb{R}$ satisfies

$$E_{p,\mu}(\lambda_1, \lambda_2, \dots, \lambda_n; x^\lambda) = \Big\| x^\lambda - \sum_{i=1}^n \widetilde{a}_i x^{\lambda_i} \Big\|_{L_p(\mu)},$$

then

$$f(x) := x^\lambda - \sum_{i=1}^n \widetilde{a}_i x^{\lambda_i}$$

has exactly n sign changes on (a, b).

Hint: Since $(x^{\lambda_1}, \dots, x^{\lambda_n}, x^\lambda)$ is a Chebyshev system, it is sufficient to prove that f has at least n sign changes on (a, b). Suppose f has at most $n-1$ sign changes on $[a, b]$. Then, since $(x^{\lambda_1}, \dots, x^{\lambda_n})$ is a Chebyshev system, by E.11 of Section 3.1 there exists an element

$$h \in \operatorname{span}\{x^{\lambda_1}, \dots, x^{\lambda_n}\}$$

such that

$$|f(x)|^{p-1} \operatorname{sign}(f(x)) h(x) \geq 0$$

on $[a, b]$ with strict inequality at all but n points (at every point where f does not vanish). Using that the support of μ contains at least $n+1$ distinct points, this implies

$$\int_a^b |f|^{p-1} \operatorname{sign}(f) h \, d\mu > 0,$$

which contradicts E.7 h] of Section 2.2. □

c] Suppose $p = \infty$, $\operatorname{supp}(\mu) = [a, b]$, $\lambda_1, \lambda_2, \dots, \lambda_n, \lambda$ are arbitrary fixed distinct real numbers if $a > 0$, or fixed distinct nonnegative real numbers if $a = 0$. Show that if $(\widetilde{a}_i)_{i=1}^n \subset \mathbb{R}$ satisfies

$$E_{\infty,\mu}(\lambda_1, \lambda_2, \dots, \lambda_n; x^\lambda) = \Big\| x^\lambda - \sum_{i=1}^n \widetilde{a}_i x^{\lambda_i} \Big\|_{L_\infty(\mu)},$$

then

$$f(x) := x^\lambda - \sum_{i=1}^n \widetilde{a}_i x^{\lambda_i}$$

has exactly n sign changes on (a, b).

Hint: Use Theorem 3.1.6. □

d] Best Approximation to x^λ from Certain Classes of Müntz Polynomials. Let $1 < p \le \infty$. Suppose the support of μ contains at least $n+1$ distinct points if $1 < p < \infty$ or $\operatorname{supp}(\mu) = [a,b]$ if $p = \infty$. Let $\lambda > \gamma_1 > \gamma_2 > \cdots$ be arbitrary fixed real numbers if $a > 0$, or fixed nonnegative real numbers if $a = 0$. Suppose we wish to minimize

$$E_{p,\mu}(\lambda_1, \lambda_2, \dots, \lambda_n; x^\lambda)$$

for all sets of n distinct real numbers $\lambda_1 > \lambda_2 > \cdots > \lambda_n$ satisfying

$$\{\lambda_1, \lambda_2, \dots, \lambda_n\} \subset \{\gamma_1, \gamma_2, \dots\}.$$

Show that the minimum occurs if and only if

$$\{\lambda_1, \lambda_2, \dots, \lambda_n\} = \{\gamma_1, \gamma_2, \dots, \gamma_n\}.$$

Hint: Let

$$E_{p,\mu}(\lambda_1, \lambda_2, \dots, \lambda_n; x^\lambda) = \left\| x^\lambda - \sum_{i=1}^{n} \widetilde{a}_i x^{\lambda_i} \right\|_{L_p(\mu)},$$

where $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ is a set of n distinct real numbers for which the minimum is taken; see part a]. By parts b] and c]

$$f(x) := x^\lambda - \sum_{i=1}^{n} \widetilde{a}_i x^{\lambda_i}$$

has exactly n sign changes $x_1, x_2, \dots, x_n$ on (a,b). Let

$$g \in \operatorname{span}\{x^{\gamma_1}, x^{\gamma_2}, \dots, x^{\gamma_n}\}$$

interpolate x^λ at the points $x_1, x_2, \dots, x_n$. Now use Theorem 3.2.5 to finish the proof. □

E.6 Strictly Totally Positive Kernels (Karlin [68]). A (continuous) function $K(s,t)$ is an *STP kernel* on $[a,b] \times [c,d]$ if

$$\begin{vmatrix} K(s_0,t_0) & \dots & K(s_0,t_n) \\ \vdots & \ddots & \vdots \\ K(s_n,t_0) & \dots & K(s_n,t_n) \end{vmatrix} > 0$$

for all $a \le s_0 < \cdots < s_n \le b$, $c \le t_0 < \cdots < t_n \le d$, and for all $n > 0$.

a] Observe that E.3 b] implies that

$$K(s,t) = \frac{1}{s+t} \quad \text{is STP on } [a,b] \times [a,b]\,, \quad a > 0\,.$$

Observe also that E.2 b] implies that

$$K(s,t) = e^{st} \quad \text{is STP on } (-\infty,\infty) \times (-\infty,\infty)\,.$$

b] Suppose K is STP on $[a,b] \times [c,d]$, and $(f_0, \dots, f_n)$ is a Chebyshev system on $[a,b]$. Show that if

$$v_i(x) = \int_a^b K(t,x) f_i(t)\, dt\,, \qquad i = 0, 1, \dots, n\,,$$

then $(v_0, \dots, v_n)$ is a Chebyshev system on $[c,d]$.

c] **Variation Diminishing Property.** Suppose K is STP on $[a,b] \times [c,d]$ and suppose $f \in C[a,b]$. Let

$$g(x) := \int_a^b K(t,x) f(t)\, dt\,.$$

Then g has no more sign changes on $[c,d]$ than f has on $[a,b]$.

d] The Laplace transform of a function $f \in C[0,\infty) \cap L_1[0,\infty)$

$$L(f)(x) := \int_0^\infty f(t) e^{-tx}\, dt$$

has no more sign changes on $[0,\infty)$ than f does.

Proof. This follows from parts a] and c]. It may also be proved directly by induction as follows. Suppose f has exactly n sign changes on $[0,\infty)$, one at x_0. Then $g(x) := (x_0 - x) f(x)$ has exactly $n-1$ sign changes on $[0,\infty)$. Now observe that

$$e^{-x_0 x} \frac{d}{dx} (e^{x_0 x} L(f)(x)) = L(g)(x)\,,$$

so $L(f)$ has at most one more sign change on $[0,\infty)$ than $L(g)$ does. □

e] Use part d] and E.2 a] to reprove that

$$\left(\frac{1}{x+\alpha_1}\,, \frac{1}{x+\alpha_2}\,, \dots \right)\,, \qquad -a < \alpha_1 < \alpha_2 < \cdots$$

satisfies Descartes' rule of signs on $[a,b]$. □

Proof. Observe that

$$\int_0^\infty e^{-\alpha_i t} e^{-tx}\, dt = \frac{1}{x+\alpha_i}$$

for every $x \in (-\alpha_i, \infty)$. □

E.7 Descartes' Rule of Signs for Polynomials.

a] Prove by induction that $\sum_{k=0}^{n} a_k x^k \in \mathcal{P}_n$ has no more zeros in $(0,\infty)$ (repeated zeros are counted according to their multiplicities) than the number of sign changes in $(a_0, a_1, \ldots, a_n)$.

b] Let $\alpha > 0$. Let $p(x) = \sum_{k=0}^{n} a_k x^k$ and $q(x) := (x-\alpha)p(x) = \sum_{k=0}^{n+1} b_k x^k$. Show that if the number of sign changes in $\{a_0, a_1, \ldots, a_n\}$ is m, then the number of sign changes in $\{b_0, b_1, \ldots, b_{n+1}\}$ is at least $m+1$.

c] Give another proof of a] based on b].

d] In part a] the number of sign changes in $(a_0, a_1, \ldots, a_n)$ exceeds the number of positive zeros by an even integer.

Hint: See Pólya and Szegő [76]. □

Refinements of the above exercise are presented in E.6 of Appendix 1, where Cauchy indices are discussed.

The first part of the following exercise is a version of a result from Zielke [79]:

E.8 The Effect of Differentiation on Weak Markov Systems. The system $(f_0, \ldots, f_n)$ is called a *weak Chebyshev system* on $[a,b]$ if $f_i \in C[a,b]$ for each i and every $f \in \operatorname{span}\{f_0, \ldots, f_n\}$ has at most n sign changes on $[a,b]$ (so the only difference between a Chebyshev system and a weak Chebyshev system is that in the definition of the latter, zeros without sign change are not counted).

Analogously, the system $(g_0, \ldots, g_n)$ is called a *weak Markov system* on $[a,b]$ if $g_i \in C[a,b]$ for each i and $(g_0, \ldots, g_m)$ is a weak Chebyshev system on $[a,b]$ for every $m = 0, 1, \ldots, n$ (so the only difference between a Markov system and a weak Markov system is that in the definition of the latter, zeros without sign change are not counted).

a] Suppose $(1, f_1, \ldots, f_n)$ is a weak Markov system of C^1 functions on $[a,b]$. Show that $(f_1', \ldots, f_n')$ is a weak Markov system on $[a,b]$.

Outline. Proceed by induction on n. If $n = 0$, then the statement is obvious. Suppose that the statement is true for $n-1$. By the inductive hypothesis $(f_1', \ldots, f_{n-1}')$ is a weak Markov system on $[a,b]$, hence Rolle's theorem implies that $(1, f_1, \ldots, f_{n-1})$ is a Markov system on $[a,b]$.

Suppose that $g \in \operatorname{span}\{f_1', \ldots, f_{n-1}'\}$ is of the form

$$g := \sum_{i=1}^{n} a_i f_i', \qquad a_i \in \mathbb{R}$$

and g has at least n sign changes on $[a,b]$. Then there exist $n+2$ distinct

points

$$a < x_1 < x_2 < \cdots < x_{n+2} < b$$

and $\epsilon = \pm 1$ such that

$$F := \epsilon \sum_{i=1}^{n} a_i f_i$$

satisfies

$$(-1)^i (F(x_{i+1}) - F(x_i)) > 0\,, \qquad i = 1, 2, \ldots, n+1\,.$$

Since $(1, f_1, \ldots, f_{n-1})$ is a Markov system on $[a, b]$, by Proposition 3.1.2, there exist functions

$$G_\delta \in \operatorname{span}\{1, f_1, \ldots, f_{n-1}\}$$

such that

$$G_\delta(x_i) = F(x_i) + \delta(-1)^i\,, \qquad i = 2, 3, \ldots, n+1$$

for every $\delta > 0$. Then by the inductive hypothesis, G'_δ has at most $n-2$ sign changes on $[a, b]$. It follows that if $\delta > 0$ is sufficiently small, then

$$(-1)(G_\delta(x_2) - G_\delta(x_1)) < 0$$

and

$$(-1)^{n+1}(G_\delta(x_{n+2}) - G_\delta(x_{n+1})) < 0\,,$$

otherwise G'_δ would have at least n sign changes on $[a, b]$. Now show that for sufficiently small $\delta > 0$

$$F - G_\delta \in \operatorname{span}\{1, f_1, \ldots, f_n\}$$

has at least $n+1$ sign changes, which is a contradiction. □

b] Suppose that $(1, f_1, \ldots, f_n)$ is an ECT system on $[a, b]$ and suppose that each $f_i \in C^n[a, b]$ (see E.3 of Section 3.1). Show that $(f'_1, \ldots, f'_n)$ is also an ECT system on $[a, b]$ with each $f'_i \in C^{n-1}[a, b]$.

Hint: Use the definition given in E.3 of Section 3.1. □

c] Suppose $(1, f_1, \ldots, f_n)$ is a weak Markov system on $[a, b]$ with each $f_i \in C^1[a, b]$. Show that $(1, f_1, \ldots, f_n)$ is a Markov system on $[a, b]$.

Hint: Use Rolle's theorem and part a]. □

3.3 Chebyshev Polynomials in Chebyshev Spaces

Suppose

$$H_n := \operatorname{span}\{f_0, f_1, \dots, f_n\}$$

is a Chebyshev space on $[a,b]$, and A is a compact subset of $[a,b]$ with at least $n+1$ points. We can define the *generalized Chebyshev polynomial*

$$T_n := T_n\{f_0, f_1, \dots, f_n; A\}$$

for H_n on A by the following three properties:

$$T_n \in \operatorname{span}\{f_0, f_1, \dots, f_n\} \tag{3.3.1}$$

there exists an alternation sequence $(x_0 < x_1 < \cdots < x_n)$ for T_n on A, that is,

$$\operatorname{sign}(T_n(x_{i+1})) = -\operatorname{sign}(T_n(x_i)) = \pm\|T_n\|_A \tag{3.3.2}$$

for $i = 0, 1, \dots, n-1$, and

$$\|T_n\|_A = 1 \quad \text{with} \quad T_n(\max A) > 0\,. \tag{3.3.3}$$

Of course the existence and uniqueness of such a T_n has to be proved. Note that if together with $\operatorname{span}\{f_0, \dots, f_n\}$, $\operatorname{span}\{f_0, \dots, f_{n-1}\}$ is also a Chebyshev space, then Theorem 3.1.6 implies that

$$T_n = c\left(f_n - \sum_{k=0}^{n-1} a_k f_k\right),$$

where the numbers $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ are chosen to minimize

$$\left\| f_n - \sum_{k=0}^{n-1} a_k f_k \right\|_A, \tag{3.3.4}$$

satisfies properties (3.3.1) and (3.3.2), and the normalization constant $c \in \mathbb{R}$ can be chosen so that T_n satisfies property (3.3.3) as well. In E.1 we outline the proof of the existence and uniqueness of a T_n satisfying properties (3.3.1) to (3.3.3) without assuming that $\operatorname{span}\{f_0, \dots, f_{n-1}\}$ is a Chebyshev space.

Note that if $(f_0, \dots, f_n)$ is a Descartes system on $[a,b]$, then the normalization constant (that is, the lead coefficient) c in T_n is positive. This follows from E.4 of Section 3.2.

On intervals, with $f_i(x) := x^i$, the definition (3.3.1) to (3.3.3) gives the usual Chebyshev polynomials; see E.7 of Section 2.1.

The Chebyshev polynomials T_n for H_n on A encode much of the information of how the space H_n behaves with respect to the uniform norm on A. Many extremal problems are solved by the Chebyshev polynomials.

When $(f_0, f_1, \dots)$ is a Markov system on $[a,b]$ we can introduce the sequence $(T_n)_{n=0}^{\infty}$ of *associated Chebyshev polynomials*

$$T_n := T_n\{f_0, f_1, \dots, f_n; [a,b]\}$$

for H_n on $[a,b]$. Then $(T_0, T_1, \dots)$ is a Markov system on $[a,b]$ again with the same span. (One reason for not always choosing this as a canonical basis is that it is never a Descartes system.)

The denseness of Markov spaces in $C[a,b]$ is intimately tied to the location of the zeros of the associated Chebyshev polynomials; see Section 4.1.

An example of an extremal problem solved by the Chebyshev polynomials is the following:

Theorem 3.3.1. *Suppose* $H_n := \operatorname{span}\{f_0, \dots, f_n\}$ *is a Chebyshev space on* $[a,b]$ *with associated Chebyshev polynomial*

$$T_n := T_n\{f_0, f_1, \dots, f_n; [a,b]\}$$

and each f_i *is differentiable at* b. *Then*

$$\max\{|p'(b)| : p \in H_n\,,\ \ \|p\|_{[a,b]} \le 1\,,\ \ p(b) = T_n(b)\}$$

is attained by T_n.

Proof. Suppose $p \in H_n$, $\|p\|_{[a,b]} \le 1$, and $p(b) = T_n(b)$. We need to show that $|p'(b)| \le |T_n'(b)|$. Let $a \le \zeta_0 < \zeta_1 < \cdots < \zeta_n \le b$ be the points of alternation for T_n, that is,

$$T_n(\zeta_i) = \pm(-1)^i\,, \qquad i = 0, 1, \dots, n\,.$$

Note that $T_n - p$ has at least n zeros in $[\zeta_0, \zeta_n]$, one in each $[\zeta_{i-1}, \zeta_i]$, $i = 1, 2, \dots, n$ (we count each internal zero without sign change twice, as in E.10 of Section 3.1). So if $b \ne \zeta_n$, then $T_n - p$ has $n+1$ zeros on $[a,b]$ including the zero at b, hence $p = T_n$, and the proof is finished. We may thus assume that $T_n(b) = 1$. Assume that $|p'(b)| > |T_n'(b)|$. Since $T_n'(b) \ge 0$, without loss of generality we may assume that $p'(b) > T_n'(b)$, otherwise we study $-p$. Then $T_n - p$ has two zeros on $[\zeta_n, b]$, and hence has $n+1$ zeros in $[a,b]$ (again, we count each internal zero without sign change twice). Thus by E.10 of Section 3.1 we have $p = T_n$, which contradicts the assumption $p'(b) > T_n'(b)$. □

An extension of the above theorem to interior points is considered in E.3.

Theorem 3.3.2. *Suppose* $(f_0, \dots, f_{n-1}, g)$ *and* $(f_0, \dots, f_{n-1}, h)$ *are both Chebyshev systems on* $[a,b]$ *with associated Chebyshev polynomials*

$$T_n := T_n\{f_0, f_1, \dots, f_{n-1}, g; [a,b]\}$$

and

$$S_n := T_n\{f_0, f_1, \dots, f_{n-1}, h; [a,b]\},$$

respectively. Suppose $(f_0, f_1, \dots, f_{n-1}, g, h)$ *is also a Chebyshev system. Then the zeros of* T_n *and* S_n *interlace (there is exactly one zero of* S_n *between any two consecutive zeros of* T_n*).*

Proof. Since $(f_0, \dots, f_{n-1}, g, h)$ is a Chebyshev system on $[a,b]$, $T_n \pm S_n$ has at most $n+1$ zeros. However, between any two consecutive alternation points of T_n, of which there are $n+1$, there is a zero of $T_n \pm S_n$ (which may be at an internal alternation point of T_n only if it is a zero without sign change, which is then counted twice). Likewise, there is a zero of $T_n \pm S_n$ between any two consecutive alternation points of S_n. Thus between any three successive alternation points of say T_n there can be at most three zeros of $T_n \pm S_n$. However, if S_n had two zeros between two consecutive zeros of T_n, then there would be three consecutive alternation points of either T_n or S_n with at least four zeros of either $T_n + S_n$ or $T_n - S_n$ between them, which is impossible. □

Theorem 3.3.3. *Suppose* $(f_0, f_1, \dots)$ *is a Markov system on* $[a,b]$ *with associated Chebyshev polynomials*

$$T_n := T_n\{f_0, f_1, \dots, f_n; [a,b]\}.$$

Then the zeros of T_n *and* T_{n-1} *strictly interlace (there is exactly one zero of* T_{n-1} *strictly between any two consecutive zeros of* T_n*).*

Proof. The proof is analogous to that of Theorem 3.3.2. □

Theorem 3.3.4 (Lexicographic Property). *Let* $(f_0, f_1, \dots)$ *be a Descartes system on* $[a,b]$*. Suppose* $\lambda_0 < \lambda_1 < \cdots < \lambda_n$ *and* $\gamma_0 < \gamma_1 < \cdots < \gamma_n$ *are nonnegative integers satisfying*

$$\lambda_i \le \gamma_i, \qquad i = 0, 1, \dots, n.$$

Let

$$T_n := T_n\{f_{\lambda_0}, f_{\lambda_1}, \dots, f_{\lambda_n}; [a,b]\}$$

and

$$S_n := T_n\{f_{\gamma_0}, f_{\gamma_0}, \dots, f_{\gamma_n}; [a,b]\}$$

denote the associated Chebyshev polynomials.

Let

$$\alpha_1 < \alpha_2 < \cdots < \alpha_n \quad \textit{and} \quad \beta_1 < \beta_2 < \cdots < \beta_n$$

denote the zeros of T_n *and* S_n*, respectively. Then*

$$\alpha_i \leq \beta_i\,, \qquad i = 1, 2, \ldots, n$$

with strict inequality if $\lambda_i \neq \gamma_i$ *for at least one index* i*. (In other words, the zeros of* T_n *lie to the left of the zeros of* S_n*.)*

Proof. It is clearly sufficient to prove the theorem in the case that $\lambda_i = \gamma_i$ for each $i \neq m$, and $\lambda_m < \gamma_m < \lambda_{m+1}$ for a fixed index m, and then to proceed by a sequence of pairwise comparisons (if $m = n$, then λ_{m+1} is meant to be replaced by ∞). So suppose

$$T_n := T_n\{f_{\lambda_0}, \ldots, f_{\lambda_m}, \ldots, f_{\lambda_n}; [a,b]\} = \sum_{i=0}^{n} c_i f_{\lambda_i}$$

and

$$S_n := T_n\{f_{\lambda_0}, \ldots, f_{\gamma_m}, \ldots, f_{\lambda_n}; [a,b]\} = d_m f_{\gamma_m} + \sum_{\substack{i=0 \\ i \neq m}}^{n} d_i f_{\lambda_i}$$

with $\lambda_m < \gamma_m < \lambda_{m+1}$. Then by Theorem 3.3.2 the zeros of S_n and T_n interlace and all that remains to prove is that the largest zero of S_n is larger than the largest zero of T_n. That is, we must show that $\alpha_n < \beta_n$. For this we argue as follows. It follows from Theorem 3.2.5 that the lead coefficient of T_n is less than the lead coefficient of S_n ($c_n < d_n$ provided $m < n$). Since both T_n and S_n have an alternation sequence of length $n+1$ on $[a,b]$, and since

$$\|T_n\|_{[a,b]} = \|S_n\|_{[a,b]} = 1\,, \qquad T_n(b) > 0\,, \quad S_n(b) > 0\,,$$

it follows from E.1 b] that

$$S_n - T_n \in \operatorname{span}\{f_{\lambda_0}, f_{\lambda_1} \ldots, f_{\lambda_n}, f_{\gamma_m}\}$$

has $n+1$ zeros $x_1 \leq x_2 \leq \cdots \leq x_{n+1}$ on $[a,b]$. Therefore, it follows from E.4 c] of Section 3.2 that $(S_n - T_n)(x) > 0$ on (x_{n+1}, b) and $(S_n - T_n)(x) < 0$ on (x_n, x_{n+1}). Hence the assumption $\beta_n \leq \alpha_n$ would imply that $S_n - T_n$ has at least $n+2$ zeros on $[a,b]$ (counting each internal zero without sign change twice), which is a contradiction. (Draw a picture and use the alternation characterization of the Chebyshev polynomials T_n and S_n to make the proof of the above statement transparent.) So $\beta_n > \alpha_n$, indeed, and the proof is finished. □

Comments, Exercises, and Examples.

If $H_n := \text{span}(f_0, \dots, f_n)$ is a Chebyshev space on $[a,b]$, A is a compact subset of $[a,b]$, and $p \in (0,\infty)$, then

$$T_n = c\left(f_n - \sum_{k=0}^{n-1} a_k f_k\right)$$

with $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ minimizing

$$\int_A \left| f_n - \sum_{k=0}^{n-1} a_k f_k \right|^p$$

is called an L_p Chebyshev polynomial for H_n on A. When $A = [a,b]$ and $p \in (1,\infty]$, the properties of the zeros of these L_p Chebyshev polynomials are explored in Pinkus and Ziegler [79], where much of the material of this section may be found. For example, an L_p analog of Theorem 3.3.2 still holds.

E.1 Existence and Uniqueness of Chebyshev Polynomials. Let $A \subset [a,b]$ be a compact set containing at least $n+1$ points. Let $(f_0, \dots, f_n)$ be a Chebyshev system on $[a,b]$.

a] Existence of Chebyshev Polynomials. Show that there exists a T_n satisfying properties (3.3.1) to (3.3.3).

Hint: If A contains exactly $n+1$ points, then the existence of T_n is just the interpolation property of a Chebyshev space formulated in Proposition 3.1.2 b]. So assume that A contains at least $n+2$ points. Then there is a $\delta > 0$ so that $A \cap [a,c]$ contains at least $n+1$ points for every $c \in (b-\delta, b)$. Show that for every $c \in (b-\delta, b)$, there is a $g_c \in \text{span}\{f_0, \dots, f_n\}$ for which

$$\sup\{f(b) : f \in \text{span}\{f_0, f_1, \dots, f_n\},\ \|f\|_{[a,c]} = 1\}$$

is attained. Use a variational method to show that g_c satisfies properties (3.3.1) to (3.3.3) with A replaced by $A \cap [a,c]$.

Now let $(c_k)_{k=1}^{\infty}$ be a sequence of numbers from $(b-\delta, b)$ that converges to b. Let $g_{c_k} \in \text{span}\{f_0, \dots, f_n\}$ satisfy properties (3.3.1) to (3.3.3) with A replaced by $A \cap [a,c_k]$. Show that there is a subsequence of $(g_{c_k})_{k=1}^{\infty}$ that converges to a $g \in \text{span}\{f_0, \dots, f_n\}$ uniformly on $[a,b]$. Show that $T_n := g$ satisfies properties (3.3.1) to (3.3.3). □

b] A Lemma for Part c]. Suppose $f, g \in C[a,b]$ with $\|f\|_A = \|g\|_A > 0$ and there are alternation sequences

$$(x_1 < x_2 < \cdots < x_{n+1}) \quad \text{and} \quad (y_1 < y_2 < \cdots < y_{n+1})$$

for f and g, respectively, on A.

Suppose also that

$$\operatorname{sign}(f(x_1)) = \operatorname{sign}(g(y_1)).$$

Show that $f - g$ has at least $n+1$ zeros on $[a,b]$.

c] **Uniqueness of Chebyshev Polynomials.** Show that the Chebyshev polynomials

$$T_n\{f_0, f_1, \ldots, f_n; A\}$$

satisfying properties (3.3.1) to (3.3.3) are unique.

Hint: Use part a] and E.10 of Section 3.1. □

E.2 More on Chebyshev Polynomials. Let $H_n := \operatorname{span}\{f_0, \ldots, f_n\}$ be a Chebyshev space on $[a,b]$ with associated Chebyshev polynomial denoted by $T_n := T_n\{f_0, \ldots, f_n; [a,b]\}$. Show the following statements.

a] If $1 \in H_n$, then $|T_n(a)| = |T_n(b)| = 1$.

b] If $1 \in H_n$, then T_n is monotone between two successive points of its alternation sequence.

Note that the conclusions of parts a] and b] do not necessarily hold in general.

c] If $T_n =: \sum_{i=0}^{n} a_i f_i$, $a_i \in \mathbb{R}$, then the coefficient sequence of T_n/a_m solves

$$\min_{b_i \in \mathbb{R}} \Big\| f_m + \sum_{\substack{i=0 \\ i \neq m}}^{n} b_i f_i \Big\|_{[a,b]}$$

uniquely, provided that $\{f_0, \ldots f_{m-1}, f_{m+1}, \ldots f_n\}$ is also a Chebyshev system on $[a,b]$. (So this applies to ordinary polynomials on $[0,1]$ but not on $[-1,1]$.)

d] Suppose $(f_0, \ldots, f_n)$ is a Descartes system on $[a,b]$ with associated Chebyshev polynomial $T_n := T_n\{f_0, \ldots, f_n; [a,b]\} =: \sum_{i=0}^{n} a_i f_i$, $a_i \in \mathbb{R}$. Show that $a_n > 0$ and $a_i a_{i+1} < 0$ for each $i = 0, 1, \ldots, n-1$.

Hint: Use E.4 a] of Section 3.2. □

E.3 Extension of Theorem 3.3.1. *Let $H_n := \operatorname{span}\{f_0, \ldots, f_n\}$ be a Chebyshev space on $[a,b]$ with associated Chebyshev polynomial*

$$T_n := T_n\{f_0, f_1, \ldots, f_n; [a,b]\}$$

and suppose each f_i is differentiable at $x_0 \in [a,b]$.

a] *If $T_n'(x_0) > 0$, then*

$$\max\{p'(x_0) : p \in H_n,\ \|p\|_{[a,b]} \le 1,\ p(x_0) = T_n(x_0)\}$$

is attained only by T_n.

b] *If* $T_n'(x_0) < 0$, *then*

$$\min\{p'(x_0) : p \in H_n\,,\ \ \|p\|_{[a,b]} \le 1\,,\ \ p(x_0) = T_n(x_0)\}$$

is attained and only by T_n.

Hint: Consider the number of zeros of $T_n - p$. □

E.4 More Lexicographic Properties of Müntz Spaces. Let $[a,b] \subset [0,\infty)$. Suppose

$$\lambda_0 < \lambda_1 < \cdots < \lambda_n \quad \text{and} \quad \gamma_0 < \gamma_1 < \cdots < \gamma_n$$

are arbitrary real numbers if $a > 0$, or arbitrary nonnegative numbers if $a = 0$. Suppose $\lambda_i \le \gamma_i$ for each i with strict inequality for at least one index i. Let

$$H_n := \operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \dots, x^{\lambda_n}\} \quad \text{and} \quad G_n := \operatorname{span}\{x^{\gamma_0}, x^{\gamma_1}, \dots, x^{\gamma_n}\}.$$

Denote the associated Chebyshev polynomials for H_n and G_n on $[a,b]$ by

$$T_{n,\lambda} := T_n\{x^{\lambda_0}, x^{\lambda_1}, \dots, x^{\lambda_n}; [a,b]\}$$

and

$$T_{n,\gamma} := T_n\{x^{\gamma_0}, x^{\gamma_1}, \dots, x^{\gamma_n}; [a,b]\},$$

respectively.

a] Show that $\lambda_n \ge 0$ implies $T_{n,\lambda}(b) = 1$.

Hint: $T_{n,\lambda}(b) \ne 1$ would imply that $T_{n,\lambda}'$ has at least $n+1$ distinct zeros in $[a,\infty)$ if $\lambda_0 > 0$ and at least n distinct zeros if $\lambda_0 = 0$. □

b] Let $x_0 = a$ or $x_0 = b$. If $x_0 = a = 0$, then assume that $\lambda_0 = 0$ and $\lambda_1 = 1$. Show that

$$\max\{|p'(x_0)| : p \in H_n\,,\ \ \|p\|_{[a,b]} \le 1\}$$

is attained uniquely by $\pm T_{n,\lambda}$.

c] Let $x_0 \in [0,\infty) \setminus [a,b]$. If $x_0 = 0$, then assume that $\lambda_0 = 0$. Show that

$$\max\{|p(x_0)| : p \in H_n\,,\ \ \|p\|_{[a,b]} \le 1\}$$

is attained uniquely by $\pm T_{n,\lambda}$.

Hint for b] and c]: First prove that an extremal $p^* \in H_n$ exists. Then show, by a variational method, that p^* equioscillates $n+1$ times between ± 1 on $[a,b]$. □

d] Let $\lambda_n \ge 0$, $\gamma_n \ge 0$, and $a > 0$. Show that

$$|T_{n,\lambda}'(b)| < |T_{n,\gamma}'(b)|.$$

Show also that if $a > 0$ and there exists an index k, $0 \le k \le n$, such that $\lambda_k = \gamma_k = 0$, then

$$|T'_{n,\lambda}(a)| > |T'_{n,\gamma}(a)|\,.$$

e] Let $\lambda_n \ge 0$, $\gamma_n \ge 0$, and $a > 0$. Show that

$$|T_{n,\lambda}(x_0)| < |T_{n,\gamma}(x_0)|\,, \qquad x_0 \in (b, \infty)\,.$$

Show also that if $\lambda_0 \ge 0, \gamma_0 \ge 0$, and $a > 0$, then

$$|T_{n,\lambda}(x_0)| > |T_{n,\gamma}(x_0)|\,, \qquad x_0 \in [0, a)$$

(when $x_0 = 0$, we need the assumption $\lambda_0 = \gamma_0 = 0$).

Hint for d] and e]: Suppose to the contrary that one of the inequalities of parts d] and e] fails. Assume, without loss of generality, that there is an index m such that $\lambda_i = \gamma_i$ whenever $i \ne m$, and $\lambda_m < \gamma_m$. Note that by a], $\lambda_n \ge 0$ and $\gamma_n \ge 0$ imply $T_{n,\lambda}(b) = T_{n,\gamma}(b) = 1$. Also, by E.1 a], $\lambda_k = \gamma_k = 0$ implies $T_{n,\lambda}(a) = T_{n,\gamma}(a) = (-1)^n$. Now use Theorem 3.3.4 to show that

$$T_{n,\lambda} - T_{n,\gamma} \in \operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}, x^{\gamma_m}\}$$

has at least $n+2$ zeros in $(0, \infty)$ (in the cases when $\lambda_n \ge 0$ and $\gamma_n \ge 0$ are assumed) or in $[0, \infty)$ (in the cases when $\lambda_k = \gamma_k = 0$ is assumed). This contradiction finishes the proof. □

f] Let $0 < a < b$. Show that if $\lambda_n \ge 0$ and $\gamma_n \ge 0$, then

$$\max_{p \in H_n} \frac{|p'(b)|}{\|p\|_{[a,b]}} < \max_{q \in G_n} \frac{|q'(b)|}{\|q\|_{[a,b]}}\,.$$

Show also that if there exists an index k, $0 \le k \le n$, such that $\lambda_k = \gamma_k = 0$, then

$$\max_{p \in H_n} \frac{|p'(a)|}{\|p\|_{[a,b]}} > \max_{q \in G_n} \frac{|q'(a)|}{\|q\|_{[a,b]}}\,.$$

Hint: Combine parts b] and d]. □

g] Let $0 < a < b$. Let $\lambda_n \ge 0$ and $\gamma_n \ge 0$. Show that

$$\max_{p \in H_n} \frac{|p(x_0)|}{\|p\|_{[a,b]}} < \max_{q \in G_n} \frac{|q(x_0)|}{\|q\|_{[a,b]}}\,, \qquad x_0 \in (b, \infty)\,.$$

Show also that

$$\max_{p \in H_n} \frac{|p(x_0)|}{\|p\|_{[a,b]}} > \max_{q \in G_n} \frac{|q(x_0)|}{\|q\|_{[a,b]}}\,, \qquad x_0 \in [0, a)$$

(when $x_0 = 0$, we need the assumption $\lambda_0 = \gamma_0 = 0$).

Hint: Combine parts c] and e]. □

h] Extend the validity, with $<$ and $>$ replaced by $\leq$ and $\geq$, respectively, of the inequalities of parts f] and g] to the case when the interval $[a,b] \subset (0,\infty)$ is replaced by $[0,b]$.

i] Suppose there exists an index k, $0 \leq k \leq n$, such that $\lambda_k = \gamma_k = 0$. Let $x_0 \in (0,a)$. Show that the second inequalities of parts e] and g] hold true.

Hint: Modify the arguments given in the hints to parts d], e], f], and g]. Note that $1 \in H_n \cap G_n$ ensures, as in E.1 a], that

$$T_{n,\lambda}(b) = T_{n,\gamma}(b) = 1 \qquad \text{and} \qquad T_{n,\lambda}(a) = T_{n,\gamma}(a) = (-1)^n .$$

□

E.5 Lexicographic Properties of $(\sinh \lambda_0 t, \ldots, \sinh \lambda_n t)$**.** Let

$$0 < \lambda_0 < \lambda_1 < \cdots < \lambda_n \quad \text{and} \quad 0 < \gamma_0 < \gamma_1 < \cdots < \gamma_n .$$

Suppose $\lambda_i \leq \gamma_i$ for each i. Let

$$H_n := \operatorname{span}\{\sinh \lambda_0 t, \sinh \lambda_1 t, \ldots, \sinh \lambda_n t\}$$

and

$$G_n := \operatorname{span}\{\sinh \gamma_0 t, \sinh \gamma_1 t, \ldots, \sinh \gamma_n t\} .$$

Denote the associated Chebyshev polynomials for H_n and G_n on $[0,1]$ by

$$T_{n,\lambda} := T_n\{\sinh \lambda_0 t, \sinh \lambda_1 t, \ldots, \sinh \lambda_n t; [0,1]\}$$

and

$$T_{n,\gamma} := T_n\{\sinh \gamma_0 t, \sinh \gamma_1 t, \ldots, \sinh \gamma_n t; [0,1]\} ,$$

respectively.

a] Let

$$\alpha_1 < \alpha_2 < \ldots < \alpha_n \quad \text{and} \quad \beta_1 < \beta_2 < \cdots < \beta_n$$

denote the zeros of $T_{n,\lambda}$ and $T_{n,\gamma}$, respectively. Show that

$$\alpha_i \leq \beta_i, \qquad i = 1, 2, \ldots, n$$

(in other words, the zeros of $T_{n,\lambda}$ lie to the left of the zeros of $T_{n,\gamma}$).

Outline. By E.2 d] of Section 3.2, $(\sinh \lambda_0 t, \ldots, \sinh \lambda_n t)$ is a Descartes system on $(0,\infty)$. Hence, by Theorem 3.3.3, the zeros of

$$T_{n,\lambda,\delta} := T_n\{\sinh \lambda_0 t, \sinh \lambda_1 t, \ldots, \sinh \lambda_n t; [\delta,1]\}$$

on $[\delta,1]$ lie to the left of the zeros of

$$T_{n,\gamma,\delta} := T_n\{\sinh\gamma_0 t, \sinh\gamma_1 t, \ldots, \sinh\gamma_n t; [\delta,1]\}$$

on $[\delta,1]$ for every $\delta \in (0,1)$. Show that

$$\lim_{\delta\to 0} \|T_{n,\lambda} - T_{n,\lambda,\delta}\| = \lim_{\delta\to 0} \|T_{n,\gamma} - T_{n,\gamma,\delta}\| = 0\,;$$

hence the desired result follows by a continuity argument. □

b] Show that

$$\max\{|p'(0)| : p \in H_n\,,\ \ \|p\|_{[0,1]} \le 1\}$$

is attained uniquely by $\pm T_{n,\lambda}$.

c] Show that

$$T_{n,\lambda}(1) = T_{n,\gamma}(1) = 1\,.$$

Hint: $T_{n,\lambda}(1) \neq 1$ would imply that

$$T'_{n,\lambda} \in \operatorname{span}\{\cosh\lambda_0 t, \cosh\lambda_1 t, \ldots, \cosh\lambda_n t\}$$

has at least $n+1$ distinct zeros in $(0,\infty)$. This is impossible, since by E.2 e] of Section 3.2, $(\cosh\lambda_0 t, \ldots, \cosh\lambda_n t)$ is a Descartes (hence Chebyshev) system on $(0,\infty)$. □

d] Show that

$$|T'_{n,\lambda}(0)| \ge |T'_{n,\gamma}(0)|\,.$$

Hint: Suppose to the contrary that the above inequality fails to hold. Assume, without loss of generality, that there is an index m such that $\lambda_i = \gamma_i$ whenever $i \neq m$, and $\lambda_m < \gamma_m$. Obviously

$$T_{n,\lambda}(0) = T_{n,\gamma}(0) = 0\,.$$

Part c] implies that $T_{n,\lambda}(1) = T_{n,\gamma}(1) = 1$. Now use part a] and the above observation to show that

$$T_{n,\lambda} - T_{n,\gamma} \in \operatorname{span}\{\sinh\lambda_0 t, \sinh\lambda_1 t, \ldots, \sinh\lambda_n t, \sinh\gamma_m t\}$$

has at least $n+2$ zeros in $(0,\infty)$. This contradicts E.2 d] of Section 3.2. □

e] Show that

$$\max_{0\neq p\in H_n} \frac{|p'(0)|}{\|p\|_{[0,1]}} \ge \max_{0\neq q\in G_n} \frac{|q'(0)|}{\|q\|_{[0,1]}}\,.$$

Hint: Combine parts b] and d]. □

The result of the following exercise has been observed independently by Lubinsky and Ziegler [90] and Kroó and Szabados [94]. Various coefficient estimates for polynomials are discussed in Milovanović, Mitrinović, and Rassias [94]. An estimate for the coefficients of polynomials having a given number of terms is obtained in Baishanski and Bojanic [80]. Approximation by such polynomials is studied in Baishanski [83].

E.6 Coefficient Bounds for Polynomials in a Special Basis. Show that

$$|c_m| \le 2^{-n} \binom{2n}{2m} \|p\|_{[-1,1]}, \qquad m = 0, 1, \dots, n$$

for every polynomial p of the form

$$p(x) = \sum_{m=0}^{n} c_m (1-x)^m (x+1)^{n-m}, \qquad c_m \in \mathbb{R}.$$

Outline. By E.5 a] and b] of Section 2.3

$$T_n(x) = \sum_{m=0}^{n} d_{m,n} (1-x)^m (x+1)^{n-m},$$

where

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

If $|c_m| > d_{m,n}\|p\|_{[-1,1]}$ for some index m, then the polynomial

$$q(x) = \frac{T_n(x)}{d_{m,n}} - \frac{p(x)}{c_m}$$

has at least n distinct zeros in $(-1,1)$. However,

$$q(x) = \sum_{\substack{j=0 \\ j\neq m}}^{n} a_j (1-x)^j (x+1)^{n-j} = (1-x)^n \sum_{\substack{j=0 \\ j\neq m}}^{n} a_j \left(\frac{1+x}{1-x}\right)^{n-j}$$

can have at most $n-1$ distinct zeros in $(-1,1)$ since

$$(u^0, u^1, \dots u^{n-m-1}, u^{n-m+1}, u^{n-m+2}, \dots, u^n)$$

is a Chebyshev system on $(0,\infty)$ by E.1 a] of Section 3.1. □

E.7 On the Zeros of the Chebyshev Polynomials for Müntz Spaces. Let $0 =: \lambda_0 < \lambda_1 < \cdots < \lambda_n$, and let

$$H_n := \operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \dots, x^{\lambda_n}\}.$$

Denote the associated Chebyshev polynomials for H_n on $[0,1]$ by

$$T_n := T_n\{x^{\lambda_0}, x^{\lambda_1}, \dots, x^{\lambda_n}; [0,1]\}.$$

a] Let $\epsilon \in \left(0, \frac{1}{2}\right)$. Suppose $\alpha < \beta$ are two consecutive zeros of T_n lying in $[\epsilon, 1-\epsilon]$. Show that

$$\frac{\epsilon^2}{n} \le \beta - \alpha \,.$$

Hint: It is clear that $Q_n(x) := x(1-x)T_n(x)$ has two consecutive zeros $\gamma < \delta$ in $[\alpha, 1]$ such that

$$\delta - \gamma \ge \frac{\epsilon}{n} \,.$$

Show that if

$$\beta - \alpha < \frac{\epsilon^2}{n} \,,$$

then

$$R_n(x) := T_n(x) - T_n\left(\tfrac{\gamma x}{\alpha}\right) \in H_n$$

has at least $n+1$ zeros on $[0, \alpha/\gamma]$, which is a contradiction. □

b] Denote the zeros of T_n in $(0,1)$ by $x_1 < x_2 < \cdots < x_n$. Show that

$$\log x_{k+1} - \log x_k \le \log x_k - \log x_{k-1} \,, \qquad k = 2, 3, \ldots, n-1 \,.$$

Hint: Use a zero counting argument, as in the hint to part a]. □

3.4 Müntz-Legendre Polynomials

We examine in some detail the system

$$(x^{\lambda_0}, x^{\lambda_1}, \ldots)$$

on $[0,1]$ which we call a *Müntz system*. In particular, we explicitly construct orthogonal "polynomials" for this system. This allows us to derive various extremal properties of these systems and leads to a very simple proof of the classical Müntz-Szász theorem in Section 4.2.

We adopt the following definition for x^λ:

$$x^\lambda = e^{\lambda \log x} \,, \quad x \in (0, \infty) \,, \quad \lambda \in \mathbb{C} \tag{3.4.1}$$

with value at 0 defined to be the limit of x^λ as $x \to 0$ from $(0, \infty)$ whenever the limit exists.

Given a sequence $\Lambda := (\lambda_i)_{i=0}^{\infty}$ of complex numbers, an element of $\operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}\}$ is called a *Müntz polynomial* or a *Λ-polynomial*. We denote the set of all such polynomials by $M_n(\Lambda)$, that is,

$$M_n(\Lambda) := \operatorname{span}\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}\} \,, \tag{3.4.2}$$

where the linear span is over $\mathbb{R}$ or $\mathbb{C}$ according to context. Let

$$M(\Lambda) := \bigcup_{n=0}^{\infty} M_n(\Lambda) = \text{ span}\{x^{\lambda_0}, x^{\lambda_1}, \dots\}. \tag{3.4.3}$$

For the $L_2[0,1]$ theory of Müntz systems, we consider

$$\Lambda = (\lambda_i)_{i=0}^{\infty}, \quad \text{Re}(\lambda_i) > -1/2, \quad \text{and} \quad \lambda_i \neq \lambda_j, \ i \neq j, \tag{3.4.4}$$

where $\text{Re}(\lambda)$ denotes the real part of λ. This ensures that the Λ-polynomials

$$p(x) = \sum_{k=0}^{n} a_k x^{\lambda_k}, \qquad a_k \in \mathbb{C}$$

are in $L_2[0,1]$. We can then define the orthogonal Λ-polynomials with respect to Lebesgue measure. We call these *Müntz-Legendre polynomials*. Although we often assume (3.4.4), the following definition requires neither the distinctness of the numbers λ_i nor the assumption $\text{Re}(\lambda_i) > -1/2$.

Definition 3.4.1 (Müntz-Legendre Polynomials). Let $\Lambda := (\lambda_i)_{i=0}^{\infty}$ be a sequence of complex numbers. We define the nth *Müntz-Legendre polynomial* on $(0,\infty)$ by

$$\begin{aligned} L_n(x) &:= L_n\{\lambda_0, \dots, \lambda_n\}(x) \\ &:= \frac{1}{2\pi i} \int_{\Gamma} \prod_{k=0}^{n-1} \frac{t + \overline{\lambda}_k + 1}{t - \lambda_k} \frac{x^t \, dt}{t - \lambda_n}, \quad n = 0, 1, \dots, \end{aligned} \tag{3.4.5}$$

where the positively oriented, simple closed contour Γ surrounds the zeros of the denominator in the integrand, and $\overline{\lambda}_k$ denotes the conjugate of λ_k.

The orthogonality of the above functions with respect to the Lebesgue measure is proved in Theorem 3.4.3. However, first we give a simple explicit representation of L_n in the case that the numbers λ_i are distinct. This is deduced immediately from evaluating the above integral by the residue theorem.

Proposition 3.4.2. *Let $\Lambda := (\lambda_i)_{i=0}^{\infty}$ be a sequence of distinct complex numbers. Then*

$$L_n\{\lambda_0, \dots, \lambda_n\}(x) = \sum_{k=0}^{n} c_{k,n} x^{\lambda_k}, \qquad x \in (0,\infty) \tag{3.4.6}$$

with

$$c_{k,n} := \frac{\prod_{j=0}^{n-1} (\lambda_k + \overline{\lambda}_j + 1)}{\prod_{j=0, j\neq k}^{n} (\lambda_k - \lambda_j)},$$

where $L_n\{\lambda_0, \dots, \lambda_n\}(x)$ is defined by (3.4.5).

So $L_n\{\lambda_0, \ldots, \lambda_n\}$ is indeed a Λ polynomial provided the numbers λ_i are distinct. Its value at $x = 0$ is defined if for all i either $\mathrm{Re}(\lambda_i) > 0$ or $\lambda_i = 0$.

From either Definition 3.4.1 or the above proposition it is obvious that the order of $\lambda_0, \ldots, \lambda_{n-1}$ in $L_n\{\lambda_0, \ldots, \lambda_n\}$ does not make any difference, as long as λ_n is kept last. For example,

$$L_2\{\lambda_0, \lambda_1, \lambda_2\} = L_2\{\lambda_1, \lambda_0, \lambda_2\}$$

while both, in general, are different from $L_2\{\lambda_0, \lambda_2, \lambda_1\}$. For a fixed sequence Λ, we let $L_n(\Lambda)$, or simply L_n, denote the nth Müntz-Legendre polynomial $L_n\{\lambda_0, \ldots, \lambda_n\}$, whenever there is no ambiguity.

An analog of Proposition 3.4.2 can be established even if the numbers λ_i are not distinct, however, in the nondistinct case, $L_n(\Lambda)$ does not belong to the space $M_n(\Lambda)$; see E.7 b]. In the very special case that all the indices are the same we recover the Laguerre polynomials; see E.1.

The orthogonality of $\{L_n\}_{n=0}^{\infty}$ is the content of the main theorem of this section.

Theorem 3.4.3 (Orthogonality). *Let $\Lambda = (\lambda_i)_{i=0}^{\infty}$ be a sequence of complex numbers with $\mathrm{Re}(\lambda_i) > -1/2$ for $i = 0, 1, \ldots$. The functions L_n defined by (3.4.5) satisfy*

$$\int_0^1 L_n(x)\overline{L_m(x)}\,dx = \frac{\delta_{n,m}}{1 + \lambda_n + \overline{\lambda}_n} \tag{3.4.7}$$

for all nonnegative integers n and m. (Here $\delta_{n,m}$ is the Kronecker symbol.)

Proof. We may assume that the numbers λ_i are distinct. Note that

$$L_n\{\lambda_0, \lambda_1, \ldots, \lambda_n\}(x)$$

is uniformly continuous in $\lambda_0, \ldots, \lambda_n$ for x in closed subintervals of $(0, 1]$, and the nondistinct λ_i case can be handled by a limiting argument. We may further suppose that $m \le n$. Since $\mathrm{Re}(\lambda_i) > -1/2$, we can pick a simple closed contour Γ such that Γ lies completely to the right of the vertical line $\mathrm{Re}(t) = -1/2$ and Γ surrounds all zeros of the denominator of the integrand in (3.4.5). When $t \in \Gamma$, we have $\mathrm{Re}(t + \overline{\lambda}_m) > -1$, and

$$\int_0^1 x^{t+\overline{\lambda}_m}\,dx = \frac{1}{t + \overline{\lambda}_m + 1}$$

for every $m = 0, 1, \ldots$. Hence Fubini's theorem yields

$$\int_0^1 L_n(x)x^{\overline{\lambda}_m}\,dx = \frac{1}{2\pi i}\int_\Gamma \prod_{k=0}^{n-1}\frac{t + \overline{\lambda}_k + 1}{t - \lambda_k}\frac{dt}{(t - \lambda_n)(t + \overline{\lambda}_m + 1)}\,.$$

Notice that for $m < n$, the new factor, $t + \overline{\lambda}_m + 1$, in the denominator can be cancelled, and for $m = n$ the new pole $-(\overline{\lambda}_n + 1)$ is outside Γ since $\mathrm{Re}(-\overline{\lambda}_n - 1) < -1/2$. Changing the contour from Γ to $|t| = R$ with $R > \max\{\lambda_0| + 1, \dots, |\lambda_n| + 1\}$, gives

$$\int_0^1 L_n(x) x^{\overline{\lambda}_m}\, dx = \frac{1}{2\pi i} \int_{|t|=R} \prod_{k=0}^{n-1} \frac{t + \overline{\lambda}_k + 1}{t - \lambda_k} \frac{dt}{(t - \lambda_n)(t + \overline{\lambda}_m + 1)} - \frac{\delta_{m,n}}{-\overline{\lambda}_n - 1 - \lambda_n} \prod_{k=0}^{n-1} \frac{-\overline{\lambda}_n + \overline{\lambda}_k}{-\overline{\lambda}_n - 1 - \lambda_k}\,.$$

On letting $R \to \infty$, we see that the integral on the right-hand side is actually 0, which gives

$$\int_0^1 L_n(x) x^{\overline{\lambda}_m}\, dx = \frac{\delta_{n,m}}{\overline{\lambda}_n + \lambda_n + 1} \prod_{k=0}^{n-1} \frac{\overline{\lambda}_n - \overline{\lambda}_k}{\overline{\lambda}_n + \lambda_k + 1}\,.$$

Therefore Proposition 3.4.2 and $m \le n$ yield

$$\begin{aligned} \int_0^1 L_n(x) \overline{L_m(x)}\, dx &= \int_0^1 L_n(x) \sum_{k=0}^{m} \overline{c}_{k,m} x^{\lambda_k}\, dx \\ &= \overline{c}_{m,m} \int_0^1 L_n(x) x^{\overline{\lambda}_m}\, dx = \frac{\delta_{m,n}}{\lambda_n + \overline{\lambda}_n + 1}\,, \end{aligned}$$

and the proof is finished. □

An alternative proof of orthogonality is suggested in E.3. If we let

$$L_n^* := (1 + \lambda_n + \overline{\lambda}_n)^{1/2} L_n\,, \tag{3.4.8}$$

then we get an orthonormal system, that is,

$$\int_0^1 L_n^*(x) \overline{L_m^*(x)}\, dx = \delta_{m,n}\,, \qquad m, n = 0, 1, \dots\,.$$

We call these L_n^* the *orthonormal Müntz-Legendre polynomials*.

There is also a Rodrigues-type formula for the Müntz-Legendre polynomials (see E.2). Let

$$p_n(x) = \sum_{k=0}^{n} \frac{x^{\lambda_k}}{\prod_{j=0, j\ne k}^{n} (\lambda_k - \lambda_j)}\,.$$

Then

$$L_n(x) = (D_{\lambda_0} D_{\lambda_1} \cdots D_{\lambda_{n-1}})(p_n)(x)\,,$$

where the differential operators D_λ are defined by

$$D_\lambda(f)(x) := x^{-\overline{\lambda}} \frac{d}{dx}(x^{1+\overline{\lambda}} f(x))\,.$$

The following is a differential recurrence formula for $(L_n)_{n=0}^{\infty}$:

Theorem 3.4.4. *For a fixed sequence $\Lambda := (\lambda_i)_{i=0}^{\infty}$ of complex numbers, let L_n be defined by (3.4.5). The identity*

$$xL_n'(x) - xL_{n-1}'(x) = \lambda_n L_n(x) + (1 + \overline{\lambda}_{n-1})L_{n-1}(x) \tag{3.4.9}$$

holds for every $x \in (0, \infty)$ and $n = 1, 2, \ldots$.

Proof. From (3.4.5) we get

$$(x^{-\lambda_n} L_n(x))' = \frac{1}{2\pi i} \int_\Gamma \frac{\prod_{k=0}^{n-2}(t + \overline{\lambda}_k + 1)}{\prod_{k=0}^{n-1}(t - \lambda_k)} (t + \overline{\lambda}_{n-1} + 1) x^{t-\lambda_n - 1}\, dt\,.$$

On multiplying both sides by $x^{\lambda_n + \overline{\lambda}_{n-1}+1}$, we obtain

$$\begin{aligned} x^{\lambda_n + \overline{\lambda}_{n-1}+1}&(x^{-\lambda_n} L_n(x))' \\ &= \frac{1}{2\pi i} \int_\Gamma \frac{\prod_{k=0}^{n-2}(t + \overline{\lambda}_k + 1)}{\prod_{k=0}^{n-1}(t - \lambda_k)} (t + \overline{\lambda}_{n-1} + 1) x^{t+\overline{\lambda}_{n-1}}\, dt\,, \end{aligned}$$

and again by the definition of L_{n-1},

$$x^{\lambda_n + \overline{\lambda}_{n-1}+1}(x^{-\lambda_n} L_n(x))' = (x^{\overline{\lambda}_{n-1}+1} L_{n-1}(x))'\,.$$

We finish the proof by simplifying by the product rule and dividing both sides by $x^{\overline{\lambda}_{n-1}}$. □

Corollary 3.4.5. *For a fixed sequence $\Lambda := (\lambda_i)_{i=0}^{\infty}$ of complex numbers, let L_n and L_n^* be defined by (3.4.5) and (3.4.8), respectively. Then for every $x \in (0, \infty)$ and for every $n = 0, 1, \ldots$,*

a] $$xL_n'(x) = \lambda_n L_n(x) + \sum_{k=0}^{n-1} (\lambda_k + \overline{\lambda}_k + 1) L_k(x)\,,$$

b] $$xL_n^{*\prime}(x) = \lambda_n L_n^*(x) + \sqrt{\lambda_n + \overline{\lambda}_n + 1} \sum_{k=0}^{n-1} \sqrt{\lambda_k + \overline{\lambda}_k + 1}\, L_k^*(x)\,,$$

and

c] $$xL_n''(x) = (\lambda_n - 1)L_n'(x) + \sum_{k=0}^{n-1} (\lambda_k + \overline{\lambda}_k + 1) L_k'(x)\,.$$

Proof. The first identity follows from Theorem 3.4.4 on expressing

$$xL_n'(x) - xL_0'(x)$$

as a telescoping sum. From this and the relation

$$L_k^* = (\lambda_k + \overline{\lambda}_k + 1)^{1/2} L_k$$

we get part b]. Differentiating the identity of part a] gives part c]. □

The values and derivative values of the Müntz-Legendre polynomials at 1 can now all be calculated.

Corollary 3.4.6. *For a fixed sequence $\Lambda := (\lambda_i)_{i=0}^{\infty}$ of complex numbers, let L_n be defined by (3.4.5). Then*

a] $L_n(1) = 1$,

b] $L_n'(1) = \lambda_n + \sum_{k=0}^{n-1} (\lambda_k + \overline{\lambda}_k + 1)$,

and

c] $L_n''(1) = (\lambda_n - 1)L_n'(1) + \sum_{k=0}^{n-1} (\lambda_k + \overline{\lambda}_k + 1)L_k'(1)$.

Proof. It suffices to show that $L_n(1) = 1$; the rest follows from Corollary 3.4.5. Notice that

$$L_n(1) = \frac{1}{2\pi i} \int_\Gamma \prod_{k=0}^{n-1} \frac{t + \overline{\lambda}_k + 1}{t - \lambda_k} \frac{dt}{t - \lambda_n}.$$

Now, since Γ surrounds all zeros of the denominator, and the degree of the denominator is one higher than that of the numerator, we can evaluate the integral on circles of radius $R \to \infty$ to get the result. □

Comments, Exercises, and Examples.

Müntz polynomials are just exponential polynomials $\sum a_k e^{-\lambda_k t}$ under the change of variables $x = e^{-t}$ and have received considerable scrutiny (Schwartz [59] is a monograph on this topic). The orthogonalizations of Müntz systems exist in the Russian literature (see, for example, Badalyan [55] and Taslakyan [84]; it has been further explored in McCarthy, Sayre, and Shawyer [93]). Borwein, Erdélyi, and Zhang [94b] contains most of the content of this section.

Various properties of the Müntz-Legendre polynomials are examined in the exercises. Note that if $\lambda_0 < \lambda_1 < \lambda_2 < \cdots$, then the Müntz system $(x^{\lambda_0}, x^{\lambda_1}, \dots)$ is a Descartes system on $[a, b]$, $a > 0$, and so we can apply Theorem 3.3.4 to the associated Chebyshev polynomials on $[a, b]$ to deduce how the zeros shift when the exponents are varied lexicographically. Similar results are given for the Müntz-Legendre polynomials in E.7.

E.1 Laguerre Polynomials.

a] Let $L_n\{\lambda_0, \dots, \lambda_n\}(x)$ be defined by (3.4.5). If $\lambda_0 = \cdots = \lambda_n = \lambda$, then

$$L_n\{\lambda_0, \lambda_1, \dots, \lambda_n\}(x) = x^\lambda \mathcal{L}_n(-(1 + \lambda + \overline{\lambda}) \log x),$$

where $\mathcal{L}_n$ is the nth Laguerre polynomial orthonormal with respect to the weight e^{-x} on $[0, \infty)$ with $\mathcal{L}_n(0) = 1$ (see E.7 of Section 2.3, where $\mathcal{L}_n$ is denoted by L_n as is standard).

Proof. Since $\lambda_k = \lambda$, (3.4.5) yields

$$L_n\{\lambda_0, \lambda_1, \ldots, \lambda_n\}(x) = \frac{1}{2\pi i} \int_\Gamma \frac{x^t (t + \overline{\lambda} + 1)^n}{(t - \lambda)^{n+1}}\, dt\,,$$

where the contour Γ can be taken to be any circle centered at λ. By the residue theorem,

$$\begin{aligned} L_n\{\lambda_0, \lambda_1, \ldots, \lambda_n\}(x) &= \frac{1}{n!} \left[\frac{d^n}{dt^n}(x^t (t + \overline{\lambda} + 1)^n\right]_{t=\lambda} \\ &= \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^\lambda (\log x)^k [n(n-1)\cdots(k+1)](\lambda + \overline{\lambda} + 1)^k \\ &= x^\lambda \sum_{k=0}^{n} \frac{1}{k!} \binom{n}{k} (1 + \lambda + \overline{\lambda})^k \log^k x\,. \end{aligned}$$

See also part b]. □

b] Let

$$\mathcal{L}_n(x) := \sum_{k=0}^{n} \frac{1}{k!} \binom{n}{k} (-x)^k.$$

Then $(\mathcal{L}_n)_{n=0}^\infty$ is an orthonormal sequence of polynomials on $[0, \infty)$ with respect to the inner product

$$\langle f, g \rangle = \int_0^\infty f(x) g(x) e^{-x}\, dx\,.$$

Deduce the orthonormality from a] and Theorem 3.4.3 by substituting

$$y = -(1 + \lambda + \overline{\lambda}) \log x\,.$$

E.2 Rodrigues-Type Formula. Let $\Lambda = (\lambda_i)_{i=0}^\infty$ be a sequence of distinct complex numbers. Let L_n be defined by (3.4.5).

a] Let

$$p_n(x) := \sum_{k=0}^{n} \frac{x^{\lambda_k}}{\prod_{j=0,\, j\neq k}^{n} (\lambda_k - \lambda_j)}\,.$$

Show that

$$p_n(x) = \frac{1}{2\pi i} \int_\Gamma \frac{x^t\, dt}{\prod_{j=0}^{n} (t - \lambda_j)}\,,$$

where Γ is any contour surrounding $\lambda_0, \lambda_1, \ldots, \lambda_n$. Use this to show that

$$p_n^{(k)}(1) = 0\,, \qquad k = 0, 1, \ldots, n-1\,.$$

b] Show that

$$L_n(x) = (D_{\lambda_0} D_{\lambda_1} \cdots D_{\lambda_{n-1}})(p_n)(x)\,,$$

where

$$D_\lambda(f)(x) := x^{-\overline{\lambda}}(x^{1+\overline{\lambda}} f(x))'\,.$$

c] If $0 = \lambda_0 < \lambda_1 < \cdots$, then

$$p_n(0) = (-1)^n \prod_{j=1}^{n} \lambda_j^{-1}$$

and $(-1)^n p_n$ is strictly decreasing on $[0,1]$.

E.3 Another Proof of Orthogonality. Deduce the orthogonality of the sequence $(L_n)_{n=0}^{\infty}$ on $[0,1]$ from Theorem 3.4.4 by using integration by parts and induction.

E.4 Integral Recursion. For a given sequence $\Lambda := (\lambda_i)_{i=0}^{\infty}$ of complex numbers satisfying (3.4.4), let L_n be defined by (3.4.5). Show that

$$L_n(x) = L_{n-1}(x) - (\lambda_n + \overline{\lambda}_{n-1} + 1)x^{\lambda_n} \int_x^1 t^{-\lambda_n - 1} L_{n-1}(t)\,dt\,, \quad x \in (0,1]\,.$$

Hint: Use Theorem 3.4.4. □

E.5 On the Maximum of L_n on $[0,1]$. If $\Lambda = (\lambda_i)_{i=0}^{\infty}$ is a sequence of nonnegative numbers satisfying

$$\lambda_n \geq \sum_{k=0}^{n-1} (1 + 2\lambda_k)\,, \qquad n = 1, 2, \ldots \tag{3.4.10}$$

and L_n is defined by (3.4.5), then

$$|L_n(x)| < L_n(1) = 1\,, \qquad x \in [0,1)\,, \quad n = 2, 3, \ldots\,.$$

Hint: Use Theorem 3.4.4. □

If $\lambda_k = \rho^k$, then (3.4.10) holds if and only if $\rho \geq 2 + \sqrt{3}$.

E.6 The Reproducing Kernel. Let $\Lambda = (\lambda_i)_{i=0}^{\infty}$ be as in (3.4.4), and let L_n and L_n^* be defined by (3.4.5) and (3.4.8), respectively. Then for every $p \in M_n(\Lambda)$, we have

$$p(x) = \int_0^1 K_n(x,t) p(t)\,dt\,,$$

where

$$K_n(x,t) := \sum_{k=0}^{n} L_k^*(x)\overline{L_k^*(t)}$$

is the nth reproducing kernel (see also E.5 of Section 2.2).

E.7 On the Zeros of Müntz-Legendre Polynomials. Assume that

$$(\lambda_0, \lambda_1, \dots, \lambda_n) \subset \left(-\tfrac{1}{2}, \infty\right).$$

a] For a function $f \in C(0,1)$ let $S^-(f)$ and $Z(f)$ denote the number of sign changes and the number of zeros, respectively, of f in $(0,1)$ (we count each zero without sign change twice). Let Φ and $\Psi \in C(0,1)$. Show that if

$$n \le S^-(\alpha\Phi + \beta\Psi) \le Z(\alpha\Phi + \beta\Phi) \le n+1$$

for every real α and β with $\alpha^2 + \beta^2 > 0$, then the zeros of Φ and Ψ strictly interlace.

Proof: This result is due to Pinkus and Ziegler [79]. □

b] Assume that

$$\{\lambda_0, \lambda_1, \dots, \lambda_n\} = \{\widetilde{\lambda}_0, \widetilde{\lambda}_1, \dots, \widetilde{\lambda}_m\},$$

where the numbers $\widetilde{\lambda}_0, \widetilde{\lambda}_1, \dots, \widetilde{\lambda}_m$ are distinct, and let m_j, $j = 0, 1, \dots, m$, be the number of indices $i = 0, 1, \dots, n$ for which $\lambda_i = \widetilde{\lambda}_j$. Show that $L_n\{\lambda_0, \dots, \lambda_n\}$ is in the space

$$H_n := \operatorname{span}\{x^{\lambda_j}(\log x)^i : j = 0, 1, \dots, m, \;\; i = 0, 1, \dots, m_j - 1\},$$

which is a Chebyshev space on $(0, \infty)$.

Hint: Use the definition and the residue theorem. □

c] Show that $\{L_k\{\lambda_0, \dots, \lambda_k\}\}_{k=0}^n$ is a basis for the Chebyshev space H_n defined in part b].

Hint: Use Theorem 3.4.3 (orthogonality). □

d] Show that $L_n := L_n\{\lambda_0, \dots, \lambda_n\}$ has exactly n distinct zeros in $(0,1)$ and L_n changes sign at each of these zeros.

Hint: Assume to the contrary that the number of sign changes of L_n is less than n. Use part c] to find a function $p \in \operatorname{span}\{L_k\}_{k=0}^{n-1}$ that changes sign exactly at those points in $(0,1)$ where L_n changes sign. Then $\int_0^1 L_n p \neq 0$, which contradicts Theorem 3.4.3. □

e] Suppose $\lambda_n < \lambda_n^*$. Show that the zeros of

$$\Phi := L_n\{\lambda_0, \lambda_1, \dots, \lambda_{n-1}, \lambda_n\}$$

and

$$\Psi := L_n\{\lambda_0, \lambda_1, \dots, \lambda_{n-1}, \lambda_n^*\}$$

in $(0,1)$ strictly interlace.

Hint: Note that Theorem 3.4.3 (orthogonality) implies

$$\int_0^1 (\alpha\Phi + \beta\Psi)p = 0$$

for every $p \in H_{n-1}$, where H_{n-1} is defined by part b] with respect to the sequence $(\lambda_0, \lambda_1, \dots, \lambda_{n-1})$. Use the hint given to part d] to show that $\alpha\Phi + \beta\Psi$ has at least n sign changes in $(0,1)$ whenever α and β are real with $\alpha^2 + \beta^2 > 0$. Use part b] to obtain that $\alpha\Phi + \beta\Psi$ cannot have more than $n+1$ zeros in $(0,1)$ whenever α and β are real with $\alpha^2 + \beta^2 > 0$. Finish the proof by part a]. □

f] Let $\lambda_0, \dots, \lambda_{k-1}, \lambda_{k+1}, \dots, \lambda_n$ be fixed distinct numbers. Suppose

$$(\lambda_{k,i})_{i=1}^{\infty} \subset (-1/2, \infty)$$

is a sequence with $\lim_{i\to\infty} \lambda_{k,i} = \infty$. Show that the largest zero of

$$L_{n,k,i} := L_n\{\lambda_0, \dots, \lambda_{k-1}, \lambda_{k,i}, \lambda_{k+1}, \dots, \lambda_n\}$$

in $(0,1)$ tends to 1.

Outline. Assume, without loss, that $\lambda_{k,i}$ is greater than each of the numbers λ_j, $j = 0, 1, \dots, n$, $j \neq k$. Let

$$g_i(x) := \lambda_{k,i}(L_{n,k,i}(x) - c_{k,n}^{(i)}\, x^{\lambda_{k,i}}),$$

where

$$c_{k,n}^{(i)} = \frac{\prod_{j=0}^{n-1}(\lambda_{k,i} + \lambda_j + 1)}{\prod_{j=0, j\neq k}^{n}(\lambda_{k,i} - \lambda_j)}$$

is the coefficients of $x^{\lambda_{k,i}}$ in $L_{n,k,i}$. Use (3.4.6) to show that the functions g_i converge uniformly on $[\delta, 1]$, $\delta \in (0,1)$, to a function

$$0 \not\equiv g \in H_{n-1} := \operatorname{span}\{x^{\lambda_0}, \dots, x^{\lambda_{k-1}}, x^{\lambda_{k+1}}, \dots, x^{\lambda_n}\}.$$

Use $L_{n,k,i}(1) = 1$ (see Corollary 3.4.6) and the explicit formula for $c_{k,n}^{(i)}$ to show that $g(1) \leq 0$ and that the functions

$$L_{n,k,i}(x) = (L_{n,n,i}(x) - c_{k,n}^{(i)}\, x^{\lambda_{k,i}}) + c_{k,n}^{(i)}\, x^{\lambda_{k,i}}$$

converge to $g(x)$, as $i \to \infty$, for every $x \in (0,1)$.

Now assume that the statement of part f] is false. Then there is an $\epsilon \in (0,1)$ and a subsequence $(\lambda_{k,i_j})_{j=1}^{\infty}$ of $(\lambda_{k,i})_{i=1}^{\infty}$ such that the Müntz-Legendre polynomials L_{n,k,i_j} have no zeros in $[1-\epsilon, 1]$. Deduce from this and

$L'_{n,k,i_j}(1) > 0$, $\lambda_{n,i_j} > 0$ (see Corollary 3.4.6 a]), that g_{i_j} is nondecreasing on $[1-\epsilon, 1]$ whenever $\lambda_{n,i_j} > 0$.

Therefore g is nondecreasing on $[1-\epsilon, 1]$, which, together with $0 \neq g \in H_{n-1}$ and $g(1) \leq 0$, implies that $g(1-\epsilon) < 0$. Hence $L_{n,k,i}(1-\epsilon) < 0$ if i is large enough. Since $L_{n,k,i}(1) = 1$ (see Corollary 3.4.6), each $L_{n,k,i}$ has a zero in $(1-\epsilon, 1)$, provided i is large enough, which contradicts our assumption. □

g] Let Φ and Ψ be as in part e]. Let

$$x_1 < x_2 < \cdots < x_n \quad \text{and} \quad x_1^* < x_2^* < \cdots < x_n^*$$

be the zeros of Φ and Ψ, respectively, in $(0,1)$. Show that $\lambda_n < \lambda_n^*$ implies that

$$x_j < x_j^*, \qquad j = 1, 2, \ldots, n.$$

Hint: Combine parts e] and f]. □

h] Let $\lambda_k \neq \lambda_n$. Show that the zeros of

$$\Phi := L_n\{\lambda_0, \ldots, \lambda_{k-1}, \lambda_k, \lambda_{k+1} \ldots, \lambda_{n-1}, \lambda_n\}$$

and

$$\Psi := L_n\{\lambda_0, \ldots, \lambda_{k-1}, \lambda_n, \lambda_{k+1} \ldots, \lambda_{n-1}, \lambda_k\}$$

in $(0,1)$ strictly interlace.

Hint: Use part a] and arguments similar to those given in the hints to part e]. Note that

$$\Phi(1) = \Psi(1) = 1 \qquad \text{and} \qquad \Psi'(1) - \Phi'(1) = \lambda_n - \lambda_k \neq 0$$

(see Corollary 3.4.6 a]) imply that $\alpha\Phi + \beta\Psi$ is not identically 0 whenever α and β are real with $\alpha^2 + \beta^2 > 0$. □

i] Let Φ and Ψ be as in part h]. Let

$$x_1 < x_2 < \cdots < x_n \quad \text{and} \quad x_1^* < x_2^* < \cdots < x_n^*$$

be the zeros of Φ and Ψ, respectively, in $(0,1)$. Show that $\lambda_k < \lambda_n$ implies that

$$x_j < x_j^*, \qquad j = 1, 2, \ldots, n.$$

Hint: By part h] it is sufficient to prove that $x_n \leq x_n^*$. Let H_n be the Chebyshev space defined in part b]. Corollary 3.4.6 implies that

$$\Phi(1) = \Psi(1) = 1 \qquad \text{and} \qquad \Psi'(1) - \Psi'(1) = \lambda_n - \lambda_k > 0.$$

Deduce from this and part i] that $x_n^* < x_n$ would imply that $0 \neq \Psi - \Phi \in H_n$ has at least $n+1$ distinct zeros in $(0,1]$, which is a contradiction. □

j] Lexicographic Property. Suppose

$$\max_{0\le i\le n} \lambda_i \le \min_{0\le j\le n} \mu_j$$

and $\lambda_i < \mu_j$ for some indices i and j. Let

$$x_1 < x_2 < \cdots < x_n \quad \text{and} \quad x_1^* < x_2^* < \cdots < x_n^*$$

be the zeros of

$$L_n\{\lambda_0, \lambda_1, \dots, \lambda_n\} \quad \text{and} \quad L_n\{\mu_0, \mu_1, \dots, \mu_n\},$$

respectively, in $(0,1)$. Show that

$$x_j < x_j^*, \qquad j = 1, 2, \dots, n.$$

Hint: Repeated applications of parts g] and i] give the desired result. □

k] Let $\lambda_0 < \lambda_n$. Let

$$x_1 < x_2 < \cdots < x_n \quad \text{and} \quad x_1^* < x_2^* < \cdots < x_n^*$$

be the zeros of

$$L_n\{\lambda_0, \lambda_1, \dots, \lambda_n\} \quad \text{and} \quad L_n\{\lambda_n, \lambda_{n-1}, \dots, \lambda_0\},$$

respectively, in $(0,1)$. Show that $(x_j)_{j=1}^n$ and $(x_j^*)_{j=1}^n$ strictly interlace and

$$x_j < x_j^*, \qquad j = 1, 2, \dots, n.$$

Hint: Use parts h] and i] and the comment given after Proposition 3.4.2. □

l] Show that the zeros of

$$\Phi := L_{n-1}\{\lambda_0, \dots, \lambda_{n-1}\} \quad \text{and} \quad \Psi := L_n\{\lambda_0, \dots, \lambda_n\}$$

in $(0,1)$ strictly interlace.

Hint: Use part a] and arguments similar to those given in the hints for part e]. □

E.8 A Global Estimate for the Zeros. Let $(\lambda_i)_{i=1}^n \subset (-1/2, \infty)$. Assume that $x_1 < x_2 < \cdots < x_n$ are the zeros of $L_n\{\lambda_0, \dots, \lambda_n\}$ in $(0,1)$. Then

$$\exp\left(-\frac{4n+2}{1+2\lambda_*}\right) < x_1 < x_2 < \cdots < x_n < \exp\left(\frac{-j_1^2}{(1+2\lambda^*)(4n+2)}\right),$$

where $\lambda_* := \min\{\lambda_0, \dots, \lambda_n\}$, $\lambda^* := \max\{\lambda_0, \dots, \lambda_n\}$ and $j_1 > 3\pi/4$ is the smallest positive zero of the Bessel function

$$J_0(z) := \sum_{k=0}^{\infty} \frac{(-z^2)^k}{(k!\,2^k)^2}\,.$$

Proof. Let $\mathcal{L}_n$ be the nth Laguerre polynomial with respect to the weight e^{-x} on $[0,\infty)$, and let the zeros of $\mathcal{L}_n$ be $z_1 < z_2 < \cdots < z_n$. Then by [Szegő [75], pp. 127–131])

$$\frac{j_1^2}{4n+2} < z_1 < z_2 < \cdots < z_n < 4n+2\,,$$

where the upper estimate is asymptotically sharp, and the lower estimate is sharp up to a multiplicative constant (not exceeding $4^4/(9\pi^2)$). Now use E.1 and E.7 j]. □

E.9 The Order of the Zero at 1 of Certain Polynomials. This exercise, due in part to Kós, gives precise estimates on the maximum order of the zero at 1 of a polynomial whose coefficients are bounded in modulus by the leading coefficient.

a] Suppose $a_0, a_1, \dots, a_{n-1}$ are complex numbers with modulus at most 1, and suppose $a_n = 1$. Then the multiplicity of the zero of

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

at 1 is at most $5\sqrt{n}$.

Proof. If p has a zero at 1 of multiplicity m, then for every polynomial f of degree less than m, we have

$$a_0 f(0) + a_1 f(1) + \cdots + a_n f(n) = 0\,. \tag{3.4.11}$$

We construct a polynomial f of degree at most $5\sqrt{n}$, for which

$$f(n) > |f(0)| + |f(1)| + \cdots + |f(n-1)|\,.$$

Equality (3.4.11) cannot hold with this f, so the multiplicity of the zero of p at 1 is at most the degree of f.

Let T_ν be the νth Chebyshev polynomial defined by (2.1.1). Let $k \in \mathbb{N}$, and let

$$g := T_0 + T_1 + \cdots + T_k \in \mathcal{P}_k\,.$$

Note that $g(1) = k+1$. Also, for $0 < y \le \pi$,

$$g(\cos y) = 1 + \cos y + \cos 2y + \cdots + \cos ky = \frac{\sin(k+\frac{1}{2})y + \sin\frac{1}{2}y}{2\sin\frac{1}{2}y}\,.$$

Hence, for $-1 \le x < 1$,

$$|g(x)| \le \frac{\sqrt{2}}{\sqrt{1-x}}\,.$$

Let $f(x) := g^4(\frac{2x}{n} - 1)$. Then $f(n) = g^4(1) = (k+1)^4$ and

$$|f(0)| + |f(1)| + \cdots + |f(n-1)| \le \sum_{j=1}^{n} \frac{4}{\left(\frac{2j}{n}\right)^2} < n^2 \sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2 n^2}{6}\,.$$

If $k := \lfloor (\pi^2/6)^{1/4}\sqrt{n} \rfloor$, then

$$f(n) > |f(0)| + |f(1)| + \cdots + |f(n-1)|\,.$$

In this case the degree of f is $4k \le 5\sqrt{n}$. □

The result of part a] is essentially sharp.

b] For every $n \in \mathbb{N}$, there exists a polynomial

$$p_n(x) = a_0 + a_1 x + \cdots + a_{2n^2} x^{2n^2}$$

such that $a_{2n^2} = 1$, $|a_0|, |a_1|, \ldots, |a_{2n^2-1}|$ are real numbers with modulus at most 1, and p_n has a zero at 1 with multiplicity at least n.

Proof. Define

$$L_n(x) := \frac{(n!)^2}{2\pi i} \int_\Gamma \frac{x^t\, dt}{\prod_{k=0}^{n}(t-k^2)}\,, \qquad n = 0, 1, \ldots\,,$$

where the simple closed contour Γ surrounds the zeros of the denominator of the integrand. Then L_n is a polynomial of degree n^2 with a zero of multiplicity at least n at 1. (This can easily be seen by repeated differentiation and then evaluation of the above contour integral by expanding the contour to infinity.)

Also, by the residue theorem,

$$L_n(x) = 1 + \sum_{k=1}^{n} c_{k,n} x^{k^2}$$

where

$$c_{k,n} = \frac{(-1)^n (n!)^2}{\prod_{j=0,\, j\ne k}^{n}(k^2 - j^2)} = \frac{(-1)^k 2 (n!)^2}{(n-k)!(n+k)!}\,.$$

It follows that

$$|c_{k,n}| \le 2\,, \qquad k = 1, 2, \ldots\,, n\,.$$

Hence,

$$P_n(x) := \frac{L_n(x) + L_n(x^2)}{2}$$

is a polynomial of degree $2n^2$ with a zero of order n at 1. Also P_n has constant coefficient 1 and each of its other coefficients is a real number of modulus less than 1. Now let $p_n(x) := x^{2n^2} P_n(1/x)$. □

c] For every $n \in \mathbb{N}$, there exists a polynomial

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

such that $a_n = 1$, $a_0, a_1, \ldots, a_{n-1}$ are real numbers of modulus less than 1, and p has a zero at 1 with multiplicity at least $\lfloor \sqrt{n/2} \rfloor$.

3.5 Chebyshev Polynomials in Rational Spaces

There are very few situations where Chebyshev polynomials can be explicitly computed. Indeed, only the classical case of Section 2.1 is well known.

However, the explicit formulas for the Chebyshev polynomials for the trigonometric rational system

$$\left(1\,,\ \frac{1 \pm \sin\theta}{\cos\theta - a_1}\,,\ \frac{1 \pm \sin\theta}{\cos\theta - a_2}\,,\ \ldots\,,\ \frac{1 \pm \sin\theta}{\cos\theta - a_n}\right)\,, \quad \theta \in [0, 2\pi) \tag{3.5.1}$$

and therefore also for the rational system

$$\left(1\,,\ \frac{1}{x - a_1}\,,\ \frac{1}{x - a_2}\,,\ \ldots\,,\ \frac{1}{x - a_n}\right)\,, \quad x \in [-1, 1] \tag{3.5.2}$$

with distinct real poles outside $[-1, 1]$ are implicitly contained in Achiezer [56].

The case (3.5.1) does not perfectly fit our discussion of Section 3.3 because of the periodicity or because $[0, 2\pi)$ is not a compact subset of $\mathbb{R}$. This leads to nonuniqueness of the Chebyshev polynomials. Note that ordinary polynomials arise as a limiting case of the span of system (3.5.2) on letting all the poles tend to $\pm\infty$.

We are primarily interested in the linear span of (3.5.2) and its trigonometric counterpart obtained with the substitution $x = \cos\theta$. Let

$$\mathcal{P}_n(a_1, a_2, \ldots, a_n) := \left\{ \frac{p(x)}{\prod_{k=1}^n |x - a_k|} : p \in \mathcal{P}_n \right\} \tag{3.5.3}$$

and

$$\mathcal{T}_n(a_1, a_2, \ldots, a_n) := \left\{ \frac{t(\theta)}{\prod_{k=1}^n |\cos\theta - a_k|} : t \in \mathcal{T}_n \right\}\,, \tag{3.5.4}$$

where $(a_k)_{k=1}^n \subset \mathbb{C}\backslash[-1, 1]$ is a fixed sequence of poles.

When the poles $a_1, a_2, \dots, a_n$ are distinct and real, (3.5.3) and (3.5.4) are simply the real spans of the systems

$$\left(1, \frac{1}{x-a_1}, \frac{1}{x-a_2}, \dots, \frac{1}{x-a_n}\right) \qquad \text{on } [-1,1] \tag{3.5.5}$$

and

$$\left(1, \frac{1 \pm \sin\theta}{\cos\theta - a_1}, \frac{1 \pm \sin\theta}{\cos\theta - a_2}, \dots, , \frac{1 \pm \sin\theta}{\cos\theta - a_n}\right) \qquad \text{on } [0, 2\pi)\,, \tag{3.5.6}$$

respectively.

We can construct Chebyshev polynomials of the first and second kinds, which are analogous to T_n and U_n of Section 2.1, for the spaces $\mathcal{P}_n(a_1, a_2, \dots, a_n)$ and $\mathcal{T}_n(a_1, a_2, \dots, a_n)$ as follows. Given a sequence $(a_k)_{k=1}^n \subset \mathbb{C}\backslash[-1,1]$, we define the sequence $(c_k)_{k=1}^n$ by

$$a_k = \tfrac{1}{2}(c_k + c_k^{-1})\,, \qquad |c_k| < 1\,, \tag{3.5.7}$$

that is,

$$c_k := a_k - \sqrt{a_k^2 - 1}\,, \qquad |c_k| < 1\,. \tag{3.5.8}$$

Note that

$$\left(a_k + \sqrt{a_k^2 - 1}\right)\left(a_k - \sqrt{a_k^2 - 1}\right) = 1\,.$$

In what follows, $\sqrt{a_k^2 - 1}$ is always defined by (3.5.8) (this specifies the choice of root). Let $D := \{z \in \mathbb{C} : |z| < 1\}$, let

$$M_n(z) := \left(\prod_{k=1}^{n} (z - c_k)(z - \overline{c}_k)\right)^{1/2}, \tag{3.5.9}$$

where the square root is defined so that $M_n^*(z) := z^n M_n(z^{-1})$ is an analytic function in a neighborhood of the closed unit disk $\overline{D}$, and let

$$f_n(z) := \frac{M_n(z)}{z^n M_n(z^{-1})}\,. \tag{3.5.10}$$

Note that f_n^2 is actually a finite Blaschke product (see E.12 of Section 4.2). Also, $f_n(z^{-1}) = f_n(z)^{-1}$ whenever $|z| = 1$.

The *Chebyshev polynomials of the first kind* for the spaces

$$\mathcal{P}_n(a_1, a_2, \dots, a_n) \quad \text{and} \quad \mathcal{T}_n(a_1, a_2, \dots, a_n)$$

are now defined by

$$T_n(x) := \tfrac{1}{2}(f_n(z) + f_n(z)^{-1}), \text{ where } x := \tfrac{1}{2}(z + z^{-1}), \ |z| = 1 \tag{3.5.11}$$

and

$$\widetilde{T}_n(\theta) := T_n(\cos\theta), \qquad \theta \in \mathbb{R}, \tag{3.5.12}$$

respectively.

The *Chebyshev polynomials of the second kind* for these two spaces are defined by

$$U_n(x) := \frac{f_n(z) - f_n(z)^{-1}}{z - z^{-1}}, \text{ where } x := \tfrac{1}{2}(z + z^{-1}), \quad |z| = 1 \tag{3.5.13}$$

and

$$\widetilde{U}_n(\theta) := U_n(\cos\theta)\sin\theta, \qquad \theta \in \mathbb{R}, \tag{3.5.14}$$

respectively.

As we will see, these Chebyshev polynomials preserve many of the elementary properties of the classical trigonometric and algebraic Chebyshev polynomials. This is the content of the next three results.

Theorem 3.5.1 (Chebyshev Polynomials of the First and Second Kinds in Trigonometric Rational Spaces). *Given* $(a_k)_{k=1}^n \subset \mathbb{C}\backslash[-1,1]$, *let* $\widetilde{T}_n$ *and* $\widetilde{U}_n$ *be defined by (3.5.12) and (3.5.14), respectively. Then the following statements hold:*

a] $\widetilde{T}_n \in \mathcal{T}_n(a_1, a_2, \dots, a_n)$ *and* $\widetilde{U}_n \in \mathcal{T}_n(a_1, a_2, \dots, a_n)$.

b] $\|\widetilde{T}_n\|_{\mathbb{R}} = 1$ *and* $\|\widetilde{U}_n\|_{\mathbb{R}} = 1$.

c] *There exist* $0 = \theta_0 < \theta_1 < \cdots < \theta_n = \pi$ *such that*

$$\widetilde{T}_n(\theta_j) = \widetilde{T}_n(-\theta_j) = (-1)^j, \quad j = 0, 1, \dots, n.$$

d] *There exist* $0 < \tau_1 < \tau_2 < \cdots < \tau_n < \pi$ *such that*

$$\widetilde{U}_n(\tau_j) = -\widetilde{U}_n(-\tau_j) = (-1)^{j-1}, \quad j = 1, 2, \dots, n.$$

e] *For every* $\theta \in \mathbb{R}$,

$$\widetilde{T}_n^2(\theta) + \widetilde{U}_n^2(\theta) = 1.$$

Proof. Observe that there are polynomials $p_1 \in \mathcal{P}_n$ and $p_2 \in \mathcal{P}_{n-1}$ such that

$$\begin{aligned} \widetilde{T}_n(\theta) = T_n(\cos\theta) &= \frac{e^{-in\theta}M_n^2(e^{i\theta}) + e^{in\theta}M_n^2(e^{-i\theta})}{2M_n(e^{i\theta})M_n(e^{-i\theta})} \\ &= \frac{p_1(\cos\theta)}{\prod_{k=1}^n |\cos\theta - a_k|} \end{aligned} \tag{3.5.15}$$

and

$$\widetilde{U}_n(\theta) = U_n(\cos\theta)\sin\theta = \frac{e^{-in\theta}M_n^2(e^{i\theta}) - e^{in\theta}M_n^2(e^{-i\theta})}{2iM_n(e^{i\theta})M_n(e^{-i\theta})} = \frac{p_2(\cos\theta)\sin\theta}{\prod_{k=1}^n |\cos\theta - a_k|}. \tag{3.5.16}$$

Thus a] is proved.

Since $|c_k| < 1$ and f_n^2 is a finite Blaschke product, we have

$$|f_n(z)| = 1 \quad \text{whenever} \quad |z| = 1\,. \tag{3.5.17}$$

Now b] follows immediately from (3.5.10) to (3.5.14).

Note that $\widetilde{T}_n(\theta)$ is the real part, and $\widetilde{U}_n(\theta)$ is the imaginary part of $f_n(e^{i\theta})$, that is,

$$f_n(e^{i\theta}) = \widetilde{T}_n(\theta) + i\widetilde{U}_n(\theta)\,, \qquad \theta \in \mathbb{R}\,,$$

which together with (3.5.17) implies e].

To prove parts c] and d], we note that $\widetilde{T}_n(\theta) = \pm 1$ if and only if $f_n(e^{i\theta}) = \pm 1$, and $\widetilde{U}_n(\theta) = \pm 1$ if and only if $f_n(e^{i\theta}) = \pm i$. Since $|c_k| < 1$ for $k = 1, 2, \ldots, n$, f_n^2 has exactly $2n$ zeros in the open unit disk D. Since f_n^2 is analytic in a region containing the closed unit disk $\overline{D}$, c] and d] follow by the argument principle (see, for example, Ash [71]). □

With the transformation $x = \cos\theta = \frac{1}{2}(z + z^{-1})$ and $z = e^{i\theta}$, Theorem 3.5.1 can be reformulated as follows:

Theorem 3.5.2 (Chebyshev Polynomials in Algebraic Rational Spaces). *Given $(a_k)_{k=1}^n \subset \mathbb{C} \setminus [-1,1]$, let T_n and U_n be defined by (3.5.11) and (3.5.13), respectively. Then*

a] $T_n \in \mathcal{P}_n(a_1, a_2, \ldots, a_n)$ *and* $U_n \in \mathcal{P}_n(a_1, a_2, \ldots, a_n)$.

b] $\|T_n\|_{[-1,1]} = 1$ *and* $\|\sqrt{1-x^2}\,U_n(x)\|_{[-1,1]} = 1$.

c] *There exist* $1 = x_0 > x_1 > \cdots > x_n = -1$ *such that*

$$T_n(x_j) = (-1)^j\,, \quad j = 0, 1, 2, \ldots, n\,.$$

d] *There exist* $1 > y_1 > y_2 > \cdots > y_n > -1$ *such that*

$$\sqrt{1-y_j^2}\,U_n(y_j) = (-1)^{j-1}\,, \quad j = 1, 2, \ldots, n\,.$$

e] *For every* $x \in [-1,1]$,

$$(T_n(x))^2 + (\sqrt{1-x^2}\,U_n(x))^2 = 1\,.$$

Parts c] and d] of Theorems 3.5.1 and 3.5.2 establish the equioscillation property of the Chebyshev polynomials, which also extends to certain linear combinations of Chebyshev polynomials. In the trigonometric polynomial case this is the fact that $\cos\alpha\cos n\theta+\sin\alpha\sin n\theta=\cos(n\theta-\alpha)$ equioscillates $2n$ times on the unit circle $[0,2\pi]$. Our next theorem characterizes the Chebyshev polynomials for $\mathcal{T}_n(a_1,a_2,\dots,a_n)$ and records a monotonicity property that we require later.

Theorem 3.5.3 (Chebyshev Polynomials in Trigonometric Rational Spaces). *Let $(a_k)_{k=1}^n\subset\mathbb{C}\setminus[-1,1]$. Then (i) and (ii) below are equivalent:*

(i) There is an $\alpha\in\mathbb{R}$ such that

$$V=(\cos\alpha)\,\widetilde{T}_n+(\sin\alpha)\,\widetilde{U}_n\,,$$

where $\widetilde{T}_n$ and $\widetilde{U}_n$ are defined by (3.5.12) and (3.5.14).

(ii) $V\in\mathcal{T}_n(a_1,a_2,\dots,a_n)$ has uniform norm 1 on $\mathbb{R}$, and it equioscillates $2n$ times on $\mathbb{R}$ (mod 2π). That is, there exist

$$0\le\theta_0<\theta_1<\cdots<\theta_{2n-1}<2\pi$$

so that

$$V(\theta_j)=\pm(-1)^j\,,\quad j=0,1,\dots,2n-1\,.$$

Furthermore, if V is of the form in (i) (or characterized by (ii)), then

$$V'=(\cos\alpha)\,\widetilde{T}_n'+(\sin\alpha)\,\widetilde{U}_n'$$

does not vanish between any two consecutive alternation points of V (that is, between θ_{j-1} and θ_j for $j=1,2,\dots,2n-1$ and between θ_{2n-1} and $2\pi+\theta_0$).

Proof. $(i)\Rightarrow(ii)$. By Theorem 3.5.1 e] and Cauchy's inequality, we have

$$|(\cos\alpha)\,\widetilde{T}_n+(\sin\alpha)\,\widetilde{U}_n|^2\le(\cos^2\alpha+\sin^2\alpha)(\widetilde{T}_n^2+\widetilde{U}_n^2)=1 \tag{3.5.18}$$

on the real line. From Theorem 3.5.1 c], d], and e], we obtain that $\widetilde{T}_n/\widetilde{U}_n$ oscillates between $+\infty$ and $-\infty$ exactly $2n$ times on $\mathbb{R}$ (mod 2π), and hence it takes the value $\cot\alpha$ exactly $2n$ times. At each such point, (3.5.18) becomes an equality, namely, $(\cos\alpha)\,\widetilde{T}_n+(\sin\alpha)\,\widetilde{U}_n=\pm1$ with different signs for every two consecutive such points.

$(ii)\Rightarrow(i)$. Let V be as specified in part (ii) of the theorem. Let θ_0 be a point where V achieves its maximum on $\mathbb{R}$, so $V(\theta_0)=1$. We want to show that V is equal to $p:=\widetilde{T}_n(\theta_0)\widetilde{T}_n+\widetilde{U}_n(\theta_0)\widetilde{U}_n$. Since $V(\theta_0)=p(\theta_0)=1$ and $V'(\theta_0)=p'(\theta_0)=0$, $V-p$ has a zero at θ_0 with multiplicity at least 2. There are at least $2n-1$ more zeros (we count multiplicities) of $V-p$

in $\mathbb{R} \pmod{2\pi}$, with one between any two consecutive alternation points of p if the first zero of p to the right of θ_0 is greater than the first zero of V to the right of θ_0. If the first zero of V to the right of θ_0 is greater than or equal to the first zero of p to the right of θ_0, then there is one zero of $p - V$ between any two consecutive alternation points of V. In any case $V - p$ has at least $2n + 1$ zeros in $\mathbb{R} \pmod{2\pi}$. Hence $V - p$ is identically 0.

To prove the final part of the theorem let $V \in \mathcal{T}_n(a_1, a_2, \dots, a_n)$ be such that $\|V\|_{\mathbb{R}} = 1$ and V equioscillates $2n$ times on $\mathbb{R} \pmod{2\pi}$ between ± 1. Assume there is a $\theta_0 \in [0, 2\pi)$ such that $|V(\theta_0)| < 1$ and $V'(\theta_0) = 0$. Then $V(\theta_0) \neq 0$; otherwise the numerator of V would have at least $2n + 1$ zeros in $\mathbb{R} \pmod{2\pi}$, which is a contradiction. Observe that there is a trigonometric polynomial $t \in \mathcal{T}_{2n}$ such that

$$V^2(\theta) - V^2(\theta_0) = \frac{t(\theta)}{\prod_{k=1}^{n} (\cos\theta - a_k)(\cos\theta - \overline{a}_k)} .$$

This t has at least $4n + 1$ zeros in $\mathbb{R} \pmod{2\pi}$, which is a contradiction again. Therefore $V'(\theta) \neq 0$ if $|V(\theta)| < 1$, which means that V is strictly monotone between any two of its consecutive alternation points. □

Under some assumptions on $(a_k)_{k=1}^n$ it is easy to write down the explicit partial fraction decompositions for T_n and U_n.

Theorem 3.5.4. *Let $(a_k)_{k=1}^n \subset \mathbb{C} \setminus [-1, 1]$ be a sequence of distinct numbers such that its nonreal elements are paired by complex conjugation. Let T_n and U_n be the Chebyshev polynomials of the first and second kinds defined by (3.5.11) and (3.5.13), respectively. Then*

$$T_n(x) = A_{0,n} + \frac{A_{1,n}}{x - a_1} + \cdots + \frac{A_{n,n}}{x - a_n} \tag{3.5.19}$$

and

$$U_n(x) = \frac{B_{1,n}}{x - a_1} + \frac{B_{2,n}}{x - a_2} + \cdots + \frac{B_{n,n}}{x - a_n} , \tag{3.5.20}$$

where

$$A_{0,n} = \frac{(-1)^n}{2}(c_1^{-1} c_2^{-1} \cdots c_n^{-1} + c_1 c_2 \cdots c_n) ,$$

$$A_{k,n} = \left(\frac{c_k - c_k^{-1}}{2}\right)^2 \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{1 - c_k c_j}{c_k - c_j} , \qquad k = 1, 2, \dots, n ,$$

and

$$B_{k,n} = \frac{c_k - c_k^{-1}}{2} \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{1 - c_k c_j}{c_k - c_j} , \quad k = 1, 2, \dots, n .$$

Proof. It follows from Theorems 3.5.1 a] and 3.5.2 a] that T_n and U_n can be written as the partial fraction forms above. Now it is quite easy to calculate the coefficients $A_{k,n}$ and $B_{k,n}$. For example,

$$\begin{aligned} A_{0,n} &= \lim_{x\to\infty} T_n(x) = \lim_{z\to 0} \frac{1}{2}\left(\frac{M_n(z)}{z^n M_n(z^{-1})} + \frac{z^n M_n(z^{-1})}{M_n(z)}\right) \\ &= \frac{(-1)^n}{2}(c_1^{-1}c_2^{-1}\cdots c_n^{-1} + c_1 c_2 \cdots c_n) \end{aligned}$$

and for $k = 1, 2, \ldots, n$,

$$\begin{aligned} A_{k,n} &= \lim_{x\to a_k} (x - a_k) T_n(x) \\ &= \lim_{z\to c_k} \frac{1}{4}(z - c_k)(1 - c_k^{-1} z^{-1}) \left(\frac{M_n(z)}{z^n M_n(z^{-1})} + \frac{z^n M_n(z^{-1})}{M_n(z)}\right) \\ &= \left(\frac{c_k - c_k^{-1}}{2}\right)^2 \prod_{\substack{j=1 \\ j\neq k}}^{n} \frac{1 - c_k c_j}{c_k - c_j}, \qquad k = 1, 2, \ldots, n. \end{aligned}$$

The coefficients $B_{k,n}$ can be calculated in the same fashion. □

Comments, Exercises, and Examples.

The explicit formulas of this section are tremendously useful. They allow, for example, derivation of sharp Bernstein-type inequalities for rational functions; see Section 7.1. Various further properties of these Chebyshev polynomials for rational function spaces are explored in the exercises, which follow, Borwein, Erdélyi, and Zhang [94b]. In particular, the orthogonalization of such rational systems on $[-1, 1]$ with respect to the weight $w(x) = (1 - x^2)^{-1/2}$ can be made explicit in terms of the Chebyshev polynomials. Various other aspects of these orthogonalizations may be found in Achiezer [56], Bultheel et al. [91], and Van Assche and Vanherwegen [92].

E.1 Further Properties of $\widetilde{T}_n$ and T_n. Given $(a_k)_{k=1}^n \subset \mathbb{C} \setminus [-1, 1]$, let $(c_k)_{k=1}^n$ be defined by

$$c_k := a_k - \sqrt{a_k^2 - 1}, \qquad |c_k| < 1,$$

as before. We introduce the *Bernstein factors*

$$B_n(x) := \sum_{k=1}^{n} \operatorname{Re}\left(\frac{\sqrt{a_k^2 - 1}}{a_k - x}\right)$$

and

$$\widetilde{B}_n(\theta) := B_n(\cos\theta) == \sum_{k=1}^{n} \operatorname{Re}\left(\frac{\sqrt{a_k^2 - 1}}{a_k - \cos\theta}\right),$$

where the choice of $\sqrt{a_k^2-1}$ is determined by the restriction $|c_k|<1$. Note that for $x\in[-1,1]$, we have

$$\operatorname{Re}\left(\frac{\sqrt{a_k^2-1}}{a_k-x}\right)=\operatorname{Re}\left(\frac{c_k^{-1}-c_k}{\frac{1}{2}c_k^{-1}-x}\right)\geq\frac{(1-|c_k|^2)(1-|c_k|)^2}{|1-2c_kx|^2}>0\,.$$

The following result generalizes the trigonometric identities

$$(\cos nt)'=-n\sin nt\,,\qquad(\sin nt)'=n\cos nt\,,$$

and

$$((\cos nt)')^2+((\sin nt)')^2=n^2\,,$$

which are limiting cases (if $n\in\mathbb{N}$ and $t\in\mathbb{R}$ are fixed, then $\lim\widetilde{B}_n(t)=n$ as all $a_k\to\pm\infty$).

a] Show that, on the real line,

$$\widetilde{T}_n'=-\widetilde{B}_n\widetilde{U}_n\,,\qquad\widetilde{U}_n'=\widetilde{B}_n\widetilde{T}_n\,,$$

and

$$(\widetilde{T}_n')^2+(\widetilde{U}_n')^2=\widetilde{B}_n^2\,.$$

Hint: For example,

$$\begin{aligned}\widetilde{T}_n'(\theta)&=\frac{1}{2}\left(f_n'(e^{i\theta})-\frac{f_n'(e^{i\theta})}{f_n^2(e^{i\theta})}\right)ie^{i\theta}\\&=\frac{-e^{i\theta}f_n'(e^{i\theta})}{f_n(e^{i\theta})}\,\frac{f_n(e^{i\theta})-f_n(e^{i\theta})^{-1}}{2i}=-\widetilde{B}_n(\theta)\widetilde{U}_n(\theta)\,.\end{aligned}$$

□

b] If $V:=(\cos\alpha)\,\widetilde{T}_n+(\sin\alpha)\,\widetilde{U}_n$ for some $\alpha\in\mathbb{R}$, then

$$(V')^2+\widetilde{B}_n^2V^2=\widetilde{B}_n^2$$

holds on the real line.

c] **The Derivative of T_n at ±1.** Let T_n be defined by (3.5.11). Then

$$T_n'(1)=\left(\sum_{k=1}^n\operatorname{Re}\left(\frac{1+c_k}{1-c_k}\right)\right)^2$$

and

$$T_n'(-1)=(-1)^n\left(\sum_{k=1}^n\operatorname{Re}\left(\frac{1-c_k}{1+c_k}\right)\right)^2\,.$$

d] Contour Integral for T_n. Show that

$$T_n(x) = \frac{1}{2\pi i}\int_\gamma \left(\prod_{j=1}^n \frac{(t-c_j)(t-\overline{c}_j)}{(1-c_jt)(1-\overline{c}_jt)}\right)^{1/2} \frac{t-x}{t^2-2tx+1}\,dt$$

for every $x \in [-1,1]$, where γ is a circle centered at the origin with radius $1 < r < \min\{|c_j^{-1}| : 1 \le j \le n\}$, and the square root in the integrand is an analytic function of t in a neighborhood of γ.

Hint: Cauchy's integral formula and the map $x = \frac{1}{2}(z+z^{-1})$ give

$$\begin{aligned} T_n(x) &= \frac{1}{2}\left(\frac{M_n(z)}{z^nM_n(z^{-1})} + \frac{z^nM_n(z^{-1})}{M_n(z)}\right) \\ &= \frac{1}{2\pi i}\int_\gamma \frac{1}{2}\frac{M_n(t)}{t^nM_n(t^{-1})}\left(\frac{1}{t-z}+\frac{1}{t-z^{-1}}\right)dt \\ &= \frac{1}{2\pi i}\int_\gamma \frac{M_n(t)}{t^nM_n(t^{-1})}\frac{t-x}{t^2-2tx+1}\,dt\,, \end{aligned}$$

where M_n is defined by (3.5.9). □

E.2 Orthogonality. Given $(a_k)_{k=1}^\infty \subset \mathbb{R}\setminus[-1,1]$, let $(c_k)_{k=1}^\infty$ be defined by

$$a_k = \tfrac{1}{2}(c_k + c_k^{-1})\,, \qquad c_k = a_k - \sqrt{a_k^2-1}\,, \quad c_k \in (-1,1)$$

and let $(T_n)_{n=0}^\infty$ be defined by (3.5.11).

a] Show that

$$\int_{-1}^1 T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}(-1)^{n+m}(1+c_1^2\cdots c_m^2)c_{m+1}\cdots c_n$$

for all integers $0 \le m \le n$. (The empty product is understood to be 1.)

b] Given $a \in \mathbb{R}\setminus[-1,1]$, let $c \in (-1,1)$ be defined by

$$a = \tfrac{1}{2}(c+c^{-1})\,, \qquad c = a - \sqrt{a^2-1}\,, \quad c \in (-1,1)\,.$$

Show that

$$\int_{-1}^1 T_n(x)\frac{1}{x-a}\frac{dx}{\sqrt{1-x^2}} = \frac{2\pi}{c-c^{-1}}\prod_{j=1}^n \frac{c-c_j}{1-cc_j}\,.$$

c] Show that

$$\int_{-1}^{1} T_n(x)\frac{dx}{\sqrt{1-x^2}} = (-1)^n \pi c_1 c_2 \cdots c_n$$

and

$$\int_{-1}^{1} T_n(x)\frac{1}{x-a_k}\frac{dx}{\sqrt{1-x^2}} = 0\,, \qquad k = 1,2,\ldots,n\,.$$

Given a sequence $(a_k)_{k=1}^{\infty} \subset \mathbb{R}\backslash[-1,1]$, we define

$$R_0 := 1\,, \qquad R_n := T_n + c_n T_{n-1}$$

and

$$R_0^* := \frac{1}{\sqrt{\pi}}\,, \qquad R_n^* := \sqrt{\frac{2}{\pi(1-c_n^2)}}\,(T_n + c_n T_{n-1})\,.$$

The following part of this exercise indicates that these simple linear combinations of T_n and T_{n-1} give the orthogonalization of the rational system

$$\left(1\,,\ \frac{1}{x-a_1}\,,\ \frac{1}{x-a_2}\,,\ \ldots\right)$$

whenever $(a_k)_{k=1}^{\infty} \subset \mathbb{R}\setminus[-1,1]$ is a sequence of distinct real numbers.

d] Show that, for all nonnegative integers n and m,

$$\int_{-1}^{1} R_n^*(x) R_m^*(x)\frac{dx}{\sqrt{1-x^2}} = \delta_{m,n}\,,$$

where $\delta_{m,n}$ is the Kronecker symbol.

Proof. Let $m \le n$. By part c],

$$\int_{-1}^{1} R_n(x)\frac{1}{x-a_k}\frac{dx}{\sqrt{1-x^2}} = 0$$

holds for $k = 1,2,\ldots,n-1$. Also

$$\begin{aligned}\int_{-1}^{1} R_n(x)\frac{dx}{\sqrt{1-x^2}} &= \int_{-1}^{1} (T_n(x) + c_n T_{n-1}(x))\frac{dx}{\sqrt{1-x^2}} \\ &= (-1)^n(c_1 c_2 \cdots c_n) + c_n(-1)^{n-1}(c_1 c_2 \cdots c_{n-1}) = 0\,.\end{aligned}$$

This implies that

$$\int_{-1}^{1} R_n(x) R_m(x)\frac{dx}{\sqrt{1-x^2}} = 0\,, \qquad m = 0,1,\ldots,n-1\,.$$

Finally, it follows from part a] that

$$\int_{-1}^{1} R_n^*(x)^2\frac{dx}{\sqrt{1-x^2}} = 1\,.$$

□

e] Assume $(a_k)_{k=1}^{\infty} \subset \mathbb{R}\setminus[-1,1]$. Then T_n and R_n have exactly n zeros in $[-1,1]$, and the zeros of T_{n-1} and T_n strictly interlace.

E.3 Extension of Theorems 3.5.1 and 3.5.3. Given $(a_k)_{k=1}^{2n} \subset \mathbb{C} \setminus \mathbb{R}$, let

$$\mathcal{T}_n(a_1, a_2, \ldots, a_{2n}) := \left\{ \frac{t(\theta)}{\prod_{k=1}^{2n} |\sin((\theta - a_k)/2)|} : t \in \mathcal{T}_n \right\}.$$

Without loss of generality we may assume that

$$\mathrm{Im}(a_k) > 0\,, \qquad k = 1, 2, \ldots, 2n\,.$$

a] Show that there is a polynomial $q_{2n} \in \mathcal{P}_{2n}^c$ of the form

$$q_{2n}(z) = \gamma \prod_{k=1}^{2n} (z - c_k)\,, \qquad |c_k| < 1\,, \quad \gamma \in \mathbb{C}$$

such that

$$|q_{2n}(e^{i\theta})| = \prod_{k=1}^{2n} |\sin((\theta - a_k)/2)|\,, \qquad \theta \in \mathbb{R}\,.$$

Hint: Use the fact that $|z - c| = |1 - \overline{c}z|$ whenever $|z| = 1$ and $c \in \mathbb{C}$. □

Associated with $q_{2n} \in \mathcal{P}_{2n}^c$ defined in part a], let

$$M_n(z) := \sqrt{q_{2n}(z)}$$

and

$$M_n^*(z) := \left(\overline{\gamma} \prod_{k=1}^{2n} (1 - \overline{c}z) \right)^{1/2},$$

where the square roots are defined so that M_n^* is analytic in a neighborhood of the closed unit disk, and M_n is analytic in a neighborhood of the complement of the open unit disk. Let

$$f_n(z) := \frac{M_n(z)}{M_n^*(z)}\,.$$

For $\theta \in \mathbb{R}$, we define

$$\widetilde{T}_n(\theta) := \mathrm{Re}(f_n(e^{i\theta})) = \frac{1}{2} \left(\frac{M_n(e^{i\theta})}{M_n^*(e^{i\theta})} + \frac{M_n^*(e^{i\theta})}{M_n(e^{i\theta})} \right)$$

and

$$\widetilde{U}_n(\theta) := \mathrm{Im}(f_n(e^{i\theta})) = \frac{1}{2i} \left(\frac{M_n(e^{i\theta})}{M_n^*(e^{i\theta})} - \frac{M_n^*(e^{i\theta})}{M_n(e^{i\theta})} \right).$$

Using the new (extended) definitions, show the following:

b] $\widetilde{T}_n \in \mathcal{T}_n(a_1, a_2, \dots, a_{2n})$ and $\widetilde{U}_n \in \mathcal{T}_n(a_1, a_2, \dots, a_{2n})$.

c] $\|\widetilde{T}_n\|_{\mathbb{R}} = 1$ and $\|\widetilde{U}_n\|_{\mathbb{R}} = 1$.

d] There are numbers $\theta_1 < \theta_2 < \cdots < \theta_{2n}$ in $[-\pi, \pi)$ such that

$$\widetilde{T}(\theta_j) = \pm(-1)^j, \qquad j = 1, 2, \dots, 2n.$$

e] There are numbers $\tau_1 < \tau_2 < \cdots < \tau_{2n}$ in $[-\pi, \pi)$ such that

$$\widetilde{U}(\tau_j) = \pm(-1)^j, \qquad j = 1, 2, \dots, 2n.$$

f] $\widetilde{T}(\theta)^2 + \widetilde{U}(\theta)^2 = 1$ for every $\theta \in \mathbb{R}$.

g] Both $\widetilde{T}_n$ and $\widetilde{U}_n$ have exactly $2n$ simple zeros in the period $[-\pi, \pi)$, and the zeros of $\widetilde{T}_n$ and $\widetilde{U}_n$ strictly interlace.

h] The statements of Theorem 3.5.3 remain valid.

E.4 Extension of the Bernstein Factor $\widetilde{B}_n$. Let

$$(a_k)_{k=1}^{2n} \subset \mathbb{C} \setminus \mathbb{R}, \qquad \operatorname{Im}(a_k) > 0.$$

With the notation of the previous exercise we define

$$\widetilde{B}_n(\theta) := \frac{e^{i\theta} f_n'(e^{i\theta})}{f_n(e^{i\theta})}, \qquad \theta \in \mathbb{R}.$$

a] Show that for every $\theta \in \mathbb{R}$,

$$\widetilde{B}_n(\theta) = \sum_{k=1}^{2n} \frac{1 - |c_k|^2}{|c_k - e^{i\theta}|^2} = \sum_{k=1}^{2n} \frac{1 - |e^{ia_k}|^2}{|e^{ia_k} - e^{i\theta}|^2}.$$

b] Show that, on the real line,

$$\widetilde{T}_n' = -\widetilde{B}_n \widetilde{U}_n, \qquad \widetilde{U}_n' = \widetilde{B}_n \widetilde{T}_n,$$

and

$$(\widetilde{T}_n')^2 + (\widetilde{U}_n')^2 = \widetilde{B}_n^2.$$

c] Show that

$$(V')^2 + \widetilde{B}_n^2 V^2 = \widetilde{B}_n^2$$

holds on the real line for every V of the form

$$V = (\cos\alpha)\, \widetilde{T}_n + (\sin\alpha)\, \widetilde{U}_n, \qquad \alpha \in \mathbb{R}.$$

E.5 Chebyshev Polynomials for $\mathcal{P}_n(a_1, a_2, \dots, a_n)$ on $\mathbb{R}$. Let

$$(a_k)_{k=1}^n \subset \mathbb{C} \setminus \mathbb{R} \qquad \text{with} \qquad \operatorname{Im}(a_k) > 0\,, \qquad k = 1, 2, \dots, n\,.$$

Let

$$M_n(z) := \left(\prod_{k=1}^{n} (z - a_k) \right)^{1/2}$$

and

$$M_n^*(z) := \left(\prod_{k=1}^{n} (z - \overline{a}_k) \right)^{1/2},$$

where the square roots are defined so that M_n^* is analytic in a neighborhood of the closed upper half-plane, and M_n is analytic in a neighborhood of the closed lower half-plane. Let

$$f_n(z) := \frac{M_n(z)}{M_n^*(z)}\,.$$

For $x \in \mathbb{R}$, we define

$$T_n(x) := \operatorname{Re}(f_n(x)) = \frac{1}{2} \left(\frac{M_n(x)}{M_n^*(x)} + \frac{M_n^*(x)}{M_n(x)} \right)$$

and

$$U_n(x) := \operatorname{Im}(f_n(x)) = \frac{1}{2i} \left(\frac{M_n(x)}{M_n^*(x)} - \frac{M_n^*(x)}{M_n(x)} \right).$$

Show the following:

a] $T_n \in \mathcal{P}_n(a_1, a_2, \dots, a_n)$ and $U_n \in \mathcal{P}_n(a_1, a_2, \dots, a_n)$.

b] $\|T_n\|_{\mathbb{R}} = 1$ and $\|U_n\|_{\mathbb{R}} = 1$.

c] There are real numbers $x_1 > x_2 > \cdots > x_{n-1}$ such that

$$T_n(x_j) = (-1)^j\,, \quad \lim_{x \to \infty} T_n(x) = 1\,, \quad \text{and} \quad \lim_{x \to -\infty} T_n(x) = (-1)^n\,.$$

d] There are real numbers $y_1 > y_2 > \cdots > y_{n-1}$ such that

$$U_n(y_j) = (-1)^{j+1} \quad \text{and} \quad \lim_{x \to \pm\infty} U_n(x) = 0\,.$$

e] $T_n(x)^2 + U_n(x)^2 = 1$ for every $x \in \mathbb{R}$.

f] Both T_n and U_n have exactly n simple zeros on $\mathbb{R}$, and the zeros of T_n and U_n strictly interlace.

g] The following statements are equivalent:

(i) There exists an $\alpha \in \mathbb{R}$ such that

$$V = (\cos\alpha)\, T_n + (\sin\alpha)\, U_n\,.$$

(ii) $V \in \mathcal{P}_n(a_1, a_2, \cdots, a_n)$ has uniform norm 1 on $\mathbb{R}$, and it equioscillates n times on the extended real line. That is, there are extended real numbers $\infty \geq z_1 > z_2 > \cdots > z_n > -\infty$ such that

$$V(z_j) = \pm(-1)^j\,, \qquad j = 1, 2, \ldots, n\,,$$

where

$$V(\infty) := \lim_{x\to\infty} V(x)\,.$$

h] With the notation of part g], V is strictly monotone on each of the intervals

$$(z_1, \infty),\ (z_2, z_1),\ \ldots,\ (z_n, z_{n-1}),\ (-\infty, z_n)\,.$$

E.6 Bernstein Factor on $\mathbb{R}$. Let $(a_k)_{k=1}^n \subset \mathbb{C} \setminus \mathbb{R}$ with

$$\mathrm{Im}(a_k) > 0\,, \qquad k = 1, 2, \ldots, n\,.$$

With the notation of E.5 let

$$B_n(x) := \frac{f_n'(x)}{f_n(x)}\,, \qquad x \in \mathbb{R}\,.$$

a] Show that

$$B_n(x) = \sum_{k=1}^{n} \frac{2\mathrm{Im}(a_k)}{|x - a_k|^2}\,, \qquad x \in \mathbb{R}\,.$$

b] Show that, on the real line,

$$T_n' = -B_n U_n\,, \qquad U_n' = B_n T_n\,,$$

and

$$(T_n')^2 + (U_n')^2 = B_n^2\,.$$

c] Show that

$$(V')^2 + B_n^2 V^2 = B_n^2$$

holds on the real line for every V of the form

$$V = (\cos\alpha)\, T_n + (\sin\alpha)\, U_n\,.$$

E.7 Coefficient Bounds in Nondense Rational Function Spaces. Suppose $(a_k)_{k=1}^{\infty} \subset \mathbb{R} \setminus [-1,1]$ is a sequence of distinct numbers satisfying

$$\sum_{j=1}^{\infty} \sqrt{1 - |a_j|^{-2}} < \infty \,.$$

Show that there are numbers $K_j > 0$ such that

$$|D_{j,n}| \le K_j \|p\|_{[-1,1]} \,, \qquad j = 0, 1, \ldots, n \,, \quad n \in \mathbb{N}$$

for every $p \in \mathcal{P}_n(a_1, a_2, \ldots, a_n)$ of the form

$$p(x) = D_{0,n} + \frac{D_{1,n}}{x - a_1} + \cdots + \frac{D_{n,n}}{x - a_n} \,, \qquad D_{j,n} \in \mathbb{R} \,.$$

Hint: Use E.2 c] of Section 3.3 and Theorem 3.5.4. □