2

Some Special Polynomials

Overview

Chebyshev polynomials are introduced and their central role in problems in the uniform norm on $[-1,1]$ is explored. Sequences of orthogonal functions are then examined in some generality, although our primary interest is in orthogonal polynomials (and rational functions). The third section of this chapter is concerned with orthogonal polynomials; it introduces the most classical of these. These polynomials satisfy many extremal properties, similar to those of the Chebyshev polynomials, but with respect to (weighted) L_2 norms. The final section of the chapter deals with polynomials with positive coefficients in various bases.

2.1 Chebyshev Polynomials

The ubiquitous Chebyshev polynomials lie at the heart of many analytic problems, particularly problems in $C[a,b]$, the space of real-valued continuous functions equipped with the uniform (supremum) norm, $\|\cdot\|_{[a,b]}$. Throughout this book, for any real- or complex-valued function f defined on $[a,b]$,

$$\|f\|_{[a,b]} := \sup_{x\in[a,b]} |f(x)| .$$

The Chebyshev polynomials are defined by

$$
\begin{aligned}
T_n(x) &:= \cos(n \arccos x)\,, && x \in [-1,1]\,,\\
&= \frac{1}{2}\left((x+\sqrt{x^2-1})^n + (x-\sqrt{x^2-1})^n\right)\,, && x \in \mathbb{C}, \qquad (2.1.1)\\
&= \frac{n}{2}\sum_{k=0}^{\lfloor n/2\rfloor} (-1)^k \frac{(n-k-1)!}{k!(n-2k)!}(2x)^{n-2k}\,, && x \in \mathbb{C}.
\end{aligned}
$$

These elementary equivalences are left for the reader (see E.1). The nth Chebyshev polynomial has the following equioscillation property on $[-1,1]$. There exist $n+1$ points $\zeta_i \in [-1,1]$ with $-1 = \zeta_n < \zeta_{n-1} < \cdots < \zeta_0 = 1$ so that

$$
T_n(\zeta_j) = (-1)^{n-j}\|T_n\|_{[-1,1]} = (-1)^{n-j}\,, \qquad j = 0,1,\ldots,n\,. \qquad (2.1.2)
$$

In other words $T_n \in \mathcal{P}_n$ takes the values $\pm\|T_n\|_{[-1,1]}$ with alternating sign the maximum possible number of times on $[-1,1]$. (These extreme points are just the points $\cos(k\pi/n)$, $k = 0,1,\ldots,n$.) The Chebyshev polynomial T_n satisfies the following extremal property:

Theorem 2.1.1. *We have*

$$
\min_{p \in \mathcal{P}^c_{n-1}} \|x^n - p(x)\|_{[-1,1]} = \|2^{1-n}T_n\|_{[-1,1]} = 2^{1-n}\,,
$$

where the minimum is uniquely attained by $p(x) = x^n - 2^{1-n}T_n(x)$.

Proof. Observe that, while the minimum is taken over $\mathcal{P}^c_{n-1}$, we need only consider $p \in \mathcal{P}_{n-1}$, since taking the real part of a $p \in \mathcal{P}^c_{n-1}$ can only improve the estimate. From the above formulas for T_n we have

$$
2^{1-n}T_n(x) = x^n + s(x)\,, \qquad s \in \mathcal{P}_{n-1}\,.
$$

Now suppose there exists $q \in \mathcal{P}_{n-1}$ with

$$
\|x^n - q(x)\|_{[-1,1]} < 2^{1-n}\,. \qquad (2.1.3)
$$

Then

$$
2^{1-n}T_n(x) - (x^n - q(x)) = s(x) + q(x) \in \mathcal{P}_{n-1}
$$

changes sign between any two consecutive extrema of T_n, hence it has at least n zeros in $(-1,1)$, and thus it must vanish identically. This contradicts (2.1.3), and we are done up to proving uniqueness (this is left as E.2). □

Comments, Exercises, and Examples.

The Chebyshev polynomials T_n are named after the versatile Russian mathematician, P. L. Chebyshev (1821–1894). The T comes from the spelling Tchebychef (or some such variant; there are many in the literature). A wealth of information on these polynomials may be found in Rivlin [90]. Throughout later sections of this book the Chebyshev polynomials will keep recurring. The initial exercises explore elementary properties of the Chebyshev polynomials.

Erdős [39] proved that for $t \in \mathcal{T}_n$ with $\|t\|_{\mathbb{R}} \leq 1$, the length of the graph of t on $[0, 2\pi]$ is the longest if and only if t is of the form $t(\theta) = \cos(n\theta + \alpha)$ with some $\alpha \in \mathbb{R}$ (see E.6). He conjectured that for any $p \in \mathcal{P}_n$ with $\|p\|_{[-1,1]} \leq 1$, the maximum arc length is attained by the nth Chebyshev polynomial T_n. This is proved in Bojanov [82b]. Kristiansen [79] also claims a proof. In E.9 the reducibility of T_n is considered, and in E.11 the basic properties of the transfinite diameter are established.

E.1 Basic Properties.

a] Establish the equivalence of the three representations of T_n given in equation (2.1.1).

Hint: $\cos n\theta = \frac{1}{2}[(\cos\theta + i\sin\theta)^n + ((\cos\theta - i\sin\theta)^n]$. To get the third representation, use E.3 b]. □

b] The zeros of T_n are precisely the points

$$x_k = \cos\tfrac{(2k-1)\pi}{2n}\,, \qquad k = 1, 2, \ldots, n\,.$$

c] The extrema of $T_n(x)$ in $[-1, 1]$ are precisely the points

$$\zeta_k = \cos\tfrac{k\pi}{n}\,, \qquad k = 0, 1, \ldots, n\,.$$

d] Observe that the zeros of T_n and T_{n+1} interlace, as do the extrema.

E.2 Uniqueness of the Minimum in Theorem 2.1.1. Prove the uniqueness of the minimum in Theorem 2.1.1.

Hint: Assume that $q \in \mathcal{P}^c_{n-1}$ and

$$\|x^n - q(x)\|_{[-1,1]} \leq 2^{1-n}\,.$$

Then

$$h(x) := 2^{1-n}T_n(x) - \mathrm{Re}(x^n - q(x))$$

defines a polynomial from $\mathcal{P}_{n-1}$ on $\mathbb{R}$ having at least n zeros (counted according to their multiplicities). Thus

$$2^{1-n}T_n(x) = \mathrm{Re}(x^n - q(x))\,, \qquad x \in \mathbb{R}\,,$$

which, together with the previous inequality, implies that $q(x)$ is real whenever $T_n(x) = \pm 1$. Now E.6 of Section 1.1 (Lagrange interpolation) yields that q has real coefficients. Hence

$$2^{1-n}T_n(x) = x^n - q(x)\,, \qquad x \in \mathbb{R}\,.$$

□

E.3 Further Properties of T_n.

a] Composition. Show that $T_{nm}(x) = T_n(T_m(x))$.

b] Three-Term Recursion. Show that

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)\,, \qquad n = 2, 3, \ldots\,.$$

c] Verify that

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ T_5(x) &= 16x^5 - 20x^3 + 5x\,. \end{aligned}$$

Note that T_n is even for n even and odd for n odd.

d] Another Formula for T_n. Show that $T_n(x) = \cosh(n \cosh^{-1}(x))$ for every $x \in \mathbb{R} \setminus [-1, 1]$.

e] Differential Equation. Show that

$$(1 - x^2)T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0\,.$$

f] An Identity. Show that

$$T_n(x) = \frac{T_{n+1}'(x)}{2n+2} - \frac{T_{n-1}'(x)}{2n-2}\,.$$

g] Orthogonality. Show that

$$\frac{2}{\pi}\int_{-1}^{1} \frac{T_n(x)T_m(x)dx}{\sqrt{1-x^2}} = \delta_{n,m} := \begin{cases} 0, & n \neq m \\ 1, & n = m > 0\,. \end{cases}$$

h] Generating Function. Show that

$$\frac{1 - yx}{1 - 2yx + y^2} = \sum_{n=0}^{\infty} T_n(x)y^n\,, \qquad x \in [-1, 1]\,, \quad |y| < 1\,.$$

Hint: Set $x = \cos\theta$ and sum. □

i] Another Representation of T_n. Show that

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

j] Another Identity. Show that

$$T_n\left(\tfrac{1}{2}(x+x^{-1})\right) = \tfrac{1}{2}(x^n + x^{-n}) \,.$$

E.4 Approximation to x^k on $[0,1]$.

a] Let $T_n^*(x) = T_n(2x-1)$ be the nth Chebyshev polynomial shifted to the interval $[0,1]$. Suppose

$$T_n^*(x) = \sum_{k=0}^{n} b_k x^k \,.$$

Show that for each $k = 0, 1, \ldots, n$,

$$\min_{c_j \in \mathbb{R}} \; \|x^k - \sum_{\substack{j=0 \\ j \neq k}}^{n} c_j x^j\|_{[0,1]} = \|b_k^{-1} T_n^*\|_{[0,1]} \,.$$

Hint: Proceed as in the proof of Theorem 2.1.1 and use E.6 of Section 1.3. □

b] Why does this not hold for T_n on $[-1,1]$?

E.5 A Composition Characterization. Suppose $(p_n)_{n=1}^{\infty}$ is a sequence of polynomials of degree n and for all positive integers n and m

$$p_n \circ p_m = p_{n \cdot m} \,.$$

Then there exists a linear transformation $w(x) = \alpha x + \beta$ so that

$$w \circ p_n \circ w^{-1} = x^n \,, \qquad n = 1, 2, \ldots$$

or

$$w \circ p_n \circ w^{-1} = T_n \,, \qquad n = 1, 2, \ldots \,.$$

This result is due to Block and Thielman [51]. The proof outlined in this exercise follows Rivlin [90].

a] Let

$$q(x) := a_0 + a_1 x + a_2 x^2\,, \quad a_2 \neq 0\,, \quad \text{and} \quad v(x) := \frac{x}{a_2} - \frac{a_1}{2a_2}\,.$$

Then

$$v^{-1}(q(v(x))) = x^2 + c \quad \text{with} \quad c := a_0 a_2 + (a_1/2) - (a_1^2/4)\,.$$

b] Let $q(x) = a_0 + a_1 x + a_2 x^2$, $a_2 \neq 0$. Then there is at most one polynomial p_n of degree exactly n so that

$$p_n(q(x)) = q(p_n(x))\,.$$

Hint: By a] we may assume $q(x) = x^2 + c$. Now suppose r, $s \in \mathcal{P}_n^c \backslash \mathcal{P}_{n-1}^c$,

$$r(x^2 + c) = r^2(x) + c$$

and

$$s(x^2 + c) = s^2(x) + c\,.$$

Then $u := r - s \in \mathcal{P}_{n-1}^c$ satisfies

$$u(x^2 + c) = u(x)(r(x) + s(x))$$

from which we deduce, by comparing degrees on both sides, that $n = 0$. (Note that the above conditions imply r and s monic.) □

c] Finish the proof of the initial statement of this exercise.

This is a special case of a more general theorem of Ritt [23] that classifies all rational functions r and s that commute in the sense that $r \circ s = s \circ r$.

d] Another Composition Characterization. Suppose $p \in \mathcal{P}_n$ has the property that the closure of the set

$$I_p := \{z \in \mathbb{C} : p^{[k]}(z) = 0 \text{ for some } k = 1, 2, \dots\}$$

is the interval $[-1, 1]$, where $p^{[k]}$ is the kth iterate of p, that is,

$$p^{[1]} := p \quad \text{and} \quad p^{[k]} := p \circ p^{[k-1]} \quad \text{for} \quad k = 2, 3, \dots\,.$$

Then $p(x) = \pm T_n(x)$.

e] Let

$$r_n(x) = \tan(n \tan^{-1}(x))\,.$$

Show that r_n is a rational function in $\mathcal{R}_{n,n}$, and observe that

$$r_n \circ r_m = r_{n \cdot m}\,.$$

E.6 Trigonometric Polynomials of Longest Arc Length. Theorem 5.1.3 (Bernstein-Szegő inequality) asserts that

$$t'(\theta)^2 + n^2 t^2(\theta) \leq n^2 \|t\|_{\mathbb{R}}^2$$

for every $t \in \mathcal{T}_n$ and $\theta \in \mathbb{R}$. Use this to prove the following result of Erdős. For $t \in \mathcal{T}_n$ with $\|t\|_{\mathbb{R}} \leq 1$, the length of the graph of t on $[0, 2\pi]$ is the longest if and only if it is of the form $t(\theta) = \cos(n\theta + \alpha)$ for some $\alpha \in \mathbb{R}$.

Hint: Suppose $t \in \mathcal{T}_n$ with $\|t\|_{\mathbb{R}} = 1$. Let $s(\theta) := \cos n\theta$. If

$$-1 < t(\theta_1) = s(\theta_2) < 1$$

holds, then by the Bernstein-Szegő inequality (see also E.5 of Section 5.1)

$$|t'(\theta_1)| \leq n(1 - t^2(\theta_1))^{1/2} = n(1 - s^2(\theta_2))^{1/2} = |s'(\theta_2)|,$$

and if equality holds for one pair of θ_1, θ_2, then it holds for all pairs, and $t(\theta) \equiv \cos(n\theta + \alpha)$ for some $\alpha \in \mathbb{R}$. Suppose $t_n(\theta) \not\equiv \cos(n\theta + \alpha)$. Let τ and σ be monotone arcs of the graphs of $y = t(\theta)$ and $y = s(\theta)$, respectively, with endpoints of each having the same ordinates y_1 and y_2. Let $|\tau|$ and $|\sigma|$ be the length of τ and σ, respectively, and let $|\tau_x|$ and $|\sigma_x|$ be the length of the projection of τ and σ, respectively, on the x-axis. Show that

$$|\tau| < |\sigma| + (|\tau_x| - |\sigma_x|)$$

by approximating τ and σ by a polygonal line corresponding to a subdivision of the interval with endpoints, y_1 and y_2 on the y-axis. □

E.7 Monic Polynomials with Minimal Norm on an Interval.

a] The unique monic polynomial $p \in \mathcal{P}_n^c$ minimizing $\|p\|_{[a,b]}$ is given by

$$p(x) = 2\left(\frac{b-a}{4}\right)^n T_n\left(\frac{2x-a-b}{b-a}\right).$$

b] Let $0 < a < b$. Find all monic polynomials $p \in \mathcal{P}_n^c$ minimizing

$$\|p\|_{[-b,-a]\cup[a,b]}.$$

(For two intervals of different lengths this is a much harder problem. The problem was originally due to Zolotarev and is solved in terms of elliptic functions. See Todd [88], Fischer [92], and Peherstorfer [87].)

E.8 Lower Bound for the Norm of Polynomials on the Unit Disk. Let D be the open unit disk of $\mathbb{C}$. Show that

$$\|a_0 + a_1 z + \cdots + a_n z^n\|_D \geq \max_{0 \leq k \leq n} |a_k|$$

for arbitrary complex numbers $a_0, a_1, \ldots, a_n$. Thus z^n plays the role of the nth Chebyshev polynomial on the unit disk.

Hint: If $p(z) := a_0 + a_1 z + \cdots + a_n z^n$, then

$$a_m = \frac{1}{2\pi i} \int_{\partial D} \frac{p(z)}{z^{m+1}}\, dz\,.$$

□

The next exercise supposes some familiarity with the rudiments of reducibility over $\mathbb{Q}$ and basic properties, such as irreducibility of cyclotomic polynomials over $\mathbb{Q}$ (see Clark [71]). Details of the following observation of Schur's are in Rivlin [90].

E.9 On the Reducibility of T_n over $\mathbb{Q}$. Let $n \in \mathbb{N}$ be fixed.

a] The zeros of $T_n(x/2)$ are all of the form

$$x_j := e^{(2j-1)i\pi/(2n)} + e^{-(2j-1)i\pi/(2n)}\,, \qquad j = 1, 2, \ldots, n\,.$$

b] If $n \geq 3$ and ζ is a primitive nth root of unity, then $\zeta + \zeta^{-1}$ is of degree $\varphi(n)/2$. (Here φ is the Euler φ function.)

c] Thus if T_n is irreducible over $\mathbb{Q}$, then n must be a power of 2.

d] For a positive integer h, let

$$F_h(x) := \prod_{\substack{j=1 \\ \gcd(2j-1,2n)=h}}^{n} (x - x_j)\,.$$

(Here $\gcd(m,n)$ denotes the greatest common divisor of m and n.) Show that if h is odd, then F_h is irreducible over $\mathbb{Q}$.

e] **The Factorization of T_n.**

$$2T_n(x/2) = \prod_{h|n\,,\ h \text{ odd}} F_h(x)\,.$$

So if n is odd, T_n has $\varphi(n)$ factors, while if n is even, then T_n has $\varphi(m)$ factors, where m is the largest odd divisor of n.

f] Let $n \geq 3$ be odd. Then $T_n(x)/x$ is irreducible over $\mathbb{Q}$ if and only if n is prime.

E.10 Chebyshev Polynomials of the Second Kind. Let the *Chebyshev polynomials of the second kind* be defined by

$$U_{n-1}(x) := \frac{1}{n}T_n'(x) = \frac{\sin n\theta}{\sin\theta}\,, \qquad x = \cos\theta\,.$$

a] $U_n(x) = 2T_n(x) + U_{n-2}(x)\,.$

b] $T_n(x) = U_n(x) - xU_{n-1}(x)\,.$

c] $U_n(x) = \sum_{k=0}^{n} x^k T_{n-k}(x)\,.$

d] $U_n(x) = \dfrac{\left(x+\sqrt{x^2-1}\right)^{n+1} - \left(x-\sqrt{x^2-1}\right)^{n+1}}{2\sqrt{x^2-1}}\,.$

e] Orthogonality. Show that

$$\frac{2}{\pi}\int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2}\,dx = \delta_{n,m} := \begin{cases} 0\,, & n \neq m \\ 1\,, & n = m > 0\,. \end{cases}$$

f] Three-Term Recursion. Show that

$$U_0(x) = 1\,, \quad U_1(x) = 2x\,,$$

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

(Note that this is the same recursion as for T_n.)

g] The Coefficients of U_n. Show that

$$U_n(x) = \sum_{k=0}^{\lfloor n/2\rfloor} (-1)^k \binom{n-k}{k} (2x)^{n-2k}\,.$$

h] Another Form of U_n. Show that

$$U_n(x) = \sum_{k=0}^{\lfloor n/2\rfloor} (-1)^k \binom{n+1}{2k+1} x^{n-2k}(x^2-1)^k\,.$$

The concepts of transfinite diameter and capacity play a central role in potential theory, harmonic analysis, and other areas of mathematics.

E.11 Transfinite Diameter. Let E be a compact subset of $\mathbb{C}$. Let

$$\Delta_n(E) := \max_{z_1,\dots,z_n \in E} \prod_{\substack{1 \le i,j \le n \\ i \ne j}} |z_i - z_j| .$$

The points z_i at which the above maximum are obtained are called nth *Fekete points* for E. If the points z_i are the nth Fekete points for E, then the polynomial

$$q_n(z) := \prod_{i=1}^{n} (z - z_i)$$

is called an nth (monic) *Fekete polynomial* for E. The *transfinite diameter* or *logarithmic capacity* of E is defined by

$$\operatorname{cap}(E) := \lim_{n \to \infty} (\Delta_n(E))^{\frac{1}{n(n-1)}} ,$$

where the limit exists by part c] (below).

a] Let $z_1, z_2, \dots, z_n$ be nth Fekete points for E. Then

$$(\Delta_n(E))^{1/2} = \operatorname{abs} \begin{vmatrix} 1 & z_1 & \dots & z_1^{n-1} \\ 1 & z_2 & \dots & z_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_n & \dots & z_n^{n-1} \end{vmatrix} .$$

Hint: See E.2 b] (Vandermonde determinant) of Section 3.2. □

b] Let $q_n(z) := \prod_{i=1}^{n} (z - z_i)$ be an nth Fekete polynomial for E. Let

$$m_n := \min_{i=1,\dots,n} |q_n'(z_i)| \quad \text{and} \quad M_n := \|q_n\|_E .$$

Then

$$M_n \le \left(\frac{\Delta_{n+1}(E)}{\Delta_n(E)} \right)^{1/2} \le m_{n+1} \le (\Delta_{n+1}(E))^{\frac{1}{n+1}} .$$

Outline. We have

$$|q_n(z)|^2 \Delta_n(E) = \prod_{i=1}^{n} |z - z_i|^2 \prod_{\substack{1 \le i,j \le n \\ i \ne j}} |z_i - z_j| \le \Delta_{n+1}(E)$$

and

$$\begin{aligned} \Delta_{n+1}(E) &= \prod_{\substack{1 \le i,j \le n+1 \\ i \ne k, j \ne k, i \ne j}} |z_i - z_j| \prod_{\substack{1 \le i \le n+1 \\ i \ne k}} |z_k - z_i|^2 \\ &\le \Delta_n(E) \prod_{\substack{1 \le i \le n+1 \\ i \ne k}} |z_k - z_i|^2 . \end{aligned}$$

From the first line above,

$$M_n^2 \Delta_n(E) \le \Delta_{n+1}(E)\,.$$

From the second line above,

$$\Delta_{n+1}(E) \le \Delta_n(E) m_{n+1}^2\,.$$

□

c] Show that $(\Delta_n(E))^{\frac{1}{n(n-1)}}$ is decreasing, so the limit exists in the definition of $\mathrm{cap}(E)$.

d] The Fekete points lie on the boundary of E. So $\mathrm{cap}(E) = \mathrm{cap}(\partial(E))$.

Hint: Use the maximum principle (see E.1 d] of Section 1.2). □

e] If $E \subset F$, where $F \subset \mathbb{C}$ is also compact, then $\mathrm{cap}(E) \le \mathrm{cap}(F)$.

f] **Chebyshev Constants.** Let

$$\mathcal{M}_n := \left\{ p \in \mathcal{P}_n^c : \;\; p(z) = \prod_{j=1}^{n} (z - z_j)\,, \quad z_j \in \mathbb{C} \right\}$$

and

$$\widetilde{\mathcal{M}}_n := \left\{ p \in \mathcal{P}_n^c : \;\; p(z) = \prod_{j=1}^{n} (z - z_j)\,, \quad z_j \in E \right\}.$$

Let

$$\mu_n(E) := \inf\{\|p\|_E : p \in \mathcal{M}_n\}$$

and

$$\widetilde{\mu}_n(E) := \inf\{\|p\|_E : p \in \widetilde{\mathcal{M}}_n\}\,.$$

Show that the infimum in the definition of $\mu_n(E)$ and $\widetilde{\mu}_n(E)$ is actually minimum. Show also that

$$\mu_{n+m}(E) \le \mu_n(E)\mu_m(E)$$

and

$$\widetilde{\mu}_{n+m}(E) \le \widetilde{\mu}_n(E)\widetilde{\mu}_m(E)$$

for any two nonnegative integers n and m. Finally show that the above inequalities imply that

$$\mu(E) := \lim_{n\to\infty} (\mu_n(E))^{1/n} \quad \text{and} \quad \widetilde{\mu}(E) := \lim_{n\to\infty} (\widetilde{\mu}_n(E))^{1/n}$$

exist.

The numbers $\mu(E)$ and $\widetilde{\mu}(E)$ are called the *Chebyshev constant* and *modified Chebyshev constant*, respectively, associated with E. Obviously $\mu(E) \le \widetilde{\mu}(E)$.

g] Transfinite Diameter and Chebyshev Constants Are the Same:

$$\operatorname{cap}(E) = \mu(E) = \widetilde{\mu}(E)\,.$$

Proof. Without loss of generality we may assume that E contains infinitely many points. Part b] yields $\widetilde{\mu}(E) \le \operatorname{cap}(E)$. Therefore, since $\mu(E) \le \widetilde{\mu}(E)$, it is sufficient to prove that $\operatorname{cap}(E) \le \mu(E)$. Note that if $p \in \mathcal{M}_n$ and $z_1, z_2, \ldots, z_{n+1} \in E$ are the $(n+1)$th Fekete points for E, then

$$(\Delta_{n+1}(E))^{1/2} = \operatorname{abs} \begin{vmatrix} 1 & z_1 & \ldots & z_1^{n-1} & p(z_1) \\ 1 & z_2 & \ldots & z_2^{n-1} & p(z_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & z_{n+1} & \ldots & z_{n+1}^{n-1} & p(z_{n+1}) \end{vmatrix}\,.$$

Expanding the above determinant with respect to its last column, we obtain

$$\begin{aligned}(\Delta_{n+1}(E))^{1/2} &\le (\Delta_n(E))^{1/2} \sum_{j=1}^{n+1} |p(z_j)| \\ &\le (n+1)(\Delta_n(E))^{1/2} \|p\|_E\,,\end{aligned}$$

so

$$(\Delta_{n+1}(E))^{1/2} \le (n+1)(\Delta_n(E))^{1/2} \mu_n(E)\,.$$

For the sake of brevity let

$$c_n := \left((n+1)^2(\mu_n(E))^2\right)^{1/n} \quad \text{and} \quad d_n := (\Delta_n(E))^{\frac{2}{n(n-1)}}\,.$$

Then

$$d_{n+1}^{n+1} \le c_n d_n^{n-1}\,.$$

Since E contains infinitely many points, $c_n > 0$ and $d_n > 0$ hold for each $n = 2, 3, \ldots$. Multiplying the above inequalities for $n = 1, 2, \ldots, k$, we obtain after simplification that

$$(d_2 d_3 \cdots d_{k+1})^{\frac{1}{k-1}} (d_{k+1})^{\frac{k}{k-1}} \le (d_2)^{\frac{2}{k-1}} (c_2 c_3 \cdots c_k)^{\frac{1}{k-1}}\,.$$

Since $\lim\limits_{k\to\infty} d_k = \operatorname{cap}(E)$ and $\lim\limits_{k\to\infty} c_k = (\mu(E))^2$, we conclude

$$(\operatorname{cap}(E))^2 \le (\mu(E))^2\,,$$

which finishes the proof. □

h] Show that $\operatorname{cap}([a,b]) = \frac{1}{4}(b-a)$.

i] Show that $\operatorname{cap}(\overline{D}_\rho) = \rho$, where

$$\overline{D}_\rho := \{z \in \mathbb{C} : |z| \le \rho\}.$$

j] Show that $\operatorname{cap}(A_\alpha) = \sin(\alpha/4)$, where A_α is an arc of the unit circle C of length α, $0 \le \alpha \le 2\pi$.

Hint: Without loss of generality we may assume that the arc A_α is symmetric with respect to the x-axis and $1 \in A_\alpha$. Now use part h] and the transformation $x = \frac{1}{2}(z + z^{-1})$. □

2.2 Orthogonal Functions

The most basic properties of orthogonal functions are explored in this section. The following section specializes the discussion to polynomials.

In this section the functions are complex-valued and the vector spaces are over the complex numbers. All the results have obvious real analogs and in many later applications we will restrict to these corresponding real cases.

An *inner product* on a vector space V is a function $\langle \cdot, \cdot \rangle$ from $V \times V$ to $\mathbb{C}$ that satisfies, for all $f, g, h \in V$ and $\alpha, \beta \in \mathbb{C}$,

(2.2.1) $\langle f, f \rangle > 0$ unless $f = 0$ (*positivity*)

(2.2.2) $\langle f, g \rangle = \overline{\langle g, f \rangle}$ (*conjugate symmetry*)

(2.2.3) $\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$ (*linearity*).

A vector space V equipped with an inner product is called an *inner product space*. It is a normed linear space with the norm $\| \cdot \| := \langle \cdot, \cdot \rangle^{1/2}$.

The canonical example for us will be the space $C[a,b]$ of all complex-valued continuous functions on $[a,b]$ with the inner product

$$\langle f, g \rangle := \int_a^b f(x)\overline{g(x)}w(x)\,dx\,, \tag{2.2.4}$$

where $w(x)$ is a nonnegative integrable function on $[a,b]$ that is positive except possibly on a set of measure zero. It is a normed linear space with the norm

$$\|f\|_{L_2(w)} := \langle f, f \rangle^{1/2} = \left(\int_a^b |f(x)|^2 w(x)\,dx \right)^{1/2}. \tag{2.2.5}$$

More generally, if (X, μ) is a measure space (with μ nonnegative), then

$$\langle f, g\rangle := \int_X f(x)\overline{g(x)}\, d\mu(x) \tag{2.2.6}$$

is an inner product on the space $L_2(\mu)$ of square integrable functions. More precisely, $L_2(\mu)$ denotes the space of equivalence classes of measurable functions for which

$$\|f\|_{L_2(\mu)} := \langle f, f\rangle^{1/2} = \left\{\int_X |f(x)|^2\, d\mu(x)\right\}^{1/2}$$

is finite. The equivalence classes are defined by the equivalence relation $f \sim g$ if $f = g$ μ-almost everywhere on X.

If V is a vector space equipped with an inner product $\langle\cdot,\cdot\rangle$, then a metric ρ can be defined on V by $\rho(f, g) := \langle f - g, f - g\rangle^{1/2}$. The fact that this ρ is a metric on V is an immediate consequence of (2.2.1) and Theorem 2.2.1 b]. If this metric space (V, ρ) is complete (that is, if every Cauchy sequence in (V, ρ) converges to some $x \in V$), then V is called a *Hilbert space*.

It can be shown that $L_2(\mu)$ is a Hilbert space for every measure space (X, μ) (see Rudin [87]), while $C[a, b]$ equipped with the inner product (2.2.4), where $w(x) \equiv 1$, is not a Hilbert space (see E.1).

When we write $L_2[a, b]$ we always mean $L_2(\mu)$ where μ is the Lebesgue measure on $X = [a, b]$. The fact that the inner product gives a norm is part of the next theorem.

Theorem 2.2.1. *If $(V, \langle\cdot,\cdot\rangle)$ is an inner product space equipped with the norm $\|\cdot\| := \langle\cdot,\cdot\rangle^{1/2}$, then for all $f, g \in V$,*

a] $|\langle f, g\rangle| \le \|f\|\,\|g\|$ *Cauchy-Schwarz inequality*

b] $\|f + g\| \le \|f\| + \|g\|$ *triangle inequality*

c] $\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2$ *parallelogram law.*

Proof. Let $f, g \in V$ be arbitrary. To prove the Cauchy-Schwarz inequality, without loss of generality we may assume $\langle g, g\rangle = 1$ and we may assume $\langle f, g\rangle$ is real (why?). Let $\zeta := \langle f, g\rangle$ and note that by (2.2.1) and (2.2.3),

$$\begin{aligned} 0 \le \langle f - \zeta g, f - \zeta g\rangle &= \langle f, f\rangle - 2\zeta\langle f, g\rangle + \zeta^2\langle g, g\rangle \\ &= \|f\|^2 - \langle f, g\rangle^2, \end{aligned}$$

which finishes the proof of part a].

Using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|f + g\|^2 = \langle f + g, f + g\rangle &= \langle f, f\rangle + 2\mathrm{Re}(\langle f, g\rangle) + \langle g, g\rangle \\ &\le \|f\|^2 + 2\|f\|\,\|g\| + \|g\|^2 \\ &\le (\|f\| + \|g\|)^2, \end{aligned}$$

which is the triangle inequality.

The parallelogram law follows from

$$\begin{aligned}\|f+g\|^2 + \|f-g\|^2 \\ &= \langle f,f\rangle + 2\,\mathrm{Re}(\langle f,g\rangle) + \langle g,g\rangle + \langle f,f\rangle - 2\,\mathrm{Re}(\langle f,g\rangle) + \langle g,g\rangle\,.\end{aligned}$$

□

For the space $L_2(\mu)$ of all square integrable functions, the Cauchy-Schwarz inequality becomes

$$\left|\int_a^b f\overline{g}\,d\mu\right| \le \left(\int_a^b |f|^2\,d\mu(x)\right)^{1/2}\left(\int_a^b |g|^2\,d\mu\right)^{1/2}\,.$$

Applying this with f and g replaced by $|f|$ and $|g|$, we obtain

$$\int_a^b |fg|\,d\mu \le \left(\int_a^b |f|^2\,d\mu(x)\right)^{1/2}\left(\int_a^b |g|^2\,d\mu\right)^{1/2}\,. \tag{2.2.7}$$

A collection of vectors $\{f_\alpha : \alpha \in \mathcal{A}\}$ in an inner product space $(V, \langle\cdot,\cdot\rangle)$ is said to be *orthogonal* if

$$\langle f_\alpha, f_\beta\rangle = 0\,, \qquad \alpha,\beta \in \mathcal{A}\,,\ \ \alpha \ne \beta\,. \tag{2.2.8}$$

If $\langle f_\alpha, f_\beta\rangle = 0$, then we write $f_\alpha \perp f_\beta$. The collection is called *orthonormal* if, in addition to being orthogonal,

$$\langle f_\alpha, f_\alpha\rangle = 1\,, \qquad \alpha \in \mathcal{A}\,. \tag{2.2.9}$$

An orthogonal collection $\{f_\alpha : \alpha \in \mathcal{A}\}$ of nonzero vectors in an inner product space can always be orthonormalized as $\{\|f_\alpha\|^{-1} f_\alpha : \alpha \in \mathcal{A}\}$. The vector space over $\mathbb{C}$ generated by $\{f_\alpha : \alpha \in \mathcal{A}\}$ is denoted by

$$\mathrm{span}\{f_\alpha : \alpha \in \mathcal{A}\}\,.$$

So $\mathrm{span}\{f_\alpha : \alpha \in \mathcal{A}\}$ is just the set of all finite linear combinations

$$\left\{\sum_{i=1}^n c_i f_{\alpha_i} : \alpha_i \in \mathcal{A}\,,\ \ c_i \in \mathbb{C}\,,\ \ n \in \mathbb{N}\right\}\,.$$

Any linearly independent collection of vectors can be orthonormalized, as the next theorem shows.

Theorem 2.2.2 (Gram-Schmidt). *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space with norm $\|\cdot\| := \langle\cdot,\cdot\rangle^{1/2}$. Suppose $\{f_i\}_{i=1}^{\infty}$ is a linearly independent collection of vectors in V. Let*

$$g_1 := \frac{f_1}{\|f_1\|}$$

and (inductively) let

$$u_n := f_n - \sum_{k=1}^{n-1} \langle f_n, g_k\rangle g_k \qquad \text{and} \qquad g_n := \frac{u_n}{\|u_n\|}\,.$$

Then $\{g_n\}_{n=1}^{\infty}$ is an orthonormal collection, and for each n,

$$\operatorname{span}\{g_1, g_2, \ldots, g_n\} = \operatorname{span}\{f_1, f_2, \ldots, f_n\}\,.$$

Proof. This can be proved easily by induction where the inductive step is: for $m < n$,

$$\begin{aligned}\langle u_n, g_m\rangle &= \langle f_n, g_m\rangle - \sum_{k=1}^{n-1} \langle f_n, g_k\rangle\langle g_k, g_m\rangle \\ &= \langle f_n, g_m\rangle - \sum_{k=1}^{n-1} \langle f_n, g_k\rangle \delta_{k,m} = 0\,.\end{aligned}$$

□

The key approximation theoretic property orthonormal sets have is encapsulated in the following result:

Theorem 2.2.3 (Best Approximation by Linear Combinations). *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space with norm $\|\cdot\| := \langle\cdot,\cdot\rangle^{1/2}$. Suppose $\{f_1, \ldots, f_n\}$ is an orthonormal collection of vectors in V. Let $f \in V$. Then*

$$\min_{c_i \in \mathbb{C}} \Big\| \sum_{i=1}^{n} c_i f_i - f \Big\|$$

is attained if and only if

$$c_i = \langle f, f_i\rangle\,, \qquad i = 1, 2, \ldots, n\,.$$

In other words, the sum $\sum_{i=1}^{n} \langle f, f_i\rangle f_i$ is the best approximation to f from $\operatorname{span}\{f_1, \ldots, f_n\}$ in the norm $\langle\cdot,\cdot\rangle^{1/2}$.

Proof. Fix $f \in V$, and let c_i be as above. Let

$$g := \sum_{i=1}^{n} c_i f_i$$

and let $h \in \text{span}\{f_1, \ldots, f_n\}$. Note that

$$(g - f) \perp f_i, \qquad i = 1, 2, \ldots, n$$

since by orthonormality

$$\langle g - f, f_i \rangle = \sum_{j=1}^{n} c_j \langle f_j, f_i \rangle - c_i = 0.$$

Thus

$$(g - f) \perp (h - g)$$

and so

$$\begin{aligned} \|h - f\|^2 &= \|(h - g) + (g - f)\|^2 \\ &= \|h - g\|^2 + 2\text{Re}(\langle h - g, g - f \rangle) + \|g - f\|^2 \\ &= \|h - g\|^2 + \|g - f\|^2 \\ &\geq \|g - f\|^2 \end{aligned}$$

with strict inequality unless $h = g$. This finishes the proof. □

Note that the above theorem gives the following corollary:

Corollary 2.2.4. *If $\{f_1, \ldots, f_n\}$ is an orthonormal collection, then every*

$$g \in \text{span}\{f_1, \ldots, f_n\}$$

can be written as

$$g = \sum_{i=1}^{n} \langle g, f_i \rangle f_i .$$

Comments, Exercises, and Examples.

The theory of orthogonal functions, and in particular orthogonal polynomials, is old and far-reaching. As we will see in the next section, the names associated with the classical orthogonal polynomials including Chebyshev, Laguerre, Legendre, Hermite, Jacobi, and Stieltjes, are the "who's who" of nineteenth century analysis. Various aspects of this beautiful body of theory are explored in the exercises of this and the next section.

Much of this material is available in G. Szegő's [75] classical treatise "Orthogonal Polynomials." Of course, orthogonal polynomials are intimately connected to Fourier series and parts of harmonic and functional analysis generally. The standard functional analysis in the following exercises is available in many sources. See, for example, Rudin [73, 87].

E.1 $C[0,1]$ **Is Not a Hilbert Space.** Construct a sequence of continuous functions $(f_n)_{n=1}^{\infty}$ on $[0,1]$ for which

$$\|f_n - f\|_{L_2[0,1]} \to 0$$

with some$f \notin C[0,1]$ (in the sense that f cannot be modified on a set of measure zero to be in $C[0,1]$).

So $C[0,1]$ equipped with the inner product (2.2.4), where $[a,b] = [0,1]$ and $w(x)$ is identically 1, is not a Hilbert space. It can be shown that there is no way of putting a norm on $C[0,1]$ that preserves the uniform topology and makes $C[0,1]$ into a Hilbert space, essentially because $C[0,1]$ is not reflexive (see Rudin [73], Chapter 4). This, in fact, shows that $C[0,1]$ is not isomorphic to $L_p[0,1]$ for any $p \in (1,\infty)$. For the definition of $L_p[0,1]$, see E.7.

E.2 On $L_2(w)$**.** Consider

$$\langle f, g\rangle = \int_a^b f(x)g(x)w(x)\,dx\,.$$

What conditions on w guarantee that $\langle f, g\rangle$ is an inner product on $C[a,b]$?

E.3 Cauchy-Schwarz Inequality for Sequences. Show that

$$\left|\sum_{i=1}^{n} \alpha_i\beta_i\right|^2 \le \left(\sum_{i=1}^{n} |\alpha_i|^2\right)\left(\sum_{i=1}^{n} |\beta_i|^2\right)$$

for all $\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_n \in \mathbb{C}$. Equality holds if and only if there exists a $\gamma \in \mathbb{C}$ so that either $\alpha_i = \gamma\overline{\beta}_i$ for each i or $\beta_i = \gamma\overline{\alpha}_i$ for each i.

Hint: $\mathbb{C}^n$ is a Hilbert space with inner product

$$\langle(\alpha_1,\alpha_2,\ldots,\alpha_n),(\beta_1,\beta_2,\ldots,\beta_n)\rangle = \sum_{i=1}^{n} \alpha_i\overline{\beta}_i\,.$$

□

E.4 Bessel's Inequality. Let $(V,\langle\cdot,\cdot\rangle)$ be an inner product space with norm $\|\cdot\| := \langle\cdot,\cdot\rangle^{1/2}$. Suppose $\{f_i\}_{i=1}^{\infty}$ is a countable collection of orthonormal vectors in V.

a] Show that

$$\sum_{i=1}^{\infty} |\langle f_i, f\rangle|^2 \le \|f\|^2\,.$$

Hint: With $h := 0$ in the last expression of the proof of Theorem 2.2.3

$$\|f\|^2 = \|g\|^2 + \|g - f\|^2\,.$$

□

b] Suppose

$$f = \sum_{i=1}^{\infty} \langle f_i, f \rangle f_i$$

in the sense that the partial sums of the right-hand side converge to f in the norm $\|\cdot\|$. Show that

$$\|f\|^2 = \sum_i |\langle f_i, f \rangle|^2 .$$

E.5 The Kernel Function. Let $\{p_i\}_{i=0}^n$ be a collection of orthonormal functions in $L_2[a,b]$ with respect to the inner product defined by (2.2.6), where $X := [a,b]$. Define the *kernel function* by

$$K_n(x_0, x) := \overline{p_0(x_0)} p_0(x) + \overline{p_1(x_0)} p_1(x) + \cdots + \overline{p_n(x_0)} p_n(x) .$$

a] **Reproducing Property.** If $q \in \text{span}\{p_0, \ldots, p_n\}$, then

$$\int_a^b K_n(t,x) q(t)\, d\mu(t) = q(x) .$$

Hint: Expand q in terms of $p_0, \ldots, p_n$ as in Corollary 2.2.4. □

b] $(K_n(x_0,x_0))^{-1/2} K_n(x_0,x)$ solves the following maximization problem:

$$\max \left\{ |q(x_0)| : q \in \text{span}\{p_0, p_1, \ldots, p_n\} \text{ and } \int_a^b |q(x)|^2\, d\mu(x) = 1 \right\} .$$

Outline. Write $q = \sum_{i=0}^n c_i p_i$. Then, as in E.4 b],

$$\|q\|_{L_2(\mu)}^2 = |c_0|^2 + |c_1|^2 + \cdots + |c_n|^2 = 1 .$$

The Cauchy-Schwarz inequality of E.3 yields that

$$|q(x_0)|^2 \le \left(\sum_{i=0}^n |c_i|^2 \right) \left(\sum_{i=0}^n |p_i(x_0)|^2 \right) = K_n(x_0, x_0) .$$

However, if

$$c_i = \frac{\overline{p_i(x_0)}}{\left(\sum_{j=0}^n |p_j(x_0)|^2 \right)^{1/2}} ,$$

so

$$q(x) = \frac{K_n(x_0, x)}{(K_n(x_0, x_0))^{1/2}} ,$$

then equality holds in the above inequality. □

c] Show, as in a], that if $q \in \text{span}\{p_0, \ldots, p_n\}$ and $p_0, \ldots, p_n$ are m times differentiable at x_0, then

$$|q^{(m)}(x_0)| \le \left(\sum_{k=0}^m |p_k^{(m)}(x_0)|^2 \right)^{1/2} \|q\|_{L_2(\mu)} .$$

When does equality hold?

E.6 Completeness. Let $\{f_\alpha : \alpha \in \mathcal{A}\}$ be an orthonormal collection in a Hilbert space H. The collection $\{f_\alpha : \alpha \in \mathcal{A}\}$ is called a *maximal orthonormal set* in H if there is no $f \neq 0$ so that $\langle f, f_\alpha \rangle = 0$ for every $\alpha \in \mathcal{A}$.

The following statements are equivalent:

(1) The set of all finite linear combinations of $f_\alpha, \alpha \in \mathcal{A}$, is dense in H.

(2) $\|f\|^2 = \sum_{\alpha \in \mathcal{A}} |\langle f_\alpha, f \rangle|^2$ for all $f \in H$.

(3) $\langle f, g \rangle = \sum_{\alpha \in \mathcal{A}} \langle f_\alpha, f \rangle \overline{\langle f_\alpha, g \rangle}$ for all $f, g \in H$.

(4) $\{f_\alpha : \alpha \in \mathcal{A}\}$ is a maximal orthonormal set in H.

If any of the above holds, then the orthonormal collection is called a *complete orthonormal system.* (See, for example, Rudin [87].)

a] Deduce (1) $\Rightarrow$ (2) from Theorem 2.2.3.

b] Deduce (2) $\Rightarrow$ (3) from the simple identity

$$4\langle f, g \rangle = \|f + g\|^2 - \|f - g\|^2 + i\|f + ig\|^2 - i\|f - ig\|^2 .$$

The above identity is called *polarization.*

c] Prove (3) $\Rightarrow$ (4).

d] Prove (4) $\Rightarrow$ (1) by contradiction.

Equality (3) is called *Parseval's identity.*

The remaining exercises assumes some familiarity with measure theory.

E.7 Basic Theory of L_p Spaces. Let (X, μ) be a measure space (μ is nonnegative) and $p \in (0, \infty]$. The space $L_p(\mu)$ is defined as the collection of equivalence classes of measurable functions for which $\|f\|_{L_p(\mu)} < \infty$, where

$$\|f\|_{L_p(\mu)} := \left(\int_X |f|^p \, d\mu \right)^{1/p}, \qquad p \in (0, \infty)$$

and

$$\|f\|_{L_\infty(\mu)} := \sup\{\alpha \in \mathbb{R} : \mu(\{x \in X : |f(x)| > \alpha\}) > 0\} < \infty .$$

In any of the cases the equivalence classes are defined by the equivalence relation $f \sim g$ if $f = g$ μ-almost everywhere on X. When we write $L_p[a, b]$ we always mean $L_p(\mu)$, where μ is the Lebesgue measure on $X = [a, b]$. The notations $L_p(a, b), L_p[a, b)$, and $L_p(a, b]$ are also used analogously to $L_p[a, b]$.

a] Hölder's Inequality. Suppose $1 \le p < q \le \infty$ and $p^{-1} + q^{-1} = 1$. Show that

$$\left| \int_X f g \, d\mu \right| \le \|f\|_{L_p(\mu)} \|g\|_{L_q(\mu)}$$

for every $f \in L_p(\mu)$ and $g \in L_q(\mu)$.

If $1 < p, q < \infty$, then equality holds if and only if $\alpha |f|^p = \beta |g|^q$ μ-almost everywhere on X for some $\alpha, \beta \in \mathbb{R}$ with $\alpha^2 + \beta^2 > 0$, and there is a $c \in \mathbb{C}$ with $|c| = 1$ so that cfg is nonnegative μ-almost everywhere on X.

Hölder's inequality was proved by Rogers [1888] before Hölder [1889] proved it independently.

Hint: If the right-hand side is 0, then the inequality is obvious. If it is different from 0, then let

$$F := \frac{|f|}{\|f\|_{L_p(\mu)}} \quad \text{and} \quad G := \frac{|g|}{\|g\|_{L_q(\mu)}} .$$

If $x \in X$ is such that $0 < F(x) < \infty$ and $0 < G(x) < \infty$, then there are real numbers s and t such that

$$F(x) = e^{s/p} \quad \text{and} \quad G(x) = e^{t/q} .$$

Use the convexity of the exponential function to show that

$$e^{s/p + t/q} \le p^{-1} e^s + q^{-1} e^t .$$

Apply this with the above choices of s and t, and integrate both sides on X with respect to μ. □

b] Minkowski's Inequality for $p \in [1, \infty]$. Let $p \in [1, \infty]$. Show that

$$\|f + g\|_{L_p(\mu)} \le \|f\|_{L_p(\mu)} + \|g\|_{L_p(\mu)}$$

for every $f, g \in L_p(\mu)$.

If $1 < p, q < \infty$, then equality holds if and only if $\alpha f = \beta g$ μ-almost everywhere on X for some $\alpha, \beta \in \mathbb{R}$ with $\alpha^2 + \beta^2 > 0$.

Hint: The cases $p = 1$ and $p = \infty$ are straightforward. Let $p \in (1, \infty)$. Then

$$|f + g|^p \le |f| \, |f + g|^{p-1} + |g| \, |f + g|^{p-1}$$

and apply Hölder's inequality (part a]) to each term separately. □

By part b], $L_p(\mu)$ is a vector space and $\|\cdot\|_{L_p(\mu)}$ is a norm on $L_p(\mu)$ whenever $p \in [1,\infty]$. If $p \in (0,1)$, then $\|\cdot\|_{L_p(\mu)}$ is still called a norm in the literature, however, for $p \in (0,1)$ the subadditive property, in general, fails. In fact, if $p \in (0,1)$, then $\|\cdot\|_{L_p[a,b]}$ is superadditive for Riemann integrable functions in $L_p[a,b]$; see Pólya and Szegő [76].

c] Assume $\mu(X) < \infty$. Show that $L_q(\mu) \subset L_p(\mu)$ for every $0 < p < q \le \infty$. If $\mu(X) \le 1$, then prove that

$$\|f\|_{L_p(\mu)} \le \|f\|_{L_q(\mu)}$$

for every measurable function f.

d] Assume $f \in L_q(\mu)$ for some $q > 0$. Show that

$$\lim_{p\to\infty} \|f\|_{L_p(\mu)} = \|f\|_{L_\infty(\mu)} .$$

e] Riesz-Fischer Theorem. *Show that if* $1 \le p \le \infty$, *then* $(L_p(\mu), \rho)$ *is a complete metric space, where*

$$\rho(f,g) := \|f-g\|_{L_p(\mu)} .$$

Hint: Use the monotone convergence theorem and Minkowski's inequality (part b]); see Rudin [87] for details. □

If $p \in [1,\infty]$, then $q \in [1,\infty]$ defined by $p^{-1} + q^{-1} = 1$ is called *conjugate* to p.

f] Bounded Linear Functionals on $L_p(\mu)$. Let $1 \le p < \infty$ and $g \in L_q(\mu)$, where q is the conjugate exponent to p. Show that

$$\Phi_g(f) := \int_X f\,g\,d\mu$$

is a bounded linear functional on $L_p(\mu)$.

Hint: Use Hölder's inequality (part a]). □

g] Riesz Representation Theorem. *Suppose* $1 \le p < \infty$, μ *is (*σ*-) finite and* Φ *is a bounded linear functional on* $L_p(\mu)$. *Then there is a unique* $g \in L_q(\mu)$, *where* q *is the conjugate exponent to* p, *so that*

$$\Phi(f) = \int_X f\,g\,d\mu\,, \qquad f \in L_p(\mu)\,.$$

Moreover, if Φ *and* g *are related as above, then*

$$\|\Phi\| := \max\{\Phi(f) : f \in L_p(\mu),\ \|f\|_{L_p(\mu)} = 1\} = \|g\|_{L_q(\mu)} .$$

Proof. See, for example, Rudin [87] or Royden [88]. □

If X is a locally compact Hausdorff space, then the characterization of bounded linear functionals on the space $C_c(X)$ of continuous functions with compact support equipped with the uniform norm is also known as the Riesz representation theorem, and its proof may be found in, for example, Rudin [87].

h] Orthogonality in $L_p(\mu)$. Suppose $1 \le p < \infty, \mu$ is (σ-)finite, Y is a finite-dimensional subspace of $L_p(\mu)$. The function $f \in L_p(\mu)$ is said to be *orthogonal* to Y in $L_p(\mu)$, written $f \perp Y$, if

$$\|f\|_{L_p(\mu)} \le \|f + h\|_{L_p(\mu)}$$

for every $h \in Y$. Show that an element $f \in L_p(\mu)$ is orthogonal to Y if and only if

$$\int_X |f|^{p-1} \overline{\operatorname{sign}}(f) h \, d\mu = 0$$

for every $h \in Y$, where

$$\overline{\operatorname{sign}}(f(x)) := \begin{cases} \dfrac{\overline{f(x)}}{|f(x)|} & \text{if } f(x) \neq 0 \\ 0 & \text{if } f(x) = 0\,. \end{cases}$$

Outline. Suppose that the integral vanishes for every $h \in Y$. Let q be the conjugate exponent to p. defined by $p^{-1} + q^{-1} = 1$. Observe that

$$g := |f|^{p-1} \, \overline{\operatorname{sign}}(f) \in L_q(\mu)$$

and

$$\int_X |g|^q d\mu = \int_X |f|^p \, d\mu\,.$$

Without loss of generality we may assume that $\|f\|_{L_p(\mu)} = 1$. Then for every $h \in Y$, Hölder's inequality yields that

$$\begin{aligned} \|f\|_{L_p(\mu)} = 1 = \int_X fg \, d\mu &= \int_X (f + h) g \, d\mu \\ &\le \|f + h\|_{L_p(\mu)} \|g\|_{L_q(\mu)} = \|f + h\|_{L_p(\mu)}\,, \end{aligned}$$

proving that $f \perp Y$. (Observe that this argument is also valid for $p = 1$.)

Suppose now $f \perp Y$. Without loss of generality we may assume that $f \notin Y$. By a standard corollary to the Hahn-Banach theorem (see, for example, Rudin [87]), there exists a linear functional M on $L_p(\mu)$ such that $M(f) = 1$, $M(h) = 0$ for every $h \in Y$, and $\|M\| = \|f\|^{-1}_{L_p(\mu)}$. This M is then representable by some element $g \in L_q(\mu)$, that is,

$$M(f) = \int_X fg \, d\mu, \qquad \|g\|_{L_p(\mu)} = \|f\|^{-1}_{L_p(\mu)}$$

(see part g]). Therefore

$$\int_X f\,g\,d\mu = \|f\|_{L_p(\mu)}\,\|g\|_{L_q(\mu)}\,,$$

and by the conditions for equality to hold in Hölder's inequality,

$$g(x)f(x) \geq 0 \qquad a.e.\ [\mu]\ \text{on}\ X$$

and

$$|g(x)|^q = \lambda |f(x)|^p \qquad a.e.\ [\mu]\ \text{on}\ X$$

for a suitable constant $\lambda > 0$. Hence

$$g(x) = \lambda |f(x)|^{p-1}\,\overline{\text{sign}}(f(x))\,,$$

and so $M(h) = 0$, $h \in Y$, implies

$$\int_X |f|^{p-1}\,\overline{\text{sign}}(f)h\,d\mu = 0$$

for every $h \in Y$. □

The statement of part h] remains valid for closed subspaces Y instead of finite-dimensional subspaces, see, for example, Shapiro [71].

i] Minkowski's Inequality for $p \in (0,1)$. Show that

$$\|f+g\|_{L_p(\mu)} \leq 2^{1/p-1}\left(\|f\|_{L_p(\mu)} + \|g\|_{L_p(\mu)}\right)$$

for every $f, g \in L_p(\mu)$ and $p \in (0,1)$.

Hint: Verify that

$$\begin{aligned}
\|f+g\|_{L_p(\mu)} &\leq \left(\int_X (|f|+|g|)^p\,d\mu\right)^{1/p} \\
&\leq \left(\int_X |f|^p\,d\mu + \int_X |g|^p\,d\mu\right)^{1/p} \\
&\leq 2^{1/p-1}\left(\left(\int_X |f|^p\,d\mu\right)^{1/p} + \left(\int_X |g|^p\,d\mu\right)^{1/p}\right)
\end{aligned}$$

whenever $f, g \in L_p(\mu)$ and $p \in (0,1)$. □

Further properties of $L_p(\mu)$ spaces may be found in Rudin [87].

E.8 Fourier Series.

a] Show that

$$\left\{ \frac{e^{in\theta}}{\sqrt{2\pi}} : n \in \mathbb{Z} \right\}$$

is a maximal orthonormal collection in $L_2[-\pi, \pi]$.

Hint: The orthonormality is obvious. In order to show the maximality, first note that $L_2[-\pi, \pi]$ is a Hilbert space by the Riesz-Fischer theorem (E.7 e]). Hence, by E.6, it is sufficient to show that the set $\mathcal{T}^c$ of all complex trigonometric polynomials is dense in $L_2[-\pi, \pi]$. By the Stone-Weierstrass theorem (E.2 of Section 4.1) $\mathcal{T}^c$ is dense in $C^*[-\pi, \pi]$, where $C^*[-\pi, \pi]$ is the space of all complex-valued 2π-periodic continuous functions on $\mathbb{R}$ equipped with the uniform norm on $\mathbb{R}$. Finally, it is a standard measure theoretic argument to show that $C^*[-\pi, \pi]$ is dense in $L_2[-\pi, \pi]$; see, for example, Rudin [87]. □

The kth *Fourier coefficient* $\widehat{f}(k)$ of a function $f \in L_1[-\pi, \pi]$ is defined by

$$\widehat{f}(k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} \, d\theta \, .$$

The (formal) *Fourier series* of a function $f \in L_1[-\pi, \pi]$ is defined by

$$f \sim \sum_{k=-\infty}^{\infty} \widehat{f}(k) e^{ik\theta} \, .$$

The functions

$$S_n(\theta) := \sum_{k=-n}^{n} \widehat{f}(k) e^{ik\theta}$$

are called the nth *partial sums* of the Fourier series of f.

b] Show that if $f \in L_2[-\pi, \pi]$, then

$$\sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 = \|f\|^2_{L_2[-\pi,\pi]} < \infty \, .$$

Hint: Use part a], E.6, and E.7 e]. □

c] Show that if $f \in L_2[-\pi, \pi]$, then

$$\lim_{n \to \infty} \|f - S_n\|_{L_2[-\pi,\pi]} = 0 \, ,$$

so f is the $L_2[-\pi, \pi]$ limit of the partial sums of its formal Fourier series.

Hint: Use part b]. □

Carleson, in 1966, solved Luzin's problem by showing that $S_n \to f$ almost everywhere on $[-\pi,\pi]$ for every $f \in L_2[-\pi,\pi]$. Earlier, Kolmogorov showed that there is a function $f \in L_1[-\pi,\pi]$ so that S_n diverges everywhere on $[-\pi,\pi]$.

d] Isometry of $L_2[-\pi,\pi]$ and ℓ_2^c. Let

$$\ell_2^c := \left\{ \mathbf{x} = (x_k)_{k=-\infty}^{\infty} : x_k \in \mathbb{C}, \quad \sum_{k=-\infty}^{\infty} |x_k|^2 < \infty \right\}$$

and

$$\|\mathbf{x}\|_{\ell_2^c} := \left(\sum_{k=-\infty}^{\infty} |x_k|^2 \right)^{1/2}, \qquad \mathbf{x} = (x_k)_{-\infty}^{\infty}, \quad x_k \in \mathbb{C}.$$

Show that the function $I : L_2[-\pi,\pi] \to \ell_2^c$ defined by

$$I(f) := \widehat{f} := (\widehat{f}(k))_{-\infty}^{\infty}$$

is one-to-one and onto, and

$$\|I(f)\|_{\ell_2^c} = \|f\|_{L_2[-\pi,\pi]}.$$

Hint: Use part a] to show that I is one-to-one. Use the Riesz-Fischer theorem (E.7 e]) to show that I is onto. The norm-preserving property is the content of part b]. □

Part d] shows that the structure of $L_2[-\pi,\pi]$ is the same as that of ℓ_2^c. Hence $L_2[-\pi,\pi]$ is a separable Hilbert space, that is, it has a countable dense subset. So if $\epsilon > 0$ is fixed, then any collection $\{f_\alpha : \alpha \in \mathcal{A}\}$ from $L_2[-\pi,\pi]$ for which

$$\|f_\alpha - f_\beta\|_{L_2[-\pi,\pi]} \geq \epsilon, \qquad \alpha,\beta \in \mathcal{A}, \ \ \alpha \neq \beta$$

must be countable.

e] The Riemann-Lebesgue Lemma. *If $f \in L_1[-\pi,\pi]$, then $\widehat{f}(k) \to 0$ as $k \to \infty$.*

Hint: First prove it for step functions, then extend the result to every $f \in L_1[-\pi,\pi]$ by using the fact that step functions form a dense set in $L_1[-\pi,\pi]$. □

f] Show that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}, \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^{2k}} = r_k \pi^{2k}$$

for every $k = 3, 4, \cdots$, where r_k is a rational number.

Hint: Let f be the 2π periodic function defined by

$$f(\theta) := \left(\frac{\pi - \theta}{2} \right)^k, \qquad \theta \in [0, 2\pi).$$

Apply part b]. □

E.9 Denseness of Polynomials in $L_2(\mu)$ on $\mathbb{R}$.

a] Let μ be a finite Borel measure on $[a,b]$ and $f \in L_1(\mu)$. Show that if

$$\int_a^b f(x)e^{itx}\,d\mu(x) = 0\,, \qquad t = \frac{2\pi k}{b-a}\,, \quad k = 0, \pm 1, \pm 2, \dots\,,$$

then $f(x) = 0$ a.e. $[\mu]$ on $[a,b]$.

Outline. Use the fact that the set $\mathcal{T}^c$ of all complex trigonometric polynomials is dense in $C^*[-\pi,\pi]$ (see the hint given for E.8 a]) and standard measure theoretic arguments to show that the assumption of part a] implies

$$\int_a^b f(x)g(x)\,d\mu(x) = 0$$

for every bounded measurable function g defined on $[a,b]$. Now, choosing

$$g(x) := \overline{\operatorname{sign}}(f(x)) := \begin{cases} \dfrac{\overline{f(x)}}{|f(x)|} & \text{if } f(x) \neq 0 \\ 0 & \text{if } f(x) = 0\,, \end{cases}$$

we obtain

$$\int_{-\pi}^{\pi} |f(x)|\,d\mu(x) = 0\,,$$

and the result follows. □

b] Let μ be a finite Borel measure on $\mathbb{R}$ and $f \in L_1(\mu)$. Show that if

$$\int_{\mathbb{R}} f(x)e^{itx}\,d\mu(x) = 0\,, \qquad t \in \mathbb{R}\,,$$

then $f(x) = 0$ a.e. $[\mu]$ on $\mathbb{R}$.

Hint: Use part a] to show that the assumption of part b] implies

$$\int_{\mathbb{R}} f(x)g(x)\,d\mu(x) = 0$$

for every bounded measurable function defined on $\mathbb{R}$ (first assume that g has compact support, then eliminate this assumption). Finish the proof as in part a]. □

c] Let μ be a Borel measure on $\mathbb{R}$ satisfying

$$\int_{\mathbb{R}} e^{r|x|}\,d\mu(x) < \infty$$

with some $r > 0$. Show that the set $\mathcal{P}^c$ of all complex algebraic polynomials is dense in $L_2(\mu)$.

Outline. First observe that the assumption on μ implies $\mathcal{P}^c \subset L_2(\mu)$. The fact that $L_2(\mu)$ is a Hilbert space (see E.7 e]), Theorem 2.2.2 (Gram-Schmidt), and E.6 imply that it is sufficient to prove that if $f \in L_2(\mu)$ and

$$\int_{\mathbb{R}} f(x)x^k \, d\mu(x) = 0\,, \qquad k = 0, 1, 2, \dots\,,$$

then $f(x) = 0$ a.e. $[\mu]$ on $\mathbb{R}$. Assume that $f \in L_2(\mu)$ satisfies the above orthogonality relation. Use Theorem 2.2.1 a] (Cauchy-Schwarz inequality) to show that

$$F(t) := \int_{\mathbb{R}} f(x)e^{-itx} \, d\mu(x)$$

is well-defined on $\mathbb{R}$. For every $t_0 \in \mathbb{R}$, we have

$$\begin{aligned} f(x)e^{-itx} &= f(x)e^{-it_0x}e^{-i(t-t_0)x} \\ &= \sum_{k=0}^{\infty} (-i)^k \frac{(t-t_0)^k}{k!} f(x)e^{-it_0x}x^k\,. \end{aligned}$$

Note that if $|t - t_0| \le r/2$, then the integral of the right-hand side with respect to $\mu(x)$ on $\mathbb{R}$ can be calculated by integrating term by term since

$$\begin{aligned} &\sum_{k=0}^{\infty} \int_{\mathbb{R}} \left| (-i)^k \frac{(t-t_0)^k}{k!} f(x)e^{-it_0x}x^k \right| \, d\mu(x) \\ &\quad \le \sum_{k=0}^{\infty} \int_{\mathbb{R}} \frac{|t-t_0|^k}{k!} |f(x)||x|^k \, d\mu(x) \\ &\quad = \int_{\mathbb{R}} |f(x)| e^{|t-t_0||x|} \, d\mu(x) \\ &\quad \le \left| \int_{\mathbb{R}} |f(x)|^2 \, d\mu(x) \int_{\mathbb{R}} e^{2|t-t_0||x|} \, d\mu(x) \right|^{1/2} < \infty\,. \end{aligned}$$

Therefore, if $|t - t_0| \le r/2$, then

$$F(t) = \sum_{k=0}^{\infty} (-i)^k \frac{(t-t_0)^k}{k!} \int_{\mathbb{R}} f(x)e^{-it_0x}x^k \, d\mu(x)\,.$$

This means that F has a Taylor series expansion about every $t_0 \in \mathbb{R}$ with radius of convergence at least $r/2$. Also, with the choice $t_0 = 0$, by the assumed orthogonality relations, we have $F(t) = 0$ whenever $|t| \le r/2$. We can now deduce that $F(t) = 0$ for every $t \in \mathbb{R}$. Hence it follows from part b] that $f(x) = 0$ a.e. $[\mu]$ on $\mathbb{R}$. □

2.3 Orthogonal Polynomials

The classical orthogonal polynomials arise on orthogonalizing the sequence

$$(1, x, x^2, \dots)$$

with respect to various particularly nice weights, $w(x)$, on an interval, which, after a linear transformation, may be taken to be one of $[-1,1]$, $[0,\infty)$, or $(-\infty,\infty)$. The main examples we consider are the *Jacobi polynomials*

$$P_n^{(\alpha,\beta)}(x)\,,\ \text{where } w(x) := (1-x)^\alpha(1+x)^\beta \text{ on } [-1,1]\,,\ \ \alpha,\beta > -1\,. \tag{2.3.1}$$

When $\alpha = \beta = -1/2$ the Jacobi polynomials are the *Chebyshev polynomials of the first kind*,

$$T_n(x)\,,\ \text{where } w(x) := (1-x^2)^{-1/2} \text{ on } [-1,1]\,. \tag{2.3.2}$$

When $\alpha = \beta = 1/2$ they are the *Chebyshev polynomials of the second kind*,

$$U_n(x)\,,\ \text{where } w(x) := (1-x^2)^{1/2} \text{ on } [-1,1]\,. \tag{2.3.3}$$

Another special case of importance is $\alpha = \beta = 0$, which gives the *Legendre polynomials*,

$$P_n(x)\,,\ \text{where } w(x) = 1 \text{ on } [-1,1]\,. \tag{2.3.4}$$

The *Laguerre polynomials* are

$$L_n(x)\,,\ \text{where } w(x) := e^{-x} \text{ on } [0,\infty)\,. \tag{2.3.5}$$

The *Hermite polynomials* are

$$H_n(x)\,,\ \text{where } w(x) := e^{-x^2} \text{ on } (-\infty,\infty)\,. \tag{2.3.6}$$

The above notation is traditionally used to denote orthogonal polynomials with a standard normalization; see the exercises. It is not usually the case that this normalization gives orthonormality. All of these much studied polynomials arise naturally, as do all the special functions, in the study of differential equations. We catalog some of the special properties of these classical orthogonal polynomials in the exercises.

In general, a nondecreasing bounded function α (typically the distribution function of a finite measure) defined on $\mathbb{R}$ is called an *m-distribution* if it takes infinitely many distinct values, and its *moments*, that is, the improper Stieltjes integrals

$$\int_{-\infty}^{\infty} x^n \, d\alpha(x) = \lim_{\substack{\omega_1 \to -\infty \\ \omega_2 \to +\infty}} \int_{\omega_1}^{\omega_2} x^n \, d\alpha(x) \,,$$

exist and are finite for $n = 0, 1, \ldots$.

Theorem 2.3.1 (Existence and Uniqueness of Orthonormal Polynomials). *For every m-distribution α there is a unique sequence of polynomials $(p_n)_{n=0}^{\infty}$ with the following properties:*

(i) $$p_n(x) = \gamma_n x^n + r_{n-1}(x) \,, \qquad \gamma_n > 0 \,, \quad r_{n-1} \in \mathcal{P}_{n-1} \,,$$

(ii) $$\int_{\mathbb{R}} p_n(x) p_m(x) \, d\alpha(x) = \delta_{n,m} = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m \,. \end{cases}$$

Proof. The result follows from Theorem 2.2.2 (Gram-Schmidt). Note that the defining property of an m-distribution α ensures that

$$\langle p, q \rangle := \int_{\mathbb{R}} pq \, d\alpha$$

is an inner product on $\mathcal{P}_n^c$. □

The sequence $(p_n)_{n=0}^{\infty}$ defined by Theorem 2.3.1 is called the *sequence of orthonormal polynomials* associated with an m-distribution α. The sequence $(q_n)_{n=0}^{\infty}$ is called a *sequence of orthogonal polynomials* associated with an m-distribution α if

$$q_n = c_n p_n \,, \quad 0 \neq c_n \in \mathbb{C} \,, \quad n = 0, 1, \ldots \,,$$

where $(p_n)_{n=0}^{\infty}$ is the sequence of orthonormal polynomials associated with α. The support $\operatorname{supp}(\alpha)$ of an m-distribution α is defined as the closure of the set

$$\{x \in \mathbb{R} : \alpha \text{ is increasing at } x\} \,.$$

If α is absolutely continuous on $\mathbb{R}$, then

$$d\alpha(x) = w(x) \, dx \quad \text{with some} \quad 0 \leq w \in L_1(\infty, \infty)$$

in which case α may be identified as a nonnegative weight function $w \in L_1(-\infty, \infty)$ whose integral takes infinitely many distinct values. If (a, b) is an interval and $w \in L_1[a, b]$ has an integral that takes infinitely many distinct values, then the sequence of orthogonal (orthonormal) polynomials associated with

$$\widetilde{w}(x) = \begin{cases} w(x) & \text{if } x \in (a, b) \\ 0 & \text{if } x \notin (a, b) \end{cases}$$

is said to be *orthogonal (orthonormal) with respect to the weight w.*

One thing distinguishing orthogonal polynomials from general orthogonal systems is the existence of a three-term recursion.

Theorem 2.3.2 (Three-Term Recursion). *Suppose $(p_n)_{n=0}^{\infty}$ is a sequence of orthonormal polynomials with respect to an m-distribution α. Then*

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1} p_{n-1}(x)\,, \qquad n = 0, 1, \dots\,,$$

where

$$p_{-1} := 0\,, \quad a_{-1} = 0\,, \quad a_n = \frac{\gamma_n}{\gamma_{n+1}} > 0\,, \quad b_n \in \mathbb{R}\,, \quad n = 0, 1, \dots$$

(γ_n is the leading coefficient of p_n).

This theorem has a converse due to Favard [35]; see E.12.

Proof. Since $xp_n(x) \in \mathcal{P}_{n+1}$, we may write

$$xp_n(x) = \sum_{k=0}^{n+1} d_k p_k(x)\,, \qquad d_k \in \mathbb{R}\,. \tag{2.3.7}$$

For notational convenience, let

$$\langle p, q\rangle := \int_{\mathbb{R}} p(x)q(x)\,d\alpha(x)$$

for any two polynomials p and q. Since $\langle p_n, q\rangle = 0$ for every $q \in \mathcal{P}_{n-1}$, we have

$$\langle xp_n(x), q(x)\rangle = \langle p_n(x), xq(x)\rangle = 0$$

for every $q \in \mathcal{P}_{n-2}$. In particular,

$$\langle xp_n(x), p_k(x)\rangle = 0\,, \qquad k = 0, 1, \dots\,, n-2\,.$$

On the other hand, using (2.3.7) and the orthonormality of $(p_n)_{n=0}^{\infty}$, we obtain

$$\langle xp_n(x), p_k(x)\rangle = d_k\langle p_k, p_k\rangle = d_k\,.$$

Hence $d_k = 0$ for each $k = 0, 1, \dots\,, n-2$ and

$$xp_n(x) = d_{n+1}p_{n+1}(x) + d_n p_n(x) + d_{n-1}p_{n-1}(x)\,. \tag{2.3.8}$$

Here the lead coefficient of the left-hand side polynomial is γ_n, while the lead coefficient of the right-hand side polynomial is $d_{n+1}\gamma_{n+1}$, so

$$a_n := d_{n+1} = \gamma_n/\gamma_{n+1}\,.$$

In order to show that $a_{n-1} := d_{n-1} = \gamma_{n-1}/\gamma_n$, note that (2.3.8) and the orthonormality of $(p_n)_{n=0}^{\infty}$ imply

$$\begin{aligned}
0 &= \langle p_{n+1}, p_{n-1}\rangle \\
&= \frac{1}{d_{n+1}}\langle x p_n(x), p_{n-1}(x)\rangle - \frac{d_n}{d_{n+1}}\langle p_n, p_{n-1}\rangle - \frac{d_{n-1}}{d_{n+1}}\langle p_{n-1}, p_{n-1}\rangle \\
&= \frac{1}{d_{n+1}}\langle p_n(x), \gamma_{n-1}x^n\rangle - \frac{d_{n-1}}{d_{n+1}} \\
&= \frac{1}{d_{n+1}}\frac{\gamma_{n-1}}{\gamma_n} - \frac{d_{n-1}}{d_{n+1}}\,.
\end{aligned}$$

Hence

$$a_{n-1} := d_{n-1} = \frac{\gamma_{n-1}}{\gamma_n}\,.$$

□

Theorem 2.3.3 (Christoffel-Darboux Formula). *With the notation of the previous theorem,*

$$\sum_{k=0}^{n} p_k(x)p_k(y) = \frac{\gamma_n}{\gamma_{n+1}}\left(\frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x-y}\right)$$

for all $x \neq y \in \mathbb{C}$.

Proof. Theorem 2.3.2 (three-term recursion) yields that

$$\begin{aligned}
\Delta_k &:= p_{k+1}(x)p_k(y) - p_k(x)p_{k+1}(y) \\
&= \frac{1}{a_k}(x-y)p_k(x)p_k(y) + \frac{a_{k-1}}{a_k}(p_k(x)p_{k-1}(y) - p_{k-1}(x)p_k(y))\,.
\end{aligned}$$

So

$$a_k\frac{\Delta_k}{x-y} = p_k(x)p_k(y) + a_{k-1}\frac{\Delta_{k-1}}{x-y}\,,$$

and we sum the above from 0 to n to get the desired formula. □

Corollary 2.3.4. *In the notation of Theorem 2.3.2*

$$\sum_{k=0}^{n} p_k^2(x) = \frac{\gamma_n}{\gamma_{n+1}}\left(p'_{n+1}(x)p_n(x) - p'_n(x)p_{n+1}(x)\right)\,.$$

Proof. Let $y \to x$ in Theorem 2.3.3. □

We can deduce quite easily from this that orthogonal polynomials associated with an m-distribution α have real interlacing zeros lying in the interior of the smallest interval containing $\mathrm{supp}(\alpha)$; see E.1 and E.2.

Comments, Exercises, and Examples.

Askey, in comments following an outline of the history of orthogonal polynomials by Szegő [82, vol. III], writes:

"The classical orthogonal polynomials are mostly attributed to someone other than the person who introduced them. Szegő refers to Abel and Lagrange and Tschebyscheff in [75, chapter 5] for work on the Laguerre polynomials $L_n^0(x)$. Abel's work was published posthumously in 1881. Probably the first published work on these polynomials that uses their orthonormality was by Murphy (1833). Hermite polynomials were studied extensively by Laplace in connection with work on probability theory. Hermite's real contribution to these polynomials was to introduce Hermite polynomials in several variables. Lagrange came across the recurrence relation for Legendre polynomials."

Perhaps this is not very surprising given the many diverse ways in which these polynomials can arise.

There are many sources for the basic properties of orthogonal polynomials. In particular, Askey and Ismail [84], Chihara [78], Erdélyi et al. [53], Freud [71], Nevai [79b], [86], Szegő [75], and, in tabular form, Abramowitz and Stegun [65] are such sources. Exercises include a treatment of the elementary properties of the most familiar orthogonal polynomials. The connections linking orthogonal polynomials, the moment problem, and Favard's converse theorem to the three-term recursion are also examined in the exercises.

E.1 Simple Real Zeros. Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. Show that each p_n has exactly n simple real zeros lying in the interior of the smallest interval containing $\operatorname{supp}(\alpha)$.

Hint: Suppose the statement is false. Then p_n has at most $n-1$ sign changes on $[a,b]$, hence there exists $0 \neq q \in \mathcal{P}_{n-1}$ so that

$$p_n(x)q(x) \geq 0\,, \qquad x \in [a,b]\,.$$

Show that this contradicts the orthogonality relation

$$0 = \int_{\mathbb{R}} p_n(x)q(x)\,d\alpha(x) = \int_a^b p_n(x)q(x)\,d\alpha(x)\,.$$

□

E.2 Interlacing of Zeros. Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. Then the zeros of p_n and p_{n+1} strictly interlace. That is, there is exactly one zero of p_n strictly between any two consecutive zeros of p_{n+1}.

Hint: From Corollary 2.3.4,

$$p'_{n+1}(x)p_n(x) - p'_n(x)p_{n+1}(x)$$

is positive on $\mathbb{R}$. Since p_{n+1} has $n+1$ simple real zeros (see E.1), we see that if γ and δ are two consecutive zeros of p_{n+1}, then

$$\operatorname{sign}(p'_{n+1}(\gamma)) = -\operatorname{sign}(p'_{n+1}(\delta))\,,$$

and hence

$$\operatorname{sign}(p_n(\gamma)) = -\operatorname{sign}(p_n(\delta))\,.$$

□

E.3 Orthogonality of $(K_n(x_0,x))_{n=0}^{\infty}$. Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. Let

$$x_0 < \min\operatorname{supp}(\alpha) \quad \text{or} \quad x_0 > \max\operatorname{supp}(\alpha)\,.$$

Let $(K_n(x_0,x))_{n=0}^{\infty}$ be the sequence of associated kernel functions (as in E.5 of Section 2.2). Show that

$$\int_{\mathbb{R}} K_n(x_0,x)K_m(x_0,x)|x-x_0|\,d\alpha(x) = 0\,,$$

for any two nonnegative integers $n \neq m$.

E.4 Hypergeometric Functions. We introduce the following standard notation: the *rising factorial (or Pochammer symbol)*

$$(a)_n := a(a+1)\cdots(a+n-1)\,, \qquad (a)_0 := 1$$

for $a \in \mathbb{C}$ and $n = 1,2,\ldots$; the *binomial coefficient*

$$\binom{a}{n} := \frac{a(a-1)\cdots(a-n+1)}{n!}\,, \qquad \binom{a}{0} := 1$$

for $a \in \mathbb{C}$ and $n = 1,2,\ldots$; and the *Gaussian hypergeometric series*

$${}_2F_1(a,b\,;c\,;z) := F(a,b\,;c\,;z) := \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!}$$

for $a,b,c \in \mathbb{C}$.

a] For $\operatorname{Re}(c) > \operatorname{Re}(b) > 0$,

$$F(a,b\,;c\,;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,dt\,,$$

where Γ is, as usual, the *gamma function* defined by

$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t}\, dt\,, \qquad \mathrm{Re}(z) > 0\,.$$

Proof. See, for example, Szegő [75]. □

b] Hypergeometric Differential Equation. The function $y = F(a, b\,; c\,; z)$ satisfies

$$z(1-z)\,\frac{d^2y}{dz^2} + [c - (a+b+1)z]\,\frac{dy}{dz} - aby = 0\,.$$

Proof. See, for example, Szegő [75]. □

In E.5, E.6, and E.7 we catalog some of the basic properties of some of the classical orthogonal polynomials. Proofs are available in Szegő [75], for example.

E.5 Jacobi Polynomials.

a] Rodrigues' Formula. Let

$$P_n^{(\alpha,\beta)}(x) := (-1)^n \frac{2^{-n}}{n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\left[(1-x)^\alpha(1+x)^\beta(1-x^2)^n\right].$$

Then $(P_n^{(\alpha,\beta)})_{n=0}^\infty$ is a sequence of orthogonal polynomials on $[-1,1]$ associated with the weight function

$$w(x) := (1-x)^\alpha(1+x)^\beta\,, \qquad -1 < \alpha, \beta < \infty\,.$$

That is,

$$P_n^{(\alpha,\beta)} \in \mathcal{P}_n \quad \text{and} \quad \int_{-1}^1 P_n^{(\alpha,\beta)} P_m^{(\alpha,\beta)} (1-x)^\alpha (1+x)^\beta\, dx = 0$$

for any two nonnegative integers $n \neq m$.

In the rest of the exercise, the polynomials $P_n^{(\alpha,\beta)}$ are as in part a].

b] Normalization. We have

$$P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n} = \frac{(\alpha+1)_n}{n!}$$

and

$$\int_{-1}^1 (P_n^{(\alpha,\beta)}(x))^2 (1-x)^\alpha (1+x)^\beta\, dx$$
$$= \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}\,.$$

c] Explicit Form.

$$\begin{aligned} P_n^{(\alpha,\beta)}(x) &= \frac{1}{2^n}\sum_{m=0}^{n}\binom{n+\alpha}{m}\binom{n+\beta}{n-m}(x-1)^{n-m}(x+1)^m \\ &= \sum_{m=0}^{n}\binom{\alpha+n}{n-m}\binom{\alpha+\beta+n+m}{m}\left(\frac{x-1}{2}\right)^m \\ &= \binom{n+\alpha}{n}\,{}_2F_1\left(-n, n+\alpha+\beta+1\,;\alpha+1\,;\frac{1-x}{2}\right). \end{aligned}$$

d] Differential Equation. The function $y = P_n^{(\alpha,\beta)}(x)$ satisfies

$$(1-x^2)\frac{d^2y}{dx^2} + [\beta-\alpha-(\alpha+\beta+2)x]\frac{dy}{dx} + n(n+\alpha+\beta+1)y = 0\,.$$

e] Recurrence Relation. The sequence $(P_n^{(\alpha,\beta)}(x))_{n=0}^{\infty}$ satisfies

$$D_n P_{n+1}^{(\alpha,\beta)}(x) = (A_n + B_n x)P_n^{(\alpha,\beta)}(x) - C_n P_{n-1}^{(\alpha,\beta)}(x)\,,$$

where

$$P_0^{(\alpha,\beta)} = 1 \qquad \text{and} \qquad P_1^{(\alpha,\beta)}(x) = \tfrac{1}{2}[\alpha-\beta+(\alpha+\beta+2)x]$$

and

$$\begin{aligned} D_n &= 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta) \\ A_n &= (2n+\alpha+\beta+1)(\alpha^2-\beta^2) \\ B_n &= (2n+\alpha+\beta+2)(2n+\alpha+\beta+1)(2n+\alpha+\beta) \\ C_n &= 2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)\,. \end{aligned}$$

f] Generating Function.

$$\sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)z^n = \frac{2^{\alpha+\beta}}{R(1-z+R)^{\alpha}(1+z+R)^{\beta}}\,,$$

where $R = \sqrt{1-2xz+z^2}$.

There are various special cases, some of which we have previously defined. The Legendre polynomials P_n are defined by

$$P_n := P_n^{(0,0)}\,, \qquad n = 0, 1, \ldots\,.$$

The Chebyshev polynomials T_n defined in Section 2.1 satisfy

$$T_n = \frac{4^n}{\binom{2n}{n}} P_n^{(-1/2,\,-1/2)}\,, \qquad n = 0, 1, \ldots\,.$$

The *ultraspherical* (or *Gegenbauer*) *polynomials* $C_n^{(\alpha)}$ are defined by

$$C_n^{(\alpha)} := \frac{\Gamma(2\alpha+n)\Gamma(\alpha+\frac{1}{2})}{\Gamma(2\alpha)\Gamma(\alpha+n+\frac{1}{2})} P_n^{(\alpha-1/2\,,\,\alpha-1/2)}\,, \quad n = 0, 1, \ldots\,.$$

In terms of $C_n^{(\alpha)}$, the Chebyshev polynomials of the first and second kind are given by

$$T_n = \frac{n}{2} C_n^{(0)} \quad \text{and} \quad U_n = C_n^{(1)}\,, \qquad n = 0, 1, \ldots\,.$$

E.6 Hermite Polynomials.

a] Rodrigues' Formula. Let

$$H_n(x) := \frac{(-1)^n}{\exp(-x^2)} \frac{d^n}{dx^n} \exp(-x^2)\,.$$

Then $(H_n)_{n=0}^{\infty}$ is a sequence of orthogonal polynomials on $(-\infty, \infty)$ associated with the weight function

$$w(x) := \exp(-x^2)\,.$$

That is,

$$H_n \in \mathcal{P}_n \quad \text{and} \quad \int_{\mathbb{R}} H_n(x) H_m(x) \exp(-x^2)\, dx = 0$$

for any two nonnegative integers $n \neq m$.

In the rest of the exercise, the polynomials H_n are as in part a].

b] Normalization. We have

$$\int_{-\infty}^{\infty} (H_n(x))^2 \exp(-x^2)\, dx = \sqrt{\pi}\, 2^n n!$$

and

$$H_{2n+1}(0) = 0\,, \qquad H_{2n}(0) = (-1)^n \frac{(2n)!}{n!}\,.$$

c] Explicit Form.

$$H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} \frac{(-1)^m (2x)^{n-2m}}{m!(n-2m)!}\,.$$

d] Differential Equation. The function $y = H_n(x)$ satisfies

$$\frac{d^2y}{dx^2} - 2x\,\frac{dy}{dx} + 2ny = 0\,.$$

e] Recurrence Relation. The sequence $(H_n(x))_{n=0}^{\infty}$ satisfies

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$$

with

$$H_0(x) = 1 \qquad \text{and} \qquad H_1(x) = 2x\,.$$

f] Generating Function.

$$\sum_{n=0}^{\infty} H_n(x)\frac{z^n}{n!} = \exp(2xz - z^2)\,.$$

E.7 Laguerre Polynomials. Let $\alpha \in (-1, \infty)$.

a] Rodrigues' Formula. Let

$$L_n^{(\alpha)}(x) := \frac{1}{n!e^{-x}x^{\alpha}}\frac{d^n}{dx^n}(e^{-x}x^{\alpha+n})\,.$$

Then $(L_n^{(\alpha)})_{n=0}^{\infty}$ is a sequence of orthogonal polynomials on $[0, \infty)$ associated with the weight function

$$w(x) := x^{\alpha}\exp(-x)\,.$$

That is,

$$L_n^{(\alpha)} \in \mathcal{P}_n \quad \text{and} \quad \int_0^{\infty} L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)x^{\alpha}\exp(-x)\,dx = 0$$

for any two nonnegative integers $n \neq m$.

In the rest of the exercise, the polynomials $L_n^{(\alpha)}$ are defined as in part a].

b] Normalization. We have

$$\int_0^{\infty} (L_n^{(\alpha)}(x))^2 x^{\alpha}e^{-x}\,dx = \frac{\Gamma(\alpha + n + 1)}{n!}$$

and

$$L_n^{(\alpha)}(0) = \binom{n+\alpha}{n}\,.$$

c] Explicit Form.

$$L_n^{(\alpha)}(x) = \sum_{m=0}^{n} \frac{(-1)^m}{m!}\binom{n+\alpha}{n-m}x^m\,.$$

d] **Differential Equation.** The function $y = L_n^{(\alpha)}(x)$ satisfies

$$x\,\frac{d^2y}{dx^2} + (\alpha + 1 - x)\,\frac{dy}{dx} + ny = 0\,.$$

e] **Recurrence Relation.** The sequence $(L_n^{(\alpha)}(x))_{n=0}^{\infty}$ satisfies

$$(n+1)L_{n+1}^{(\alpha)}(x) = [(2n+\alpha+1) - x]L_n^{(\alpha)}(x) - (n+\alpha)L_{n-1}^{(\alpha)}(x)$$

with

$$L_0^{(\alpha)} = 1 \qquad \text{and} \qquad L_1^{(\alpha)}(x) = -x + \alpha + 1\,.$$

f] **Generating Function.**

$$\sum_{n=0}^{\infty} L_n^{(\alpha)}(x)z^n = \exp\left(\frac{xz}{z-1}\right)(1-z)^{-\alpha-1}\,.$$

E.8 Christoffel Numbers and Gauss-Jacobi Quadrature. Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. Let $x_{\nu,n}$, $\nu = 1, 2, \ldots, n$, denote the zeros of p_n. Let

$$\lambda_{\nu,n} := \frac{1}{p_n'(x_{\nu,n})}\int_{\mathbb{R}} \frac{p_n(x)}{x - x_{\nu,n}}\,d\alpha(x)\,, \qquad \nu = 1, 2, \ldots, n\,.$$

The numbers $\lambda_{\nu,n}$ are called the *Christoffel* or *Cotes numbers.*

a] Show that, for any $q \in \mathcal{P}_{2n-1}$,

$$\int_{\mathbb{R}} q(x)d\alpha(x) = \sum_{\nu=1}^{n} \lambda_{\nu,n}q(x_{\nu,n})\,.$$

Hint: First show the equality for every $q \in \mathcal{P}_{n-1}$ by using the Lagrange interpolation formula (E.6 of Section 1.1). If $q \in \mathcal{P}_{2n-1}$, then $q = sp_n + r$ with some $s, r \in \mathcal{P}_{n-1}$, where s is orthogonal to p_n. □

b] Show that $\lambda_{\nu,n} > 0$ for every $\nu = 1, 2, \ldots, n$.

Hint: Use part a] to show that

$$\lambda_{\nu,n} = \frac{1}{(p_n'(x_{\nu,n}))^2}\int_{\mathbb{R}} \left(\frac{p_n(x)}{x - x_{\nu,n}}\right)^2 d\alpha(x)\,.$$

□

c] Suppose $[a,b]$ is a finite interval containing $\operatorname{supp}(\alpha)$. Let $f \in C[a,b]$. Show that

$$\left| \int_a^b f(x)\, d\alpha(x) - \sum_{\nu=1}^{n} \lambda_{\nu,n} f(x_{\nu,n}) \right| \leq 2(\alpha(b+) - \alpha(a-)) \min_{p \in \mathcal{P}_{2n-1}} \|f - p\|_{[a,b]} .$$

Hint: Use parts a] and b] together with the observation

$$\sum_{\nu=1}^{n} \lambda_{\nu,n} = \int_a^b d\alpha(x) = \alpha(b+) - \alpha(a-) .$$

□

d] Suppose $\operatorname{supp}(\alpha) \subset [a,b]$, where $a, b \in \mathbb{R}$. Show that

$$\sum_{\nu=1}^{n} \lambda_{\nu,n} f(x_{\nu,n}) \underset{n\to\infty}{\longrightarrow} \int_a^b f(x) d\alpha(x)$$

for every Riemann-Stieltjes integrable function on $[a,b]$ with respect to α.

Hint: First show that f is Riemann-Stieltjes integrable on $[a,b]$ with respect to α if and only if for every $\epsilon > 0$ there are $g_1, g_2 \in C[a,b]$ so that

$$g_1(x) \leq f(x) \leq g_2(x) , \qquad x \in [a,b]$$

and

$$\int_a^b (g_2(x) - g_1(x))\, d\alpha(x) < \epsilon .$$

Finish the proof by part c] and the Weierstrass approximation theorem (see E.1 of Section 4.1). □

e] Suppose $\operatorname{supp}(\alpha)$ is compact. Let

$$Z := \{x_{\nu,n} : \nu = 1, 2, \ldots, n, \;\; n = 1, 2, \ldots\} .$$

Show that $\operatorname{supp}(\alpha) \subset \overline{Z}$, where $\overline{Z}$ denotes the closure of Z.

Hint: Use part d]. □

f] Show by an example that $\operatorname{supp}(\alpha) \neq \overline{Z}$ is possible.

E.9 Characterization of Compact Support. Using the notation of Theorem 2.3.2 and E.8, show that the following statements are equivalent:

(1) $\operatorname{supp}(\alpha)$ is compact.

(2) $\sup_{n\in\mathbb{N}} \{|a_n| + |b_n|\} < \infty$.

(3) The set $Z := \{x_{\nu,n} : \nu = 1, 2, \ldots, n, \;\; n = 1, 2, \ldots\}$ is bounded.

Outline. (1) $\Rightarrow$ (2). Note that the orthogonality of $\{p_n\}_{n=0}^{\infty}$ implies

$$a_n = \int_{\mathbb{R}} x p_{n-1}(x) p_n(x)\, d\alpha(x) \qquad \text{and} \qquad b_n = \int_{\mathbb{R}} x p_n^2(x)\, d\alpha(x)\,.$$

So $\operatorname{supp}(\alpha) \subset [-K, K]$, the Cauchy-Schwarz inequality, and the orthonormality of $(p_n)_{n=0}^{\infty}$ yield

$$\begin{aligned} |a_n| &\le K \int_{-K}^{K} |p_{n-1}(x) p_n(x)|\, d\alpha(x) \\ &\le K \left(\int_{-K}^{K} p_{n-1}^2(x)\, d\alpha(x) \right)^{1/2} \left(\int_{-K}^{K} p_n^2(x)\, d\alpha(x) \right)^{1/2} \le K \end{aligned}$$

and

$$|b_n| \le K \int_{-K}^{K} p_n^2(x) d\alpha(x) = K\,.$$

(2) $\Rightarrow$ (3). Use Theorem 2.3.2 (three-term recursion) to show that

$$x_{\nu,n} \sum_{k=0}^{n-1} p_k^2(x_{\nu,n}) \le 2 \sum_{k=0}^{n-1} a_{k+1} p_k(x_{\nu,n}) p_{k+1}(x_{\nu,n}) + \sum_{k=0}^{n-1} b_k p_k^2(x_{\nu,n})\,.$$

Hence

$$|x_{\nu,n}| \sum_{k=0}^{n-1} p_k^2(x_{\nu,n}) \le \left(2 \max_{0 \le k \le n-1} |a_k| + \max_{0 \le k \le n-1} |b_k| \right) \sum_{k=0}^{n-1} p_k^2(x_{\nu,n})\,.$$

(3) $\Rightarrow$ (1). If $Z \subset [-K, K]$, then by E.8 a]

$$\int_{\mathbb{R}} x^{2n-2}\, d\alpha(x) = \sum_{\nu=1}^{n} \lambda_{\nu,n} x_{\nu,n}^{2n-2} \le K^{2n-2} \int_{\mathbb{R}} d\alpha(x)\,,$$

which implies $\operatorname{supp}(\alpha) \subset [-K, K]$. □

E.10 A Condition for $\operatorname{supp}(\alpha) \subset [0, \infty)$**.** Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. Suppose $\operatorname{supp}(\alpha)$ is compact and

$$p_n(0) p_{n+1}(0) < 0\,, \qquad n = 0, 1, \dots\,.$$

Show that $\operatorname{supp}(\alpha) \subset [0, \infty)$.

Hint: Use the interlacing property of the zeros of p_n (E.2) to show that

$$Z := \{x_{\nu,n} : \nu = 1, 2, \dots, n,\ \ n = 1, 2, \dots\} \subset [0, \infty).$$

Now use E.8 e] to obtain $\operatorname{supp}(\alpha) \subset [0, \infty)$. □

E.11 The Solvability of the Moment Problem. Let $(\mu_n)_{n=0}^{\infty}$ be a sequence of real numbers. We would like to characterize those sequences $(\mu_n)_{n=0}^{\infty}$ for which there exists an m-distribution α so that

$$\int_{\mathbb{R}} x^n \, d\alpha(x) = \mu_n \,, \qquad n = 0, 1, \ldots \,.$$

Let

$$\mu(p) := \sum_{k=0}^{n} a_k \mu_k$$

for every $p \in \mathcal{P}_n$ of the form $p(x) = \sum_{k=0}^{n} a_k x^k$.

A polynomial p is called *nonnegative* if it takes nonnegative values on the real line. The sequence $(\mu_n)_{n=0}^{\infty}$ is called *positive definite* if

$$\mu(p) := \sum_{k=0}^{n} a_k \mu_k > 0 \,, \qquad n = 0, 1, \ldots$$

holds for every nonnegative polynomial $p \in \mathcal{P}_n$ of the form

$$p(x) = \sum_{k=0}^{n} a_k x^k \,.$$

The aim of this exercise is to outline the proof of Hamburger's characterization of the solvability of the moment problem by the positive definiteness of the sequence of moments. See part o].

a] Show that if there exists an m-distribution α so that

$$\int_{\mathbb{R}} x^n \, d\alpha(x) = \mu_n \,, \qquad n = 0, 1, \ldots \,,$$

then $(\mu_n)_{n=0}^{\infty}$ is positive definite.

Hint: An m-distribution α is increasing at infinitely many points. □

b] Show that $(\mu_n)_{n=0}^{\infty}$ is positive definite if and only if $\mu(p^2) > 0$ holds for every $0 \neq p \in \mathcal{P}_n$, $n = 0, 1, \ldots$.

Hint: Use E.3 of Section 2.4. □

c] Show that $(\mu_n)_{n=0}^{\infty}$ is positive definite if and only if

$$\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} > 0 \,, \qquad n = 0, 1, \ldots \,.$$

Hint: Use part b] and the law of inertia of Sylvester. See, for example, van der Waerden [50]. □

d] Helly's Selection Theorem. *Suppose the functions* f_n, $n = 1, 2, \dots$, *are nondecreasing on* $\mathbb{R}$, *and*

$$\sup_{n\in\mathbb{N}} \|f_n\|_{\mathbb{R}} < \infty .$$

Then there exists a subsequence of $(f_n)_{n=1}^{\infty}$ *that converges for every* $x \in \mathbb{R}$. *That is, we can select a pointwise convergent subsequence.*

Hint: See, for example, Freud [71]. □

e] Helly's Convergence Theorem. *Let* $[a, b]$ *be a finite interval. Suppose the functions* α_n, $n = 1, 2, \dots$, *are nondecreasing on* $[a, b]$ *and*

$$\sup_{n\in\mathbb{N}} \|\alpha_n\|_{[a,b]} < \infty .$$

Suppose also that $(\alpha_n(x))_{n=1}^{\infty}$ *converges to* $\alpha(x)$ *for every* $x \in [a, b]$. *Then*

$$\lim_{n\to\infty} \int_a^b f(x)\, d\alpha_n(x) = \int_a^b f(x)\, d\alpha(x)$$

for every $f \in C[a, b]$.

Hint: See, for example, Freud [71]. □

In the rest of the exercise (except for the last part) we assume that $(\mu_n)_{n=0}^{\infty}$ is positive definite. Our goal is to prove the converse of part a]. Let

$$p_n^*(x) := \begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} & 1 \\ \mu_1 & \mu_2 & \dots & \mu_n & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} & x^n \end{vmatrix} .$$

f] Show that

$$\mu(p_n^* q) = 0 , \qquad q \in \mathcal{P}_{n-1} .$$

g] Show that each p_n^* has n simple real zeros.

Hint: Use part f]. □

Let $x_{1,n} > x_{2,n} > \cdots > x_{n,n}$ be the zeros of p_n^*. Let

$$l_{\nu,n}(x) := \frac{p_n^*(x)}{p_n^{*\prime}(x_{\nu,n})(x - x_{\nu,n})} , \qquad \nu = 1, 2, \dots, n , \quad n = 1, 2, \dots$$

(see E.6 of Section 1.1), and let

$$\lambda_{\nu,n} := \mu(l_{\nu,n}) .$$

h] Show that

$$\mu(q) = \sum_{\nu=1}^{n} \lambda_{\nu,n} q(x_{\nu,n})$$

for every $q \in \mathcal{P}_{2n-1}$.

Hint: Use part f]. □

i] Show that

$$\lambda_{\nu,n} = \mu(l_{\nu,n}^2) > 0\,, \qquad \nu = 1, 2, \ldots, n\,, \quad n = 1, 2, \ldots\,.$$

Hint: Use part h]. □

For $x \in \mathbb{R}$, let

$$\alpha_n(x) := \sum_{\{\nu:\, x_{\nu,n} \le x\}} \lambda_{\nu,n}\,, \qquad n = 1, 2, \ldots\,.$$

j] Show that $0 \le \alpha_n(x) \le \mu_0$ on $\mathbb{R}$ for each n, and there is a subsequence of $(\alpha_n)_{n=1}^{\infty}$ that converges pointwise to a nondecreasing real-valued function α on $\mathbb{R}$.

Hint: Use parts h], i], and d]. □

k] Show that for every finite interval $[a, b]$,

$$\lim_{k\to\infty} \int_a^b x^m \, d\alpha_{n_k}(x) = \int_a^b x^m \, d\alpha(x)\,, \qquad m = 0, 1, 2, \ldots\,,$$

where α is defined in part j].

Hint: Use part e]. □

l] Let m be a fixed nonnegative integer and let $r := \lfloor m/2 \rfloor + 1$. Show that if $n_k \ge r + 1$, $a \le -1$ and $b \ge 1$, then

$$\left| \int_{-\infty}^{a} x^m \, d\alpha_{n_k}(x) + \int_b^{\infty} x^m \, d\alpha_{n_k}(x) \right| \le \left(\frac{1}{|a|} + \frac{1}{|b|} \right) \int_{\mathbb{R}} x^{2r} \, d\alpha_{n_k}(x) = \left(\frac{1}{|a|} + \frac{1}{|b|} \right) \mu_{2r}\,.$$

Hint: Use part h]. □

m] Show that

$$\int_{\mathbb{R}} x^m \, d\alpha(x) = \mu_m\,, \qquad m = 0, 1, 2, \ldots$$

where α is defined in part j].

Hint: Use parts k] and l]. □

n] Show that α defined in part j] is an m-distribution.

o] There exists an m-distribution α so that

$$\mu_n = \int_{\mathbb{R}} x^n \, d\alpha(x)$$

if and only if $(\mu_n)_{n=0}^{\infty}$ is positive definite, that is, if and only if

$$\begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n} \end{vmatrix} > 0 , \qquad n = 0, 1, \dots .$$

Hint: Combine parts a], c], m], and n]. □

Necessary and sufficient conditions for the uniqueness of the solution of the moment problem are given in Freud [71], for example.

E.12 Favard's Theorem. *Given $(a_n)_{n=0}^{\infty} \subset (0, \infty)$ and $(b_n)_{n=0}^{\infty} \subset \mathbb{R}$, the polynomials $p_n \in \mathcal{P}_n$ are defined by*

$$\begin{aligned} & x p_n(x) = a_n p_{n-1}(x) + b_n p_n(x) + a_{n+1} p_{n+1}(x) , \\ & p_{-1} = 0 , \qquad p_0 = \gamma_0 > 0 . \end{aligned}$$

Then there exists an m-distribution α such that

$$\int_{\mathbb{R}} p_n(x) p_m(x) \, d\alpha(x) = 0$$

for any two distinct nonnegative integers n and m. In other words, the converse of Theorem 2.3.2 is true.

In order to prove Favard's theorem, proceed as follows:

a] Show that the polynomials p_n are of the form

$$p_n(x) = \gamma_n x^n + r(x) , \qquad \gamma_n > 0 , \quad r \in \mathcal{P}_{n-1} .$$

b] Let $\widetilde{p}_n := \gamma_n^{-1} p_n$, $n = 0, 1, \dots$. The sequence $(\mu_n)_{n=0}^{\infty}$ is defined as follows. Let

$$\mu_0 := 1 , \quad \mu(q) := c \quad \text{if} \quad q = c , \quad c \in \mathbb{R} .$$

If $\mu_0, \mu_1, \dots, \mu_n$ have already been defined, then let

$$\mu(q) := \sum_{k=0}^{n} c_k \mu_k \quad \text{whenever} \quad q(x) = \sum_{k=0}^{n} c_k x^k , \quad c_k \in \mathbb{R}$$

and let

$$\mu_{n+1} := \mu(x^{n+1} - \widetilde{p}_{n+1}(x)) .$$

Show that

$$\mu(\widetilde{p}_n \widetilde{p}_m) = 0 , \qquad m = 0, 1, \ldots , n-1 , \quad n = 0, 1, \ldots$$

and

$$\mu(\widetilde{p}_n^2) > 0 , \qquad n = 0, 1, \ldots .$$

Hint: It is sufficient to prove that

$$\mu(\widetilde{p}_n(x) x^m) = 0 , \qquad m = 0, 1, \ldots , n-1 , \quad n = 0, 1, \ldots$$

and

$$\mu(\widetilde{p}_n(x) x^n) > 0 , \qquad n = 0, 1, \ldots .$$

These can be obtained from the definition of μ and from Theorem 2.3.2 (three-term recursion) by induction on n. □

c] Show that every $q \in \mathcal{P}_n$ is of the form

$$q = \sum_{k=0}^{n} c_k \widetilde{p}_k , \qquad c_k \in \mathbb{R}$$

and if $q \neq 0$, then

$$\mu(q^2) = \sum_{k=0}^{n} c_k^2 \mu(\widetilde{p}_k^2) > 0 .$$

d] Show that $(\mu_n)_{n=0}^{\infty}$ is positive definite in the sense of E.11.

Hint: Use the previous part and E.11 b]. □

e] Prove Favard's theorem.

Hint: Use part o] of E.11, parts d] and b] of this exercise, and the definition of μ. □

E.13 Christoffel Function. Let α be an m-distribution. For a fixed $n \in \mathbb{N}$, the function

$$\lambda_n(z) = \inf \left\{ \int_{\mathbb{R}} q^2(x) \, d\alpha(x) : q \in \mathcal{P}_{n-1} , \ |q(z)| = 1 \right\} , \qquad z \in \mathbb{C}$$

is called the nth *Christoffel function* associated with α.

a] Show that

$$\lambda_n(z) := \left(\sum_{k=0}^{n-1} |p_k(z)|^2 \right)^{-1} ,$$

where $(p_n)_{n=0}^{\infty}$ is the sequence of orthonormal polynomials associated with α.

Show also that the infimum in the definition of $\lambda_n(z)$ is actually a minimum, and it is attained if and only if

$$q(x) = \frac{\sum_{k=0}^{n-1} \overline{p_k(z)} p_k(x)}{\sum_{k=0}^{n-1} |p_k(z)|^2} .$$

Hint: Write

$$q = \sum_{k=0}^{n-1} c_k p_k , \qquad c_k \in \mathbb{C}$$

and observe that the orthonormality of $(p_n)_{n=0}^{\infty}$ implies

$$\int_{\mathbb{R}} q^2(x)\, d\alpha(x) = \sum_{k=0}^{n-1} |c_k|^2 .$$

Now use the Cauchy-Schwarz inequality (E.3 of Section 2.2) to find the maximum of $|q(z)|$ for polynomials $q \in \mathcal{P}_{n-1}^c$ satisfying

$$\int_{\mathbb{R}} q^2(x)\, d\alpha(x) \le 1$$

where $z \in \mathbb{C}$ is fixed.

b] Let $\lambda_{\nu,n}$, $\nu = 1, 2, \dots, n$, be the Christoffel numbers associated with an m-distribution α, that is, the coefficients in the Gauss-Jacobi quadrature formula, as in E.8. Show that

$$\lambda_{\nu,n} = \lambda_n(x_{\nu,n}) , \qquad \nu = 1, 2, \dots, n ,$$

that is, the Christoffel numbers are the values of the Christoffel function at the zeros of the nth orthonormal polynomial p_n.

Hint: Use parts E.8 a] and E.8 b]. □

c] Let $x \in \mathbb{R}$ be fixed. Show that

$$\sum_{n=0}^{\infty} p_n^2(x) < \infty$$

if and only if x is a mass point of α, that is, $\alpha(x-) < \alpha(x+)$, in which case

$$\sum_{n=0}^{\infty} p_n^2(x) = (\alpha(x+) - \alpha(x-))^{-1} .$$

Hint: Use part a] and the Weierstrass approximation theorem. See E.1 of Section 4.1. □

E.14 The Markov-Stieltjes Inequality. Let α be an m-distribution with associated orthonormal polynomials $(p_n)_{n=0}^{\infty}$. Let $x_{1,n} > x_{2,n} > \cdots > x_{n,n}$ denote the zeros of p_n. Let $x_{0,n} := \infty$ and $x_{n+1,n} := -\infty$. As in E.8 let $\lambda_{\nu,n}$, $\nu = 1, 2, \dots, n$, be the Christoffel numbers associated with α . Show that

$$\lambda_{\nu,n} \leq \int_{x_{\nu+1,n}}^{x_{\nu-1,n}} d\alpha(x)\,, \qquad \nu = 1, 2, \dots, n$$

and

$$\int_{x_{\nu,n}}^{x_{\nu-1,n}} d\alpha(x) \leq \lambda_{\nu,n} + \lambda_{\nu-1,n}\,, \qquad \nu = 2, 3, \dots, n\,.$$

Hint: Let $1 \leq k \leq n$ be fixed. Use E.7 of Section 1.1 (Hermite interpolation) to find polynomials $P \in \mathcal{P}_{2n-1}$ and $Q \in \mathcal{P}_{2n-1}$ with the following properties:

(1) $P(x_{j,n}) = Q(x_{j,n}) = 1\,, \qquad j = 1, 2, \dots, k-1\,,$

(2) $P(x_{k,n}) = 0\,, \qquad Q(x_{k,n}) = 1\,,$

(3) $P(x_{j,n}) = Q(x_{j,n}) = 0\,, \qquad j = k+1, k+2, \dots, n\,,$

(4) $P(x) \leq \chi_{(-\infty,\, x_{k,n}]}(x) \leq Q(x)\,, \qquad x \in \mathbb{R}\,,$

where

$$\chi_{(-\infty,\, x_{k,n}]}(x) := \begin{cases} 1 & \text{if} \quad -\infty < x \leq x_{k,n} \\ 0 & \text{if} \quad x_{k,n} < x < \infty\,. \end{cases}$$

Now apply E.8 (Gauss-Jacobi quadrature formula) to P and Q. □

E.15 Orthonormal Polynomials as Determinants. Suppose α is an m-distribution with moments

$$\mu_n = \int_{\mathbb{R}} x^n\, d\alpha(x)\,, \qquad n = 0, 1, \dots\,.$$

Let

$$\Delta_n := \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}\,, \qquad n = 0, 1, \dots\,.$$

a] Show that $\Delta_n > 0$, $n = 0, 1, 2, \dots$.

b] Show that the orthonormal polynomials p_n associated with α are of the form

$$p_n(x) = (\Delta_n \Delta_{n-1})^{-1/2} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_n & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} & x^n \end{vmatrix}\,.$$

c] Let $(a_n)_{n=0}^{\infty} \subset (0,\infty)$ and $(b_n)_{n=0}^{\infty} \subset \mathbb{R}$ be the coefficients in the three-term recursion for the sequence $(p_n)_{n=0}^{\infty}$ of orthonormal polynomials associated with α as in Theorem 2.3.2. Show that the monic orthogonal polynomials $\widetilde{p}_n := \gamma_n^{-1} p_n$ are of the form

$$\widetilde{p}_n(x) = \det\,(xI_n - J_n)$$

where J_n is the tridiagonal n by n *Jacobi matrix*

$$J_n := \begin{pmatrix} b_0 & a_1 & & & \\ a_1 & b_1 & a_2 & & \\ & a_2 & b_2 & a_3 & \\ & & \ddots & \ddots & \ddots \\ & & & a_n & b_n \end{pmatrix}$$

and I_n is the n by n unit matrix.

E.16 The Support of α. Let $(a_n)_{n=0}^{\infty} \subset (0,\infty)$ and $(b_n)_{n=0}^{\infty} \subset \mathbb{R}$ be the coefficient sequences in the three-term recursion for the sequence of $(p_n)_{n=0}^{\infty}$ of orthonormal polynomials associated with an m-distribution α as in Theorem 2.3.2.

a] Show that if $\mathrm{supp}(\alpha) \subset [b-a, b+a]$ with some $a > 0$ and $b \in \mathbb{R}$, then

$$a_n \le a \quad \text{and} \quad |b_n - b| \le a\,, \qquad n = 0, 1, \ldots\,.$$

Hint: Use the orthonormality of $(p_n)_{n=0}^{\infty}$ to show that

$$a_n = \int_{b-a}^{b+a} (x-b) p_{n-1}(x) p_n(x)\, d\alpha(x)$$

and

$$b_n - b = \int_{b-a}^{b+a} (x-b) p_n^2(x)\, d\alpha(x)\,.$$

Now apply the Cauchy-Schwarz inequality, and use the orthonormality of $(p_n)_{n=0}^{\infty}$ again. □

b] Show that

$$\mathrm{supp}(\alpha) \subset [-K, K]$$

where

$$K := 2\sup\{a_n : n \in \mathbb{N}\} + \sup\{|b_n| : n \in \mathbb{N}\}$$

(the suprema are taken over all nonnegative integers).

Hint: Suppose $K < \infty$; otherwise there is nothing to prove. Combine E.9, the inequality in the hint to the direction (2) ⇒ (3) of E.9, and E.8 e]. □

c] Blumenthal's Theorem. *Assume that*

$$\lim_{n\to\infty} a_n =: \frac{a}{2} > 0 \qquad \textit{and} \qquad \lim_{n\to\infty} b_n =: b \in \mathbb{R}\,.$$

Then

$$\operatorname{supp}(\alpha) = [b-a, b+a] \cup F$$

where $F \subset \mathbb{R} \setminus [b-a, b+a]$ *is a countable bounded set for which*

$$F \setminus (b-a-\epsilon, b+a+\epsilon)$$

is finite for every $\epsilon > 0$.

Proof. See Nevai [79b] or Máté, Nevai, and Van Assche [91]. □

d] Rakhmanov's Theorem. *Suppose* $supp(\alpha) \subset [b-a, b+a]$ *with some* $a > 0$ *and* $b \in \mathbb{R}$. *Suppose also that* $\alpha'(x) > 0$ *a.e. in* $[b-a, b+a]$. *Then*

$$\lim_{n\to\infty} a_n = \frac{a}{2} \quad \textit{and} \quad \lim_{n\to\infty} b_n = b\,.$$

Proof. See, for example, Máté, Nevai, and Totik [85], or Nevai [91]. □

There is an analogous theory of orthogonal polynomials on the unit circle initiated by Geronimus, Shohat, and Achiezer. An important contribution, called Szegő theory, may be found in Freud [71].

E.17 A Theorem of Stieltjes [14]. Let w be a positive continuous weight function on $[-a, a]$. Denote the nth moment by

$$\mu_n := \int_{-a}^{a} x^n w(x)\, dx\,.$$

Let $\widetilde{p}_n \in \mathcal{P}_n$ denote the nth *monic* orthogonal polynomial on $[-a, a]$ associated with the weight w. Then $(\widetilde{p}_n)_{n=0}^{\infty}$ satisfies a three-term recursion

$$\widetilde{p}_n(x) = (x - A_n)\widetilde{p}_{n-1}(x) - B_n \widetilde{p}_{n-2}(x)$$

with $\widetilde{p}_0(x) = 1$ and $\widetilde{p}_1(x) := x - A_1$; see Theorem 2.3.2.

Suppose the sequence of polynomials $(q_n)_{n=0}^{\infty}$ satisfies the same recursion commencing with $q_0(x) := 0$ and $q_1(x) := B_1$. Stieltjes' theorem (see, for example, Cheney [66]) states the following.

Theorem. *For any* $x \notin [-a, a]$,

$$\begin{aligned}\int_{-a}^{a} \frac{w(t)\, dt}{x-t} &= \frac{\mu_0}{x} + \frac{\mu_1}{x^2} + \cdots \\ &= \cfrac{B_1}{x - A_1 - \cfrac{B_2}{x - A_2 - \cfrac{B_3}{x - A_3 - \cdots}}}\,.\end{aligned}$$

Furthermore, the nth convergent $q_n/\widetilde{p}_n$ satisfies

$$\frac{q_n(x)}{\widetilde{p}_n(x)} = \cfrac{B_1}{x - A_1 - \cfrac{B_2}{\ddots \cfrac{B_n}{x - A_n}}}.$$

E.18 Completeness of Orthogonal Polynomials. Let $(p_n)_{n=0}^{\infty}$ be the sequence of orthonormal polynomials associated with an m-distribution α. If $\operatorname{supp}(\alpha) \subset [a,b]$, where $[a,b]$ is a finite interval, then $(p_n)_{n=0}^{\infty}$ is a maximal orthogonal collection in $L_2[a,b]$.

Hint: Use the Weierstrass approximation theorem (E.1 of Section 4.1) on $[a,b]$. □

E.19 Bounds for Jacobi Polynomials. For all Jacobi weight functions $w(x) = (1-x)^{\alpha}(1+x)^{\beta}$ with $\alpha \geq -1/2$ and $\beta \geq -1/2$, the inequalities

$$\max_{x\in[-1,1]} \frac{p_n^2(x)}{\sum_{k=0}^{n} p_k^2(x)} \leq \frac{4\left(2+\sqrt{\alpha^2+\beta^2}\right)}{2n+\alpha+\beta+2}$$

and

$$\max_{x\in[-1,1]} \sqrt{1-x^2}\, w(x) p_n^2(x) \leq \frac{2e\left(2+\sqrt{\alpha^2+\beta^2}\right)}{\pi}$$

hold, where $(p_n)_{n=0}^{\infty}$ is the sequence of orthonormal Jacobi polynomials associated with the weight function w.

Proof. See Erdélyi, Magnus, and Nevai [94]. □

2.4 Polynomials with Nonnegative Coefficients

A quadratic polynomial $x^2 + \alpha x + \beta$ with real coefficients has both roots in the halfplane $\{z \in \mathbb{C} : \operatorname{Re}(z) \leq 0\}$ if and only if $\beta \geq 0$ and $\alpha \geq 0$. This easy consequence of the quadratic formula gives the following lemma:

Lemma 2.4.1. *If $p \in \mathcal{P}_n$ has all its zeros in $\{z \in \mathbb{C} : \operatorname{Re}(z) \leq 0\}$, then either p or $-p$ has all nonnegative coefficients.*

The converse of this is far from true. Indeed, the following result of Meissner holds (see Pólya and Szegő [76]). We denote by $\mathcal{P}_n^+$ the set of polynomials in $\mathcal{P}_n$, that have all nonnegative coefficients.

Theorem 2.4.2. *If $p \in \mathcal{P}_n^c$ and $p(x) > 0$ for $x > 0$, then $p = s/t$, where s and t are both polynomials with all nonnegative coefficients.*

Since a polynomial p that is real-valued on the positive real axis has real coefficients, and since $p(x) > 0$ for all $x > 0$ implies that the leading coefficient of p is positive, Theorem 2.4.2 will follow immediately from the next lemma.

Lemma 2.4.3. *Suppose $\alpha, \beta \in \mathbb{R}$ and suppose $x^2 - \alpha x + \beta$ has no nonnegative root. Then $x^2 - \alpha x + \beta = p(x)/q(x)$, where $p, q \in \mathcal{P}_m$ both have all nonnegative coefficients, and where*

$$m \leq 10 \left(4 - \frac{\alpha^2}{\beta}\right)^{-1/2} .$$

Proof. The quadratic polynomial $x^2 - \alpha x + \beta$ has no positive root if and only if $\alpha^2 < 4\beta$. We set $c := \alpha^2/\beta$ and note that $c < 4$. Consider

$$\begin{aligned}(x^2 - \alpha x + \beta)(x^2 + \alpha x + \beta) &= x^4 + (2\beta - \alpha^2)x^2 + \beta^2 \\ &= x^4 + \beta(2 - c)x^2 + \beta^2 .\end{aligned}$$

If $c \leq 2$ we have the desired factorization. If $c > 2$, consider

$$\begin{aligned}(x^4 + \beta(2 - c)x^2 + \beta^2)(x^4 - \beta(2 - c)x^2 + \beta^2) \\ = x^8 + \beta^2(2 - (2 - c)^2)x^4 + \beta^4 .\end{aligned}$$

If $2 - (2 - c)^2 > 0$ we are finished. In general, we proceed as follows:
Let

$$\begin{aligned}P_n(x) :&= x^{2^{n+1}} + \beta^{2^{n-1}}(2 - (2 - (2 - \cdots 2 - (2 - c)^2)^2 \cdots)^2)x^{2^n} + \beta^{2^n} \\ &= x^{2^{n+1}} + \beta^{2^{n-1}} c_n x^{2^n} + \beta^{2^n} ,\end{aligned}$$

where c_n has n nested terms. Let

$$Q_n(x) := x^{2^{n+1}} - \beta^{2^{n-1}} c_n x^{2^n} + \beta^{2^n} .$$

Note that, since $c_{n+1} = 2 - c_n^2$

$$\begin{aligned}P_n(x)Q_n(x) &= x^{2^{(n+1)+1}} - \beta^{2^n} c_n^2 x^{2^{n+1}} + 2\beta^{2^n} x^{2^{n+1}} + \beta^{2^{n+1}} \\ &= x^{2^{n+2}} + \beta^{2^n} c_{n+1} x^{2^{n+1}} + \beta^{2^{n+1}} \\ &= P_{n+1}(x) .\end{aligned}$$

Consider the smallest n (if it exists) such that c_n is nonnegative. Then

$$(x^2 - \alpha x + \beta)(x^2 + \alpha x + \beta) = P_1(x)$$

and

$$P_1 Q_1 Q_2 \cdots Q_{n-1} = P_n \,,$$

where $Q_1 Q_2 \cdots Q_{n-1} \in \mathcal{P}^+_{2^{n+1}-4}$ since each $c_k < 0$ for $k < n$, and where $P_n \in P^+_{2^{n+1}}$ since $c_n \geq 0$. Thus, we have the desired representation

$$x^2 - \alpha x + \beta = \frac{P_n(x)}{(x^2 + \alpha x + \beta)(Q_1 Q_2 \cdots Q_{n-1})(x)} \,,$$

where n is the smallest integer such that $c_n > 0$.

Now suppose $c_1, \ldots, c_{n-1}$, c_n are all nonpositive. Then

$$c_k = -\sqrt{2 - c_{k+1}} \,, \qquad k = 1, 2, \ldots, n-1$$

and $c_1 = 2 - c$ imply

$$c > 2 + (2 + \cdots (2 + (2 + 2^{1/2})^{1/2})^{1/2} \cdots)^{1/2} =: \delta_n \,, \tag{2.4.1}$$

where the above formula contains n iterations. Since, by assumption, $c < 4$, and since $\delta_n \to 4$ as $n \to \infty$, it is clear that (2.4.1) is not satisfied for some n, and eventually some c_n is greater than zero.

The estimate on the degree requires analyzing the rate of convergence of $(\delta_n)_{n=0}^{\infty}$. Since $\delta_n = 2 + \sqrt{\delta_{n-1}}$, we have

$$4 - \delta_n = 2 - \sqrt{\delta_{n-1}} = \frac{4 - \delta_{n-1}}{2 + \sqrt{\delta_{n-1}}} \leq \frac{4 - \delta_{n-1}}{2} \,.$$

By repeated applications of the above,

$$4 - \delta_n \leq \frac{4 - \delta_0}{2^n} = \frac{1}{2^{n-1}} \,.$$

Now we can improve the above estimate as follows: We have

$$\begin{aligned}
4 - \delta_n =& \frac{4 - \delta_{n-1}}{2 + \sqrt{\delta_{n-1}}} = \frac{4 - \delta_{n-2}}{(2 + \sqrt{\delta_{n-1}})(2 + \sqrt{\delta_{n-2}})} \\
=& \frac{4 - \delta_0}{(2 + \sqrt{\delta_{n-1}})(2 + \sqrt{\delta_{n-2}}) \cdots (2 + \sqrt{\delta_0})} \\
\leq& \frac{2}{(2 + \sqrt{4 - 2^{2-n}})(2 + \sqrt{4 - 2^{3-n}}) \cdots (2 + \sqrt{4 - 2^{(n+1)-n}})} \\
\leq& \frac{2}{(2 + 2 - 2^{1-n})(2 + 2 - 2^{2-n}) \cdots (2 + 2 - 2^{n-n})} \\
\leq& 2 \cdot 4^{-n} \prod_{j=2}^{n+1} \frac{1}{1 - 2^{-j}} \leq 2 \cdot 4^{-n} \prod_{j=2}^{n+1} (1 + 2 \cdot 2^{-j}) \leq 2e 4^{-n} \,.
\end{aligned}$$

So if

$$m := 2^{n+1} \geq \frac{2\sqrt{2e}}{\sqrt{4-c}},$$

then

$$4 - \delta_n \leq 2e4^{-n} \leq 4 - c,$$

that is, $\delta_n \geq c$. □

We note that in the above proof a little additional effort yields a slightly better constant than 10.

Let $k \in \mathbb{N}$ and $\epsilon \in (0, \pi)$. It follows easily from Lemma 2.3.4 that if $p \in \mathcal{P}_k$ has no zeros in the cone

$$\{z \in \mathbb{C} : |\arg(z)| < \epsilon\},$$

then there are $s, t \in \mathcal{P}_m^+$ with $m \leq \frac{5}{4}\pi k\epsilon^{-1}$ so that $p = s/t$; see E.1 d]. The essential sharpness of this upper bound is shown by E.1 e]. An easier proof of Theorem 2.4.2 that gives a weaker bound for the degree of the numerator and denominator in the representation is given by E.1 f].

A similar sort of representation theorem due to Bernstein [15] is the following:

Theorem 2.4.4. *If $p \in \mathcal{P}_n$ and $p(x) > 0$ for $x \in (-1, 1)$, then there is a representation*

$$p(x) = \sum_{j=0}^{d} a_j (1-x)^j (1+x)^{d-j}$$

with each $a_j \geq 0$. (The smallest $d := d(p)$ for which such a representation exists is called the Lorentz degree of p.)

It suffices to prove this result for quadratic polynomials; this is left as an exercise; see E.1 f].

The proof of the following interesting result of Barnard et al. [91] is surprisingly complicated, and we do not reproduce it here.

Theorem 2.4.5. *Suppose that $p \in \mathcal{P}_n$ has all nonnegative coefficients. Suppose that the zeros of p are $z_1, z_1, \ldots, z_n \in \mathbb{C}$. For $\tau \geq 0$, let*

$$p_\tau(z) = \prod_{\substack{j=1 \\ |\arg(z_j)|>\tau}}^{n} \left(1 - \frac{z}{z_j}\right),$$

where $\arg(z)$ is defined so that $\arg(z) \in [-\pi, \pi)$. Then $p_\tau(z)$ has all nonnegative coefficients.

It follows from this result that if $p \in \mathcal{P}_n$ has all nonnegative coefficients and if $q(x) = x^2 + \alpha x + \beta$ is a quadratic polynomial with zeros forming a pair of conjugate zeros of p that have least angular distance from the positive x-axis, then p/q also has all nonnegative coefficients.

Comments, Exercises, and Examples.

Polynomials with all nonnegative coefficients have a number of distinguishing properties that are explored in the exercises. For example, only analytic functions with all nonnegative coefficients can be approximated uniformly on $[0,1]$ by such polynomials; see E.2. So a Weierstrass-type theorem does not hold for these polynomials. This is quite different from approximation by polynomials of the form

$$\sum a_{i,j}(x+1)^i(1-x)^j\,, \qquad a_{i,j} \geq 0\,. \tag{2.4.2}$$

Since every polynomial that is strictly positive on $(-1,1)$ has such a representation (E.1 b]), it follows from the Weierstrass approximation theorem that all nonnegative functions from $C[-1,1]$ are in the uniform closure.

It follows from Theorem 2.4.2 and the Weierstrass approximation theorem (see E.1 of Section 4.1) that fractions of polynomials with all nonnegative coefficients form a dense set in the uniform norm on $[0,1]$ in the set of nonnegative continuous functions on a finite closed interval $[0,1]$. Hence they have a much larger uniform closure on $[0,1]$ than that of the polynomials with all nonnegative coefficients.

Various inequalities for polynomials of the form (2.4.2) are considered in Appendix 5.

E.1 Remarks on Theorem 2.4.2.

a] Suppose $\alpha, \beta \in \mathbb{R}, \epsilon \in (0,\pi)$, and suppose $x^2+\alpha x+\beta$ has no zeros in the cone

$$\{z \in \mathbb{C}:\ |\arg(z)| < \epsilon\}\,.$$

Show that there are $p, q \in \mathcal{P}_m^+$ with $m \leq \frac{5}{2}\pi\epsilon^{-1}$ such that

$$x^2+\alpha x+\beta = \frac{p(x)}{q(x)}\,.$$

Hint: This is a reformulation of Theorem 2.4.3 by introducing the angle between the positive x-axis and the zero of the quadratic polynomial. □

b] Let $n \in \mathbb{N}$. Show that if $p \in \mathcal{P}_n^+$, then p has no zeros in the cone

$$\{z \in \mathbb{C}:\ |\arg(z)| < \pi/(2n)\}\,.$$

This is sharp, as the example $p(x) := x^n + 1$ shows.

c] Show that the result of part a] is sharp up to the constant $\frac{5}{2}$.

Hint: Let $n \in \mathbb{N}$. Consider

$$x^2+\alpha x+\beta = \left(x - \exp\left(\tfrac{2\pi i}{2n}\right)\right)\left(x - \exp\left(\tfrac{-2\pi i}{2n}\right)\right).$$

Show that if there are $p, q \in \mathcal{P}_m^+$ so that $x^2 + \alpha x + \beta = p(x)/q(x)$, then $m \geq n - 1$. □

d] Let $\epsilon \in (0, \pi)$ and $k \in \mathbb{N}$. Suppose $p \in \mathcal{P}_k$ has no zeros in the cone

$$\{z \in \mathbb{C} : \ |\arg(z)| < \epsilon\}\,.$$

Show that there are $s, t \in \mathcal{P}_m^+$ with $m \leq \frac{5}{4}\pi k \epsilon^{-1}$ so that $p = s/t$.

e] Let $\epsilon \in (0, \pi)$ and $k \in \mathbb{N}$. Let

$$f_k(x) := ((x - z_0)(x - \overline{z}_0))^k\,,$$

where $\arg(z_0) = \epsilon$. Assume that $f_k = s/t$, where $s, t \in \mathcal{P}_m^+$. Show that

$$m \geq (\log 2) k \epsilon^{-1}\,.$$

Hint: First observe that

$$s(y) \leq s(y + y\delta m^{-1}) \leq e^{\delta} s(y)$$

for every $s \in \mathcal{P}_m^+$, $y \in (0, \infty)$, and $\delta \in (0, \infty)$. Therefore

$$f_k(y + y\delta m^{-1}) \leq e^{\delta} f_k(y)$$

for every $y \in (0, \infty)$ and $\delta \in (0, \infty)$. Now let $y \geq |z_0|$ be chosen so that $|y - z_0| = \epsilon$. Applying the above inequality with this y and $\delta := m\alpha$, we obtain

$$f_k(y + \epsilon y) \leq e^{m\epsilon} f_k(y)\,,$$

hence $2^k \leq e^{m\epsilon}$, that is, $k \log 2 \leq m\epsilon$. □

f] Prove that if $r \in \mathcal{P}_n$ and $r(x) > 0$ for all $x > 0$, then there is an integer $d \geq n$ such that

$$r(x) = \frac{q(x)}{(1 + x)^{d-n}}\,,$$

where $q \in \mathcal{P}_d^+$.

Hint: Let $\alpha, \beta \in \mathbb{R}$ and $\alpha^2 < 4\beta$. Consider

$$(x^2 - \alpha x + \beta)(1 + x)^d = \sum_{j=0}^{d+2} c_j x^j$$

and compute c_j explicitly. □

g] If $p \in \mathcal{P}_n$ and $p(x) > 0$ for all $x \in (-1, 1)$, then it is of the form

$$p(x) = \sum_{j=0}^{d} a_j (1 - x)^j (1 + x)^{d-j}\,, \qquad a_j \geq 0$$

for some $d \geq n$.

Hint: Apply a] to

$$r(u) := (1 + u)^n p\left(\frac{1 - u}{1 + u}\right)\,, \qquad u := \frac{1 - x}{1 + x}\,.$$

□

E.2 Polynomials with Nonnegative Coefficients.

a] If $p \in \mathcal{P}_n^+$, then for $x > 0$

$$0 \le p'(x) < \frac{n}{x}\, p(x)\,.$$

b] If $(p_n)_{n=1}^{\infty}$ is a sequence of polynomials with $p_n \in \mathcal{P}^+ := \cup_{k=0}^{\infty} \mathcal{P}_k^+$ and $(p_n)_{n=1}^{\infty}$ converges to f uniformly on $[0,1]$, then f is the restriction to $[0,1]$ of a function analytic in $D := \{z \in \mathbb{C} : |z| < 1\}$ of the form

$$f(z) = \sum_{n=0}^{\infty} a_n z^n\,, \qquad a_n \ge 0\,.$$

Hint: Since $(p_n(1))_{n=1}^{\infty}$ converges and each p_n has nonnegative coefficients, there is a constant C such that

$$\|p_n\|_D \le p_n(1) \le C\,, \qquad n = 1, 2, \ldots\,.$$

Now Montel's theorem (see, for example, Ash [71]) implies that $(p_n)_{n=1}^{\infty}$ has a locally uniformly convergent subsequence on D. Deduce that this subsequence converges to an f with nonnegative coefficients. □

E.3 Nonnegative-Valued Polynomials and Sums of Squares.

a] Suppose $p \in \mathcal{P}_{2n}$ is nonnegative on $\mathbb{R}$. Then there exist s, $t \in \mathcal{P}_n$ such that

$$p(x) = s^2(x) + t^2(x)\,.$$

Hint: If $p \in \mathcal{P}_2$ and p is nonnegative, then for some real numbers α and β,

$$p(x) = (x - \alpha)^2 + \beta^2\,.$$

Now use the identity

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2\,.$$

□

b] If $p \in \mathcal{P}_{2n}$ is nonnegative for $x \ge 0$, then there exist s, t, u, $v \in \mathcal{P}_n$ such that

$$p(x) = s^2(x) + t^2(x) + xu^2(x) + xv^2(x)\,.$$

c] Suppose $t \in \mathcal{T}_n$ is nonnegative on $\mathbb{R}$. Show that there exists a $q \in \mathcal{P}_n^c$ such that

$$t(\theta) = |q(e^{i\theta})|^2\,, \qquad \theta \in \mathbb{R}\,.$$

Show also that if, in addition, $t \in \mathcal{T}_n$ is even, then there exists a $q \in \mathcal{P}_n$ such that the above holds.

d] If $p \in \mathcal{P}_n$ is nonnegative on $[-1,1]$, then there exist s, $t \in \mathcal{P}_n$ such that

$$p(x) = s^2(x) + (1-x^2)t^2(x)\,.$$

Hint: Write, by c],

$$p(\cos\theta) = |s(\cos\theta) + it(\cos\theta)\sin\theta|^2\,.$$

□

The above exercise follows Pólya and Szegő [76]; see also E.1 of Section 7.2 where this result is extended.

The following two exercises discuss results proved in Erdélyi and Szabados [88], [89b], and Erdélyi [91c].

E.4 Lorentz Degree of Polynomials. Given a polynomial $p \in \mathcal{P}_n$, let $d = d(p)$ be the minimal nonnegative integer for which the polynomial p is of the form

$$p(x) = \pm \sum_{j=0}^{d} a_j (1-x)^j (x+1)^{d-j}\,, \quad a_j \geq 0\,.$$

If there is no such d, then let $d(p) := \infty$. We call $d := d(p)$ the *Lorentz degree* of the polynomial p.

a] Let $p \in \mathcal{P}_n \setminus \mathcal{P}_{n-1}$ be of the form

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

For $m \geq n$, let the numbers $b_{j,m}$, $j = 0, 1, \ldots, m$, be defined by

$$\begin{aligned} p(x) &= \left(\sum_{j=0}^{n} b_j (1-x)^j (x+1)^{n-j} \right) \left(\frac{1-x}{2} + \frac{x+1}{2} \right)^{m-n} \\ &= \sum_{j=0}^{m} b_{j,m} (1-x)^j (x+1)^{m-j}\,. \end{aligned}$$

Show that if $d(p)$ is finite, then it is the smallest value of m for which each $b_{j,m}$ is nonnegative or each $b_{j,m}$ is nonpositive.

b] Show that if $p \in \mathcal{P}_1 \setminus \mathcal{P}_0$ has no zeros in $(-1,1)$, then $d(p) = 1$.

c] Show that if $p \in \mathcal{P}_2 \setminus \mathcal{P}_1$ has no zeros in the open unit disk, then $d(p) = 2$.

d] Show that if the zeros of a polynomial $p \in \mathcal{P}_2 \setminus \mathcal{P}_1$ lie on the ellipse

$$B_\epsilon := \left\{ z = x + iy : \;\; y^2 = \epsilon^2 (1-x^2), \;\; x \in (-1,1) \right\}$$

with $\epsilon \in (0,1]$, then

$$\epsilon^{-2} \leq d(p) < 2\epsilon^{-2} + 1.$$

e] Let $\epsilon \in (0,1)$ be such that ϵ^{-2} is an integer. Let

$$p_1(x) := x^2 + 2\frac{3\epsilon^2 - 1}{1-\epsilon^2}x + \frac{8\epsilon^4 - 5\epsilon^2 + 1}{1-\epsilon^2}\,.$$

Show that p_1 has its zeros on B_ϵ defined in d], and $d(p_1) = \epsilon^{-2}$.

f] Let $\epsilon \in (0,1)$ be such that $2\epsilon^{-2}$ is an integer. Let

$$p_2(x) := x^2 - 2\frac{2 - 3\epsilon^2}{(1-\epsilon^2)(2-\epsilon^2)}x + \frac{-\epsilon^8 - 5\epsilon^6 + \epsilon^4 - 8\epsilon^2 + 4}{(1-\epsilon^2)(2-\epsilon^2)^2}\,.$$

Show that p_2 has its zeros on B_ϵ defined in part d], and $d(p_2) = 2\epsilon^{-2}$.

g] Show that $d(pq) \le d(p) + d(q)$ for any two polynomials p and q.

h] Let $\epsilon \in (0,1]$. Show that if $p \in \mathcal{P}_n$ has no zeros in

$$D_\epsilon := \left\{z = x + iy : y^2 < \epsilon^2(1-x^2),\ \ x \in (-1,1),\ \ y \in \mathbb{R}\right\},$$

then

$$d(p) < 2n\epsilon^{-2} + n < 3n\epsilon^{-2}\,.$$

i] Let p be a polynomial. Show that $d(p) < \infty$ if and only if $p = 0$ or p has no zeros in $(-1,1)$.

j] Show that

$$\left|p\left(y + \tfrac{i}{2}\sqrt{\tfrac{(1-y^2)n}{d}}\right)\right| \le 2^n \left|p\left(y - \tfrac{n}{4d}\right)\right|$$

for every $p \in \mathcal{B}_d(-1,1)$, $1 \le n \le d$, and $y \in [0,1)$ (i is the imaginary unit).

Hint: Modify the proof of Lemma A.5.4. □

k] Let $b \in [0,1]$. Show that

$$p'(b) \le d\,p(b)$$

for every $p \in \mathcal{B}_d(-1,1)$, positive in $(-1,1)$.

Hint: If $q_{j,d}(x) := (1-x)^j(x+1)^{d-j}$, then

$$q'_{j,d}(b) = q_{j,d}(b)\left(\frac{d-j}{1+b} - \frac{j}{1-b}\right) \le d\,q_{j,d}(b)\,, \quad j = 0,1,\ldots,d$$

for every $b \in [0,1]$. □

l] Show that $d(p) \ge \frac{1}{17}n\epsilon^{-2}$ whenever

$$p(x) = ((x - z_0)(x - \bar{z}_0))^n\,, \quad z_0 \in B_\epsilon\,, \quad \epsilon \in (0,1]\,,$$

where B_ϵ is defined in part d].

Proof. Let $z_0 = y + i\epsilon(1-y^2)^{1/2}$, $y \in (-1,1)$. Without loss of generality it may be assumed that $0 \le y < 1$. Distinguish two cases.

Case 1: $1 - 2\epsilon^2 \le y < 1$. By part k],

$$\begin{aligned} d(p) \ge \frac{p'(1)}{p(1)} &= \frac{2n(1-y)}{(1-y)^2 + \epsilon^2(1-y^2)} \\ &= \frac{2n}{(1-y) + \epsilon^2(1+y)} > \frac{n}{2\epsilon^2}\,. \end{aligned}$$

Case 2: $0 \le y < 1 - 2\epsilon^2$. Applying part k] with

$$b := y + \epsilon(1-y^2)^{1/2} \in [0,1]$$

deduce that

$$d(p) \ge \frac{p'(b)}{p(b)} = \frac{2n(b-y)}{(b-y)^2 + \epsilon^2(1-y^2)} = \frac{n}{\epsilon(1-y^2)^{1/2}}\,. \tag{2.4.3}$$

Use part j] to obtain

$$\left|(1-y^2)\epsilon^2 - \frac{(1-y^2)n}{4d}\right|^n \le 2^n \left((1-y^2)\epsilon^2 + \frac{n^2}{16d^2}\right)^n\,, \tag{2.4.4}$$

where $d := d(p)$ and $n = \frac{1}{2}\deg(p)$. If

$$(1-y^2)\epsilon^2 \ge \frac{(1-y^2)n}{8d}\,,$$

then there is nothing to prove. Therefore assume that

$$(1-y^2)\epsilon^2 \le \frac{(1-y^2)n}{8d}\,. \tag{2.4.5}$$

Now (2.4.3) to (2.4.5) yield

$$\left(\frac{(1-y^2)n}{8d}\right)^n \le 2^n \left((1-y^2)\epsilon^2 + \frac{n^2}{16d^2}\right)^n$$

and so

$$\frac{(1-y^2)n}{8d} \le 2\left((1-y^2)\epsilon^2 + \frac{n^2}{16d^2}\right)\,.$$

Since, by (2.4.3), $n^2 d^{-2} \le (1-y^2)\epsilon^2$, the above inequality implies

$$\frac{(1-y^2)n}{8d} \le 2\,\frac{17}{16}(1-y^2)\epsilon^2$$

and so $d \ge \frac{1}{17} n\epsilon^{-2}$. □

m] Show that $p \in \mathcal{P}_n \setminus \mathcal{P}_{n-1}$ and $d(p) = n$ imply that the zeros $z_1, z_2, \ldots, z_n$ of p satisfy $|z_1 z_2 \cdots z_n| \geq 1$.

n] Show that $d(pq) < \max\{d(p), d(q)\}$ can happen.

Hint: Let

$$p(x) := (1-x)^2 - 2(1-x^2) + 4(x+1)^2 \quad \text{and} \quad q(x) := (x+1) + \tfrac{1}{2}(1-x)\,.$$

Show that $d(p) = 4$, $d(q) = 1$, and $d(pq) = 3$. □

o] Show that if $p \in \mathcal{P}_k \setminus \mathcal{P}_{k-1}$ has no zeros in $(-1, 1)$, $z \in \mathbb{C}$, $|z| > 1$, and

$$p(x) = ((x-z)(x-\overline{z}))^m q(x)\,,$$

then $d(p) = \deg(p) = k + 2m$ if m is sufficiently large. This shows that polynomials p with the property $d(p) = \deg(p)$ can have arbitrary many prescribed zeros in $\mathbb{C} \setminus (-1, 1)$.

E.5 Lorentz Degree of Trigonometric Polynomials. Given $\omega \in (0, \pi]$ and a real trigonometric polynomial $t \in \mathcal{T}_n$, let $d = d_\omega(t)$ be the minimal nonnegative integer for which t is of the form

$$t(\theta) = \pm \sum_{j=0}^{2d} a_j \sin^j \frac{\omega - \theta}{2} \sin^{2d-j} \frac{\theta + \omega}{2}\,, \qquad a_j \geq 0\,.$$

If there is no such d, then let $d_\omega(t) := \infty$. We call $d = d_\omega(t)$ the *Lorentz degree* of t.

a] Let $t \in \mathcal{T}_n \setminus \mathcal{T}_{n-1}$ be of the form

$$t(\theta) = \sum_{j=0}^{2n} b_j \sin^j \frac{\omega - \theta}{2} \sin^{2n-j} \frac{\theta + \omega}{2}\,.$$

For $m \geq n$ let the numbers $b_{j,m}$, $j = 0, 1, \ldots, 2m$, be defined by

$$\begin{aligned} t(\theta) &= \left(\sum_{j=0}^{2n} b_j \sin^j \frac{\omega - \theta}{2} \sin^{2n-j} \frac{\theta + \omega}{2} \right) \\ &\times \left(\frac{1}{\sin^2 \omega} \left(\sin^2 \frac{\omega - \theta}{2} + 2 \cos\omega \sin \frac{\omega - \theta}{2} \sin \frac{\theta + \omega}{2} + \sin^2 \frac{\theta + \omega}{2} \right) \right)^{m-n} \\ &= \sum_{j=0}^{2m} b_{j,m} \sin^j \frac{\omega - \theta}{2} \sin^{2m-j} \frac{\theta + \omega}{2}\,. \end{aligned}$$

Show that if $d_\omega(t)$ is finite, then it is the smallest value of m for which each $b_{j,m}$ is nonnegative or each $b_{j,m}$ is nonpositive.

Hint: The second factor in the representation of t is identically 1. □

We introduce the notation

$$G := \{z = x + iy : -\pi \le x < \pi,\ \ y \in \mathbb{R}\}$$

and

$$G_\omega := \{z = x + iy : \cos\omega \cosh y \ge \cos x,\ \ -\pi \le x < \pi,\ \ y \in \mathbb{R}\}\,.$$

b] Let $t \in \mathcal{T}_1 \setminus \mathcal{T}_0$. Show that $d_\omega(t) = 1$ if and only if t has its zeros in G_ω.

c] Assume $0 < \omega < \pi/2$, $t \in \mathcal{T}_1$, and $t(z) = 0$ for some

$$z := x + iy \in G \setminus (G_\omega \cup (-\omega, \omega))\,.$$

Show that

$$d_\omega(t) < \max \frac{4\sin(\omega \pm x)(\sin\omega\cosh y \mp \sin x)}{\cos\omega \sinh^2 y} - 1\,,$$

where the maximum is to be taken over both sets of signs.

d] Suppose $\pi/2 \le \omega \le \pi$, and $t \in \mathcal{T}_1$, $t(z) = 0$ for some $z \in G \setminus G_\omega$. Show that $d_\omega(t) = \infty$.

e] Show that $d_\omega(t_1 t_2) \le d_\omega(t_1) + d_\omega(t_2)$ for any two trigonometric polynomials t_1 and t_2.

f] Let $0 < \omega < \pi/2$ and $0 < \epsilon < \infty$. Show that if $t \in \mathcal{T}_n$ has no zeros in

$$E_{\omega,\epsilon} := \{z = x + iy : y^2 < \epsilon^2(\omega^2 - x^2)\,,\ \ x \in (-\omega, \omega)\,,\ \ y \in \mathbb{R}\}\,,$$

then

$$d_\omega(t) \le n\left(\frac{4}{\cos\omega}\epsilon^{-2} + 2\tan\omega + 1\right)\,.$$

g] Let p be a trigonometric polynomial and $0 < \omega < \pi/2$. Show that $d_\omega(t) < \infty$ if and only if $t = 0$ or t has no zeros in $(-\omega, \omega)$. (Note that part d] shows that this conclusion fails to hold when $\pi/2 \le \omega \le \pi$.)

h] Show that there is an absolute constant $c > 0$ (independent of n, ω and z_0) so that $d_\omega(t) \ge cn\epsilon^{-2}$ whenever

$$t(\theta) = \left(\sin\frac{\theta - z_0}{2}\sin\frac{\theta - \overline{z}_0}{2}\right)^n$$

with $z_0 \in \partial E_{\omega,\epsilon} \setminus \{-\omega, \omega\}$, $0 < \epsilon < \infty$, where $\partial E_{\omega,\epsilon}$ denotes the boundary of $E_{\omega,\epsilon}$ defined in part f].