Orthogonal polynomials and the moment problem

Henry Roland Steere

School of Mathematics

University of the Witwatersrand

Johannesburg

South Africa

Under the supervision of

Dr S. J. Johnston.

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2012.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

__________________________ __________________________

Name Student Number

______ day of _____________________ 2012

Abstract

The classical moment problem concerns distribution functions on the real line. The central feature is the connection between distribution functions and the moment sequences which they generate via a Stieltjes integral. The solution of the classical moment problem leads to the well known theorem of Favard which connects orthogonal polynomial sequences with distribution functions on the real line. Orthogonal polynomials in their turn arise in the computation of measures via continued fractions and the Nevanlinna parametrisation. In this dissertation classical orthogonal polynomials are investigated first and their connection with hypergeometric series is exhibited. Results from the moment problem allow the study of a more general class of orthogonal polynomials. q-Hypergeometric series are presented in analogy with the ordinary hypergeometric series and some results on q-Laguerre polynomials are given. Finally recent research will be discussed.

Acknowledgements

I acknowledge the support of the National Research Foundation of South Africa in funding my MSc.
I am grateful for the help and support of Dr Sarah Jane Johnston while researching the topic and compiling my dissertation.
Lastly I thank my family for their companionship.

Contents

Chapter 1

Classical orthogonal polynomials

The classical orthogonal polynomials were the first to be studied. Because they present the simplest case of orthogonality they are used in this chapter to introduce the various special characteristics of orthogonal polynomials. Hypergeometric series are presented because of their usefulness in expressing the classical polynomials. The Chebyshev polynomials are exhibited as a verification of the properties of orthogonal polynomials. The last part of the chapter presents the Jacobi, Gegenbauer, Hermite and Laguerre polynomials.

1.1 Basic theory of orthogonal polynomials

In an introduction to generalized Fourier series (cf. [27] p.43 for instance) a set $\{\phi_n\}_{n=0}^{\infty}$ of functions of a real variable is said to be orthonormal if

$$\int_a^b \phi_n \phi_m dx = \delta_{mn},$$

$$\text{where } \delta_{mn} = \begin{cases} 0, \text{ if } m \neq n, \\ 1, \text{ if } m \neq n. \end{cases}$$

Suppose that this set is a set of polynomials and the interval (a, b) is $(-1, 1)$. Let $P_0(x) = \frac{1}{\sqrt{2}}$. Define $p_1(x)$ as

$$p_1(x) = x - \frac{1}{\sqrt{2}} \int_{-1}^{1} x \frac{1}{\sqrt{2}} dx,$$

then

$$\int_{-1}^{1} p_1(x) \frac{1}{\sqrt{2}} dx = \int_{-1}^{1} x \frac{1}{\sqrt{2}} dx - \int_{-1}^{1} x \frac{1}{\sqrt{2}} dx \int_{-1}^{1} \frac{1}{2} dx = 0.$$

Define $P_1(x)$ by

$$P_1(x) = \frac{p_1(x)}{\sqrt{\int_{-1}^{1} p_1^2(x) dx}}$$

then

$$\int_{-1}^{1} P_1^2(x) dx = 1.$$

Let $p_k(x)$ be defined by

$$p_k(x) = x^k - \sum_{j=0}^{k-1} \left\{ \left[\int_{-1}^{1} x^k P_j(x) dx \right] P_j(x) \right\}$$

and let $P_k(x)$ be defined by

$$P_k(x) = \frac{p_k(x)}{\sqrt{\int_{-1}^{1} p_k^2(x) dx}}.$$

Assume that for $k, j < n$

$$\int_{-1}^{1} P_j(x) P_k(x) dx = \delta_{jk}.$$

Then

$$\begin{aligned}
&\int_{-1}^{1} p_n(x) P_k(x) dx \\
&= \int_{-1}^{1} x^n P_k(x) dx - \sum_{j=0}^{n-1} \left\{ \left[\int_{-1}^{1} x^n P_j(x) dx \right] \int_{-1}^{1} P_k(x) P_j(x) dx \right\} \\
&= \int_{-1}^{1} x^n P_k(x) dx - \int_{-1}^{1} x^n P_k(x) dx = 0
\end{aligned}$$

since $\int_{-1}^{1} P_j(x)P_k(x)dx = \delta_{jk}$ for $k < n$. Then by definition of $P_k(x)$, for $k \le n$

$$\int_{-1}^{1} P_n(x)P_k(x)dx = \delta_{nk}.$$

Because the case $k > n$ is the same as the above with the roles of n and k reversed, it is true in general that

$$\int_{-1}^{1} P_n(x)P_k(x)dx = \delta_{nk},$$

so that $\{P_n(x)\}_{n=0}^{\infty}$ is an orthonormal set of polynomials. This set of polynomials is known as the Legendre polynomials (cf. [41], p.86), and the algorithm used to obtain them is the famous Gram-Schmidt process (cf. [15], p.13).

Throughout the chapter when referring to a continuous Riemann integrable function $w(x)$ satisfying $w(x) > 0$ for $x \in (a, b)$ it will be assumed that

$$\int_a^b x^n w(x)dx < \infty$$

for $n = 0, 1, 2, \ldots$ (cf. [15], p.2).

Lemma 1.1.1. *Let $P_n(x)$ be an arbitrary real polynomial of degree n, $n = 0, 1, 2, \ldots$ and $w(x)$ be continuous and positive on (a, b). Then*

$$\int_a^b P_n^2(x)w(x)dx > 0.$$

Proof. No polynomial has infinitely many zeros. Because $w(x) > 0$ is continuous, there is an interval (between two possible zeros of $P_n(x)$, if any exist, otherwise on an arbitrary closed bounded subinterval of (a, b)) where the product $P_n^2(x)w(x)$ is greater than $\epsilon > 0$. If $\delta > 0$ is the length of this interval then the lower Riemann integral of this product is greater than or equal to $\delta\epsilon > 0$. □

Lemma 1.1.2 (cf. [15], p.2)**.** *Let $P_n(x)$ be an arbitrary real polynomial of degree n and $w(x)$ be continuous and positive on (a, b). The functional defined by*

$$\mu(P_n(x)) = \int_a^b P_n(x)w(x)dx$$

on the space of real polynomials (i.e. polynomials with real coefficients) is linear.

Proof. Let α be a real constant and $P_n(x)$, $P_m(x)$ be arbitrary real polynomials of degree n and m respectively. Then

$$\int_a^b \alpha P_n(x)w(x)dx = \alpha \int_a^b P_n(x)w(x)dx$$

and

$$\int_a^b (P_n(x) + P_m(x))w(x)dx = \int_a^b P_n(x)w(x)dx + \int_a^b P_m(x)w(x)dx$$

by the linearity of the Riemann integral. □

Definition 1.1.3 (cf. [42], p.150)**.** *Define* $\langle \cdot, \cdot \rangle$ *for the functional* μ *by*

$$\langle P_n(x), P_m(x) \rangle = \mu(P_n(x)P_m(x))$$

where $P_n(x)$ *and* $P_m(x)$ *are arbitrary real polynomials of degree* m *and* n *respectively and* $x \in \mathbb{R}$.

Lemma 1.1.4 (cf. [2], p2)**.** $\langle \cdot, \cdot \rangle$ *is an inner product on the space of real polynomials of a real variable.*

Proof. It is required to verify the following properties.

(a.) $\langle P_n(x), P_m(x) \rangle = \langle P_m(x), P_n(x) \rangle$,

(b.) $\langle \alpha_1 P_l(x) + \alpha_2 P_m(x), P_n(x) \rangle = \alpha_1 \langle P_l(x), P_n(x) \rangle + \alpha_2 \langle P_m(x), P_n(x) \rangle$,

(c.) $\langle P_n(x), P_n(x) \rangle > 0$ for $P_n(x)$ not identically zero,

where $P_l(x), P_m(x)$ and $P_n(x)$ are arbitrary polynomials of degrees l, m and n respectively and α_1, α_2 are arbitrary real numbers.
The first and second conditions are obvious consequences of Definition 1.1.3 in terms of the integral. The third condition follows from Lemma 1.1.1. □

With the inner product established it follows by application of the Gram-Schmidt process that there is a set of polynomials $\{P_n(x)\}_{n=0}^{\infty}$ where $P_n(x)$ has degree n for $n = 0, 1, 2, \ldots$ such that

$$\langle P_n(x), P_m(x) \rangle = \delta_{mn}. \tag{1.1.1}$$

Definition 1.1.5 (cf. [42], p.148)**.** *A set of polynomials $\{P_n(x)\}_{n=0}^{\infty}$, where $P_n(x)$ has degree n, satisfying*

$$\langle P_n(x), P_m(x) \rangle = h_n \delta_{nm} \tag{1.1.2}$$

where $h_n > 0$ and $\langle \cdot, \cdot \rangle$ is generated by $w(x) > 0$ is called a set of orthogonal polynomials with respect to the weight function $w(x)$.

In the general case of an orthogonal polynomial set it is not necessary that $h_n > 0$ (cf. [42], p.148). However, the most interesting work has been done under the assumption $h_n > 0$ and this assumption is used here. This assumption is entailed by the choice that $w(x) > 0$.

It is obvious from the linearity of the functional μ that any real multiple of an orthogonal polynomial is an orthogonal polynomial. In particular if in (1.1.2) $P_n(x)$ is divided by $h_n^{\frac{1}{2}}$, the resulting polynomial satisfies (1.1.1). The process of multiplying orthogonal polynomials by a real constant is called normalisation. This changes the value of h_n in (1.1.2).

Definition 1.1.6 (cf. [15], p.7)**.** *In the case where $h_n = 1$ the polynomials $P_n(x)$ are said to be orthonormal.*

Definition 1.1.7 (cf. [15], p.10)**.** *In the case where the leading coefficient of each polynomial in a set of orthogonal polynomials is 1, the polynomials are referred to as monic orthogonal polynomials.*

The following theorem will be proved in Chapter 2, Section 2.5.

Theorem 1.1.8 (cf. [15], p.12)**.** *A set of orthogonal polynomials is uniquely determined up to constant multiples.*

Polynomials constitute a vector space and the concept of a simple set arises as a natural consequence.

Definition 1.1.9 (cf. [42], p.147)**.** *If* $\{P_n(x)\}_{n=0}^{\infty}$ *is a set of polynomials such that* $P_n(x)$ *has degree* n *for each* $n = 0, 1, 2\ldots$*, then* $\{P_n(x)\}_{n=0}^{\infty}$ *is called a simple set.*

A simple set constitutes a linear basis for the space of polynomials.

Lemma 1.1.10 (cf. [42], p.147)**.** *Any polynomial can be expressed as a finite linear combination of polynomials from a simple set.*

Proof. Let $\{P_n(x)\}_{n=0}^{\infty}$ be a simple set. Let $p_m(x)$ be an arbitrary polynomial of degree m. Let A_m be the leading coefficient of $p_m(x)$ and B_m be the leading coefficient of $P_m(x)$. Then $p_m(x) - \frac{A_m}{B_m}P_m(x)$ is a polynomial of degree $m-1$ or less. This process can be applied to each new polynomial generated in this way, and the process stops once the constant term has been eliminated in this fashion, giving

$$p_m(x) - \sum_{k=0}^{m} c_k P_k(x) = 0,$$

where some of the c_k may be zero. □

If the simple set $\{P_n(x)\}_{n=0}^{\infty}$ is orthonormal then the coefficients c_k are determined by $c_k = \langle p_k(x), P_k(x)\rangle$ (cf. [2], p.18).

An equivalent and useful form of the orthogonality relation can be introduced as follows.

Theorem 1.1.11 (cf. [42], p.148)**.** *Let* $\{P_n(x)\}_{n=0}^{\infty}$ *be a simple set of polynomials. This set is an orthogonal set with respect to the weight function* $w(x)$ *continuous and positive on* (a, b) *if and only if it satisfies*

$$\int_a^b P_n(x)x^k w(x)dx = 0, \;\; k = 0, 1, 2, \ldots, n-1,$$

or, using the inner product,

$$\langle P_n(x), x^k\rangle = 0, \;\; k = 0, 1, 2, \ldots, n-1.$$

Proof. Suppose that $\{P_n(x)\}_{n=0}^{\infty}$ is an orthogonal set. By definition it is also a simple set. Consequently

$$x^k = \sum_{i=0}^{k} a_i P_i(x),$$

and

$$\int_a^b P_n(x)x^k w(x)dx = \int_a^b P_n(x) \sum_{i=0}^{k} a_i P_i(x) w(x)dx = 0$$

by linearity of μ and because $k < n$. If the condition

$$\int_a^b P_n(x)x^k w(x)dx = 0, \; k = 0, 1, 2, \ldots, n-1$$

holds, then because $\{x^k\}_{k=0}^{\infty}$ is a simple set,

$$P_k(x) = \sum_{i=0}^{k} b_i x^i,$$

and

$$\int_a^b P_n(x)P_k(x)w(x)dx = \int_a^b P_n(x) \sum_{i=0}^{k} b_i x^i w(x)dx = 0$$

for $k < n$ and then also for $k \neq n$ because if $k > n$ then n takes the place of k in the above. If $k = n$ then,

$$\int_a^b P_n^2(x)w(x)dx > 0,$$

since $w(x)$ is positive, continuous and Riemann integrable. □

If a simple set consists of polynomials with real coefficients and the polynomial which is to be expressed as a linear combination from that set also has real coefficients then the constants c_k in the expansion are real numbers.

Lemma 1.1.12 (cf [15], p.12)**.** *If $\{P_n(x)\}_{n=0}^{\infty}$ is a set of orthogonal polynomials and $p_n(x)$ is an arbitrary polynomial of degree n then*

$$\mu(P_n(x)p_n(x)) = k_n \mu(P_n(x)x^n),$$

where k_n is the leading coefficient of $p_n(x)$.

Proof. By Lemma 1.1.10

$$
\begin{aligned}
p_n(x) &= k_n x^n + \sum_{k=0}^{n-1} a_k P_k(x) \\
\mu(P_n(x)p_n(x)) &= \mu(P_n(x)\{k_n x^n + \sum_{k=0}^{n-1} a_k P_k(x)) \\
&= k_n \mu(P_n(x)x^n) + \sum_{k=0}^{n-1} a_k \mu(P_n(x)P_k(x)) \\
&= k_n \mu(P_n(x)x^n).
\end{aligned}
$$

□

Orthogonal polynomials satisfy several useful identities, one of which is the three-term recurrence relation.

Theorem 1.1.13 (cf. [5], p.244)**.** *Let $\{P_n(x)\}_{n=0}^{\infty}$ be a set of orthogonal polynomials corresponding to the functional μ (or the weight function $w(x)$ which generates μ), and let k_n be the leading coefficient of $P_n(x)$. Then there exist real sequences $\{a_n\}_{n=0}^{\infty}$, $\{b_n\}_{n=0}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$, such that for $n \geq 1$ the three-term recurrence relation*

$$P_{n+1}(x) = (a_n x + b_n)P_n(x) - c_n P_{n-1}(x),$$

$P_0(x) = k_0$, $P_{-1}(x) = 0$ holds. Here $a_n a_{n-1} c_n > 0$ for $n = 0, 1, 2 \ldots$, and if h_n is as in Definition 1.1.5 then

$$a_n = \frac{k_{n+1}}{k_n}, c_{n+1} = \frac{a_{n+1}}{a_n}\frac{h_{n+1}}{h_n}.$$

Proof. Select a_n so that $P_{n+1}(x) - a_n x P_n(x)$ has degree n or less. Then because $\{P_n(x)\}_{n=0}^{\infty}$ is a simple set there exist constants d_k such that

$$\sum_{k=0}^{n} d_k P_k(x) = P_{n+1}(x) - a_n x P_n(x).$$

Multiplying both sides of this equation by $P_k(x)$ for $k < n-1$ and applying μ gives $0 = d_k h_k$, and because $h_k > 0$ for $k \geq 0$, $d_k = 0$ for $k < n-1$. Let $c_n = -d_{n-1}$ and $b_n = d_n$ and the required recurrence results. The choice of

a_n immediately gives $a_n = \dfrac{k_{n+1}}{k_n}$. Multiply the recurrence relation by $P_{n-1}(x)$ and apply μ to obtain

$$0 = a_n \mu(P_n(x) x P_{n-1}(x)) - c_n \mu(P_{n-1}^2(x)). \tag{1.1.3}$$

The leading coefficient of $xP_{n-1}(x)$ is

$$k_{n-1} = \frac{k_{n-1}}{k_n} k_n,$$

so

$$xP_{n-1}(x) = \frac{k_{n-1}}{k_n} P_n(x) + \sum_{k=0}^{n-1} e_k P_k(x).$$

Substituting for $xP_{n-1}(x)$ in (1.1.3), gives

$$a_n \frac{k_{n-1}}{k_n} h_n = c_n h_{n-1}.$$

Recognising that $\dfrac{k_{n-1}}{k_n} = \dfrac{1}{a_{n-1}}$ and dividing by h_{n-1} gives the last part of the result. □

Alternative non-linear recurrence relations have been found for classical orthogonal polynomials (cf. [17]). The three-term recurrence relation establishes a connection between orthogonal polynomials and a particular Markoff process called a birth and death process (cf.[32], [33]). A two term recurrence relation can be constructed which connects orthogonal polynomials to inverse scattering theory (cf. [10]).

Another important relation satisfied by orthogonal polynomials is the Christoffel-Darboux formula.

Theorem 1.1.14 (cf. [5], p.246)**.** *If $\{P_n(x)\}_{n=0}^{\infty}$ is a set of orthogonal polynomials and h_i, k_i are as in Theorem 1.1.13 with $i = 0, 1, 2, \ldots$ then*

$$\sum_{i=0}^{n} \frac{P_i(x)P_i(y)}{h_i} = \frac{k_n}{k_{n+1}} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{(x-y)h_n}.$$

Proof. Multiplying the three-term recurrence relation in Theorem 1.1.13 by $P_n(y)$ gives

$$P_n(y)P_{n+1}(x) = (a_n x + b_n)P_n(x)P_n(y) - c_n P_{n-1}(x)P_n(y).$$

Swapping x with y gives

$$P_n(x)P_{n+1}(y) = (a_n y + b_n)P_n(y)P_n(x) - c_n P_{n-1}(y)P_n(x).$$

Subtract the second equation from the first and get

$$P_n(y)P_{n+1}(x) - P_n(x)P_{n+1}(y) = a_n(x-y)P_n(x)P_n(y)$$
$$+\frac{a_n}{a_{n-1}}\frac{h_n}{h_{n-1}}\left(P_{n-1}(y)P_n(x) - P_{n-1}(x)P_n(y)\right).$$

Divide by $a_n h_n(x-y)$ and take $P_n(x)P_n(y)$ to one side to obtain

$$\frac{P_n(x)P_n(y)}{h_n} = \frac{1}{a_n}\frac{1}{h_n}\frac{P_n(y)P_{n+1}(x) - P_n(x)P_{n+1}(y)}{x-y}$$
$$+\frac{1}{a_{n-1}}\frac{1}{h_{n-1}}\frac{P_{n-1}(x)P_n(y) - P_{n-1}(y)P_n(x)}{x-y}.$$

Summing the terms gives a telescoping series which establishes the result. □

Corollary 1.1.15 (cf [5], p.247)**.** *As a particular case of the Christoffel-Darboux formula*

$$\sum_{i=0}^{n}\frac{P_i^2(x)}{h_i} = \frac{k_n}{k_{n+1}}\frac{P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x)}{h_n}$$

Proof. This is obtained by taking the limit as $y \to x$ and using l'Hôpital's rule. □

If $k_n < 0$ for any n then multiplying $P_n(x)$ by -1 makes $k_n > 0$ and doesn't substantially change the three-term recurrence relation. In what follows it will be assumed that $k_n > 0$.

Because $h_n > 0$ and $k_n > 0$ for all n,

$$P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x) > 0 \tag{1.1.4}$$

for all $x \in \mathbb{R}$ (cf [5], p.247).

The zeros of real line orthogonal polynomials satisfy several useful properties.

Theorem 1.1.16 (cf [15], p.27)**.** *The zeros of real line orthogonal polynomials are real, simple and are contained in (a, b), where (a, b) is the interval of orthogonality.*

Proof. $\mu(P_n(x)) = 0$ so it can't be the case that $P_n(x) \geq 0$ on (a, b). $P_n(x)$ has at least one zero of odd multiplicity in (a, b). Let $\pi(x)$ be the polynomial

$$\pi(x) = (x - x_1)(x - x_2) \ldots (x - x_m)$$

where $x_1, x_2, \ldots, x_m$ are the zeros of odd multiplicity of $P_n(x)$ in (a, b). If $m < n$ then $\mu(P_n(x)\pi(x)) = 0$ by orthogonality. But $P_n(x)\pi(x)$ doesn't change sign on (a, b). If $P_n(x)\pi(x) \leq 0$ then $-P_n(x)\pi(x) \geq 0$, giving $\mu(-P_n(x)\pi(x)) > 0$ or $\mu(P_n(x)\pi(x)) < 0$. If $P_n(x)\pi(x) \geq 0$, it follows that $\mu(P_n(x)\pi(x)) > 0$. Either way there is a contradiction. So $m = n$ i.e. all the zeros are simple and contained in (a, b) and therefore real as well. $\square$

Theorem 1.1.17 (cf [15], p.28)**.** *If $\{P_n(x)\}_{n=0}^{\infty}$ is a set of orthogonal polynomials then the zeros of $P_n(x)$ and $P_{n+1}(x)$ interlace for all n, i.e. for every two consecutive zeros of $P_{n+1}(x)$ there is one zero of $P_n(x)$ between them.*

Proof. From (1.1.4), it follows that $P'_{n+1}(x)P_n(x) > 0$ for each zero of $P_{n+1}(x)$. Since, by Rolle's theorem, $P'_{n+1}(x)$ changes sign between each of the zeros of $P_{n+1}(x)$, $P_n(x)$ also changes sign between each of these zeros and must have a zero between each of them. $\square$

1.2 Hypergeometric series

Gauss was the first to propose the study of series of hypergeometric type. In 1812 he presented a paper which considered the series (cf. [20]),

$$1 + \frac{ab}{1!c}z + \frac{a(a+1)b(b+1)}{2!c(c+1)}z^2 + \ldots.$$

This sum converges to an analytic function (in some domain) and is denoted by ${}_2F_1(a, b; c; z)$ or

$$ {}_2F_1 \left[\begin{matrix} a, & b \\ & & ; z \\ c & \end{matrix} \right]. $$

The hypergeometric series arises in the theory of differential equations (cf. [43], [55]), and can also be used for the representation of several important sets of orthogonal polynomials.

Definition 1.2.1 (cf. [20], p. xii)**.** *The symbol* $(a)_n$ *is called Pochammer's symbol and denotes the product*

$$ (a)_n = \prod_{k=1}^{n} (a + k - 1) \tag{1.2.1} $$

where

$$ (a)_0 = 1. $$

With the Pochammer symbol defined a more concise definition of the hypergeometric series is possible.

Definition 1.2.2 (cf. [42], p.45)**.** ${}_2F_1(a, b; c; z)$ *denotes the series*

$$ \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n n!} z^n. $$

A natural question to ask is: when does the above series converge? The ratio test suffices to answer this question.

Theorem 1.2.3 (cf. [42], p.45)**.** *If* a, b *and* c *are neither negative integers nor zero, then* ${}_2F_1(a, b; c; z)$ *converges for* $|z| < 1$.

Proof. Using the ratio test gives,

$$ \lim_{n \to \infty} \left| \frac{(a)_{n+1} (b)_{n+1} z^{n+1}}{(c)_{n+1} (n+1)!} \frac{(c)_n n!}{(a)_n (b)_n z^n} \right| = \lim_{n \to \infty} \left| \frac{(a+n)(b+n)}{(c+n)(n+1)} z \right| = |z| < 1. \quad \square $$

Convergence on the unit circle requires conditions on the parameters of the series.

Theorem 1.2.4 (cf. [42], p.46)**.** *For* $|z| = 1$, ${}_2F_1(a, b; c; z)$ *converges for* $\Re(c - a - b) > 0$.

The proof of this theorem can be found in [42] on page 46.

Although not a hypergeometric function, the gamma function features commonly in identities concerning hypergeometric functions.

Definition 1.2.5 (cf. [5], p.6)**.** *For* $\Re(x) > 0$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

Using analytic continuation the gamma function can be made analytic everywhere except for poles at the negative integers (cf. [5], p.7). Using the definition of the Gamma function as an integral it is simple to obtain a well known reduction formula of the gamma function.

Theorem 1.2.6.

$$\Gamma(x + 1) = x\Gamma(x).$$

Proof. Using integration by parts

$$\int_0^\infty t^x e^{-t} dt = [t^x(-e^{-t})]_0^\infty - x \int_0^\infty t^{x-1}(-e^{-t}) dt.$$

The first term on the right disappears and the second term is $x\Gamma(x)$, as required. □

Another identity which will be used in this discussion is the reflection formula.

Theorem 1.2.7 (cf. [5], p.9)**.**

$$\Gamma(x)\Gamma(1 - x) = \frac{\pi}{\sin(\pi x)}.$$

The proof uses contour integration and the relation

$$\frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)} = \int_0^\infty \frac{s^{x-1}}{(1 + s)^{x+y}} ds.$$

Full details can be found in [5] on page 9.

As already mentioned, the hypergeometric series arises in the theory of differential equations. Specifically it arises as the solution of the so-called hypergeometric equation (cf. [43]).

Definition 1.2.8 (cf. [43], p.169)**.** *The second order differential equation,*

$$x(1-x)y'' + [\gamma - (\alpha+\beta+1)x]y' - \alpha\beta y = 0, \tag{1.2.2}$$

where α, β and γ are complex constants, is referred to as the hypergeometric equation.

The equation (1.2.2) has a regular singular point at $x = 0$. Consequently, the Frobenius method can be used to obtain a power series solution (cf. [43], [27]). In particular (cf. [43], p.169) there is a solution with exponent zero, i.e. of the form

$$y = \sum_{n=0}^{\infty} a_n x^n.$$

Following the usual steps in power series solutions (cf. [43], p.82)

$$y' = \sum_{n=1}^{\infty} n a_n x^{n-1}, \quad y'' = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2},$$

so that the equation becomes

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-1} - \sum_{n=2}^{\infty} n(n-1)a_n x^n + \gamma \sum_{n=1}^{\infty} n a_n x^{n-1}$$

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

Note that

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-1} = \sum_{n=0}^{\infty} n(n-1)a_n x^{n-1}$$

$$\gamma \sum_{n=1}^{\infty} n a_n x^{n-1} = \gamma \sum_{n=0}^{\infty} n a_n x^{n-1}$$

and

$$\sum_{n=2}^{\infty} n(n-1)a_n x^n = \sum_{n=0}^{\infty} n(n-1)a_n x^n$$

$$(\alpha+\beta+1)\sum_{n=1}^{\infty} n a_n x^n = (\alpha+\beta+1)\sum_{n=0}^{\infty} n a_n x^n.$$

Equating coefficients of x^n gives, for $n \geq 1$

$$a_n = \frac{(\alpha+n-1)(\beta+n-1)}{n(\gamma+n-1)} a_{n-1}.$$

In particular, if $a_0 = 1$ the result is

$$a_n = \frac{(\alpha)_n(\beta)_n}{n!(\gamma)_n},$$

i.e. the solution is a hypergeometric series, which explains the name given to the equation.

If a, or b is a negative integer, say $-n$ then ${}_2F_1(a,b;c;x)$ is a terminating series and represents a polynomial in x with degree n. This can be observed by realising that

$$(-n)_{n+1} = \prod_{k=0}^{n}(-n+k) = 0.$$

An important use of the hypergeometric series is its role as a representation for the classical orthogonal polynomials.

Definition 1.2.9 (cf. [42], p.73)**.** *The generalised hypergeometric function is denoted* ${}_pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z)$ *or*

$${}_pF_q\left[\begin{array}{cccc} a_1, & a_2, & \ldots, & a_p \\ b_1, & b_2, & \ldots, & b_q \end{array}; z\right] = \sum_{n=0}^{\infty} \frac{\prod_{i=1}^{p}(a_i)_n}{\prod_{j=1}^{q}(b_j)_n} \frac{z^n}{n!}.$$

Here p stipulates the number of numerator parameters and q the number of denominator parameters.

An important special case of the generalised hypergeometric function is obtained from the binomial theorem.

The binomial theorem gives the result (cf. [42], p.47),

$$\begin{aligned}(1-x)^{-a} &= \sum_{n=0}^{\infty} \frac{(-a)(-a-1)(-a-2)\dots(-a-n+1)(-1)^n x^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{a(a+1)(a+2)\dots(a+n-1)x^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(a)_n x^n}{n!} = {}_1F_0(a;-;x).\end{aligned}$$

1.3 Chebyshev polynomials

A simple trigonometric identity can be used to derive the orthogonality relation for the Chebyshev polynomials.

Lemma 1.3.1 (cf. [15], p.1).

$$\int_0^{\pi} \cos m\theta \cos n\theta d\theta = \frac{\pi}{2}\delta_{mn} \tag{1.3.1}$$

except for $n = m = 0$, in which case the integral gives π.

Proof. Using a familiar trigonometric identity,

$$\cos m\theta \cos n\theta = \frac{1}{2}\left\{\cos(m+n)\theta + \cos(m-n)\theta\right\}. \tag{1.3.2}$$

For $m \neq n$

$$\begin{aligned}&\int_0^{\pi} \cos m\theta \cos n\theta d\theta \\ &= \frac{1}{2}\left\{ \frac{1}{m+n}\sin(m+n)\theta\Big|_0^{\pi} + \frac{1}{m-n}\sin(m-n)\theta\Big|_0^{\pi}\right\} = 0,\end{aligned}$$

and for $m = n$, $\cos(m-n)\theta = \cos 0 = 1$, so the integral reduces to

$$\frac{1}{2}\int_0^{\pi} d\theta = \frac{\pi}{2}.$$

If $m = n = 0$ then

$$\int_0^{\pi} \cos 0 \cos 0 d\theta = \int_0^{\pi} d\theta = \pi. \qquad \square$$

The next step in arriving at the Chebyshev polynomials is the fact that $\cos n\theta$ is a polynomial in powers of $\cos\theta$.

Lemma 1.3.2 (cf. [15], p.2)**.** *For n a natural number, $\cos n\theta$ is a polynomial in powers of $\cos\theta$ with degree n, where degree refers to the highest power in $\cos\theta$.*

Proof. The proof uses induction. $\cos 0 = 1$ and $\cos 1\theta$ is a polynomial in $\cos\theta$ of degree 1 trivially. Suppose the statement is true for all $m < n$. By (1.3.2) $\cos n\theta = \cos(n-1+1)\theta = 2\cos(n-1)\theta\cos\theta - \cos(n-2)\theta$. $\cos(n-1)\theta$ has degree $n-1$ by the inductive hypothesis, so $\cos(n-1)\theta\cos\theta$ has degree n and the remaining term does not affect the degree. $\square$

Definition 1.3.3 (cf. [41], p.71)**.** *Using Lemmas 1.3.1 and 1.3.2, the n^{th} Chebyshev polynomial $T_n(x)$ is defined by*

$$T_n(x) = \cos n\theta$$

where $x = \cos\theta$.

Theorem 1.3.4 (cf. [15], pp.71, 252)**.** *The Chebyshev polynomials $T_n(x)$ satisfy the orthogonality relation*

$$\int_{-1}^{1} T_n(x)T_m(x)\frac{1}{\sqrt{1-x^2}}dx = \frac{\pi}{2}\delta_{mn}$$

except for $n = m = 0$ in which case the integral is equal to π.

Proof. All that is required, is to realise that the substitution $x = \cos\theta$ in (1.3.1) results in $dx = -\sin\theta d\theta$ or

$$d\theta = -\frac{1}{\sqrt{1-\cos^2\theta}}dx = -\frac{1}{\sqrt{1-x^2}}dx. \qquad \square$$

This theorem establishes the weight function,

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

which is positive on the interval $(-1, 1)$ as corresponding to the orthogonality relation of the Chebyshev polynomials.

The relationship $T_n(x) = \cos n\theta$ allows the exact determination of the zeros of $T_n(x)$ for arbitrary n.

Lemma 1.3.5 (cf. [41], pp.71, 252)**.** *The zeros of* $T_n(x)$ *are*

$$\cos\left(\frac{(2j-1)\pi}{2n}\right)$$

$j = 1, 2, \ldots, n$.

Proof. This statement follows from the fact that the zeros of $\cos n\theta$ in the interval $(0, \pi)$ occur where $n\theta = (2j-1)\frac{\pi}{2}$, or

$$\theta = \frac{(2j-1)\pi}{2n},$$

$j = 1, 2, \ldots n$. Using the relation $x = \cos\theta$ then gives the zeros of $T_n(x)$. There are n zeros (so they are simple) and they lie in $(-1, 1)$ (so they are real). □

The above result demonstrates the phenomenon of real, simple zeros contained in the interval of orthogonality, which characterises orthogonal polynomials.

The polynomials $T_n(x)$ are, strictly speaking, the Chebyshev polynomials of the first kind. The Chebyshev polynomials of the second kind are a closely related family of orthogonal polynomials, which are denoted by $U_n(x)$.

Lemma 1.3.6 (cf. [41], p.71)**.**

$$\frac{\sin(n+1)\theta}{\sin\theta}$$

is a polynomial in powers of $\cos(\theta)$ *with degree* n.

Proof. Again using induction,

$$\frac{\sin 1\theta}{\sin\theta} = 1$$

is a polynomial in $\cos\theta$ with degree 0. Supposing the statement is true for $m < n$, the elementary identity from trigonometry

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

gives

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

or

$$\frac{\sin(n+1)\theta}{\sin\theta} = 2\frac{\sin n\theta}{\sin\theta}\cos\theta - \frac{\sin(n-1)\theta}{\sin\theta}. \tag{1.3.3}$$

The first term on the right hand side of (1.3.3) is, by hypothesis, a polynomial in $\cos\theta$ of degree $n-1$ multiplied by $\cos\theta$ (i.e. has degree n). The second term on the right is a polynomial in $\cos\theta$ of degree $n-2$ and substracting does not affect the degree. $\square$

Definition 1.3.7 (cf. [15], p.5)**.** *The Chebyshev polynomial of the second kind is denoted by $U_n(x)$ and*

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}$$

for $x = \cos\theta$.

Theorem 1.3.8 (cf. [41], p.71)**.** *The Chebyshev polynomials of the second kind $U_n(x)$ satisfy the orthogonality relation*

$$\int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2}dx = \frac{\pi}{2}\delta_{mn}.$$

Proof. In a similar approach to the proof for the Chebyshev polynomials of the first kind the trigonometric representation of the polynomials is used.

$$\begin{aligned}
&\int_0^{\pi} \sin(n+1)\theta \sin(m+1)\theta d\theta \\
&= \int_0^{\pi} \frac{\sin(n+1)\theta}{\sin\theta}\frac{\sin(m+1)\theta}{\sin\theta}\sin^2\theta d\theta \\
&= \int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2}dx.
\end{aligned}$$

Because $x = \cos\theta$ and $dx = -\sin\theta d\theta$. Using the identity

$$\sin(n+1)\theta\sin(m+1)\theta = \frac{1}{2}(\cos(m-n)\theta - \cos(m+n)\theta)$$

the same reasoning as in the proof for $T_n(x)$ shows that the orthogonality condition above holds, and furthermore the exceptional case that occurred for the $T_n(x)$ where $n = m = 0$ does not occur because $\sin((n+1)\theta)$ is indexed from $n+1$ rather than n. □

Lemma 1.3.9 (cf. [41], p.71)**.** *The polynomials $T_n(x)$ and $U_n(x)$ satisfy the recurrence relations,*

$$T_{n+1}(x) = xT_n(x) - (1-x^2)U_{n-1}(x) \tag{1.3.4}$$

and

$$U_n(x) = xU_{n-1}(x) + T_n(x) \tag{1.3.5}$$

Proof. Again invoking trigonometric identities gives the results,

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

or

$$\cos(n+1)\theta = \cos n\theta\cos\theta - (1-\cos^2\theta)\frac{\sin n\theta}{\sin\theta}$$

substituting x, $T_n(x)$ and $U_n(x)$ in the above gives the first result. Also,

$$\sin(n+1)\theta = \cos\theta\sin n\theta + \sin\theta\cos n\theta$$

or

$$\frac{\sin(n+1)\theta}{\sin\theta} = \cos\theta\frac{\sin n\theta}{\sin\theta} + \cos n\theta$$

giving the second result. □

Theorem 1.3.10 (cf. [15], p.20)**.** *The Chebyshev polynomials $T_n(x)$ satisfy a three-term recurrence relation of the form*

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1.$$

Proof. Using the identity

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

which is a particular case of (1.3.2) gives the result on substituting $x = \cos\theta$ and using $T_n(x) = \cos n\theta$. $\square$

Theorem 1.3.11. *The Chebyshev polnomials of the second kind $U_n(x)$ satisfy the three-term recurrence relation*

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n \geq 1.$$

Proof. The relation (1.3.5) immediately gives

$$T_n(x) = U_n(x) - xU_{n-1}(x).$$

Substituting for $T_i(x)$, $i = n$, $n+1$ in (1.3.4) results in

$$U_{n+1}(x) - xU_n(x) = xU_n(x) - x^2U_{n-1}(x) - (1-x^2)U_{n-1}(x),$$

which, after cancelling terms, gives the result. $\square$

1.4 Other classical polynomials

The classical orthogonal polynomials are important in various applications and are the most thoroughly studied. The polynomials bear the names of the famous mathematicians who studied them: Jacobi, Legendre, Laguerre, Hermite, Gegenbauer and Chebyshev.

In the hypergeometric equation (1.2.2), setting t as the independent variable and replacing α, with $-n$, β with $n+\alpha+\beta+1$ and γ with $\alpha+1$, the resulting equation is (cf. [49], p.62)

$$t(1-t)y'' + [\alpha+1-(\alpha+\beta+2)t]y' + n(n+\alpha+\beta+1)y = 0. \qquad (1.4.1)$$

For the purposes of this section the constants α, β and γ in the above equation are real numbers. This hypergeometric equation has a solution of the form,

$${}_2F_1(-n, n+\alpha+\beta+1; \alpha+1; t).$$

Applying the substitution $\frac{1-x}{2} = t$ to the equation (1.4.1) (cf. [49], p.60) with $-\frac{1}{2}\frac{dy}{dt} = \frac{dy}{dx}$ and denoting $\frac{dy}{dx}$ by y' leads to the differential equation

$$(1-x^2)y'' + [\beta - \alpha - (\alpha+\beta+2)x]y' + n(n+\alpha+\beta+1)y = 0. \qquad (1.4.2)$$

This establishes that (1.4.2) has a solution of the form

$${}_2F_1\left(-n, n+\alpha+\beta+1; \alpha+1; \frac{1-x}{2}\right),$$

which by previous considerations is a polynomial of degree n.

Definition 1.4.1 (cf. [50], p.151)**.** *The Jacobi polynomials are given by the hypergeometric series*

$$P_n^{(\alpha,\beta)}(x) = \binom{n+\alpha}{n} F\left(-n, \alpha+\beta+n+1; \alpha+1; \frac{1-x}{2}\right), \qquad (1.4.3)$$

where $\alpha > -1$ and $\beta > -1$.

In the above (cf. [50], p.73)

$$\binom{\alpha+n}{n} = \frac{(\alpha+1)(\alpha+2)\dots(\alpha+n)}{n!}$$

is a normalisation constant and refers to the binomial coefficient which can be expressed as a quotient of gamma functions by

$$\binom{\alpha+n}{n} = \frac{\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+1)}.$$

Lemma 1.4.2 (cf. [41], p.88)**.** *Up to normalisation $P_n^{(0,0)}(x) = P_n(x)$ where $P_n(x)$ is the n^{th} Legendre polynomial. Similarly disregarding normalisation $P_n^{(-\frac{1}{2},-\frac{1}{2})}(x) = T_n(x)$ and $P_n^{(\frac{1}{2},\frac{1}{2})}(x) = U_n(x)$.*

These polynomials are known as ultraspherical or Gegenbauer polynomials (cf. [42], p.276).

Definition 1.4.3 (cf. [42], p.276)**.** *In the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ set $\alpha = \beta$. The resulting polynomial $P_n^{(\alpha,\alpha)}(x)$ is called an ultrashperical polynomial.*

Definition 1.4.4 (cf. [42], p.277)**.** *The polynomials* $\{C_n^\nu(x)\}_{n=0}^\infty$ *defined by*

$$C_n^\nu(x) = \frac{(2\nu)_n P_n^{(\nu-\frac{1}{2},\nu-\frac{1}{2})}}{(\nu+\frac{1}{2})_n}$$

are called the Gegenbauer polynomials.

Because orthogonal polynomials are determined up to constant multiples it is clear that the Gegenbauer polynomials are essentially the same as the ultraspherical polynomials.

Theorem 1.4.5 (cf. [42], p.258)**.** *The Jacobi polynomials* $P_n^{(\alpha,\beta)}(x)$*, satisfy the orthogonality relation*

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta dx = h_n\delta_{mn}, \quad h_n > 0, \tag{1.4.4}$$

where α*,* $\beta > -1$*.*

Proof. By definition the Jacobi polynomials satisfy the differential equation

$$\begin{aligned}&(1-x^2)\frac{d^2}{dx^2}P_n^{(\alpha,\beta)}(x) + [\beta-\alpha-(\alpha+\beta+2)x]\frac{d}{dx}P_n^{(\alpha,\beta)}(x)\\&+n(n+\alpha+\beta+1)P_n^{(\alpha,\beta)}(x) = 0.\end{aligned} \tag{1.4.5}$$

Using the fact that $\beta-\alpha-(\alpha+\beta+2)x = (1+\beta)(1-x)-(1+\alpha)(1+x)$ and multiplying by $(1-x)^\alpha(1+x)^\beta$, (1.4.5) can be rewritten

$$\begin{aligned}&(1-x)^{\alpha+1}(1+x)^{\beta+1}\frac{d^2}{dx^2}P_n^{(\alpha,\beta)}(x)\\&+[(1+\beta)(1-x)-(1+\alpha)(1+x)](1-x)^\alpha(1+x)^\beta\frac{d}{dx}P_n^{(\alpha,\beta)}(x)\\&+n(n+\alpha+\beta+1)(1-x)^\alpha(1+x)^\beta P_n^{(\alpha,\beta)}(x) = 0.\end{aligned} \tag{1.4.6}$$

By the product rule (1.4.6) is

$$\begin{aligned}&\frac{d}{dx}[(1-x)^{1+\alpha}(1+x)^{1+\beta}\frac{d}{dx}P_n^{(\alpha,\beta)}(x)]\\&+n(1+\alpha+\beta+n)(1-x)^\alpha(1+x)^\beta P_n^{(\alpha,\beta)}(x) = 0.\end{aligned} \tag{1.4.7}$$

Multiply (1.4.7) by $P_m^{(\alpha,\beta)}(x)$ and substract the same equation with n replaced by m and multiplied by $P_n^{(\alpha,\beta)}(x)$ to get equation (1.4.8).

[Hint: to obtain equation (1.4.8) perform the first differentiation on the right hand side]

$$[n(1+\alpha+\beta+n)-m(1+\alpha+\beta+m)](1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)$$
$$=\frac{d}{dx}[(1-x)^{1+\alpha}(1+x)^{1+\beta}\{P_n^{(\alpha,\beta)}(x)\frac{d}{dx}P_m^{(\alpha,\beta)}(x)-P_m^{(\alpha,\beta)}(x)\frac{d}{dx}P_n^{(\alpha,\beta)}(x)\}]. \tag{1.4.8}$$

Finally integrate both sides of (1.4.8) to get

$$(n-m)(1+\alpha+\beta+n+m)\int_{-1}^{1}P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)(1-x)^{\alpha}(1+x)^{\beta}dx$$
$$=\left[(1-x)^{1+\alpha}(1+x)^{1+\beta}\{P_n^{(\alpha,\beta)}(x)\frac{d}{dx}P_m^{(\alpha,\beta)}(x)-P_m^{(\alpha,\beta)}(x)\frac{d}{dx}P_n^{(\alpha,\beta)}(x)\}\right]_{-1}^{1}.$$

It follows that for $m \neq n$ (1.4.4) is 0 while for $m = n$ the integral is positive because $(1-x)^{\alpha}(1+x)^{\beta}$ is continuous and positive over $(-1,1)$. □

By examining the weight functions of the Legendre polynomials and the polynomials $T_n(x)$ and $U_n(x)$ it is established that these polynomials are special cases of the Jacobi polynomials. In fact they are ultraspherical polynomials.

An alternative approach to the Jacobi polynomials defines them in terms of the orthogonality relation. By demonstrating that they satisfy the differential equation (1.4.5) and that they are the only solution of this equation, it follows that they have the representation (1.4.3).

Theorem 1.4.6 (cf. [49], p.60)**.** *Let $P_n^{(\alpha,\beta)}(x)$ be a polynomial that satisfies the orthogonality relation (1.4.4). Then $P_n^{(\alpha,\beta)}(x)$ satisfies the differential equation,*

$$(1-x^2)y''+[\beta-\alpha-(\alpha+\beta+2)x]y'+n(n+\alpha+\beta+1)y=0,$$

Proof. First note that

$$\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}y']+n(n+\alpha+\beta+1)(1-x)^{\alpha}(1+x)^{\beta}y$$

$$= -(1-x)^\alpha(1+x)^\beta(\alpha+1)(1+x)y' + (1-x)^\alpha(1+x)^\beta(\beta+1)(1-x)y'$$

$$+(1-x^2)(1-x)\alpha(1+x)^\beta y'' + (1-x)\alpha(1+x)^\beta n(n+\alpha+\beta+1)y$$

which after collecting coefficients of $y^{(i)}$ and dividing by $(1-x)^\alpha(1+x)^\beta$ gives (1.4.2). Assume that $y = P_n^{(\alpha,\beta)}(x)$.

$$\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}y'] = -(\alpha+1)(1-x)^\alpha(1+x)^\beta(1+x)y'$$

$$+(\beta+1)(1-x)^\alpha(1+x)^\beta(1-x)y' + (1-x)^\alpha(1+x)^\beta(1-x^2)y''$$

i.e.

$$\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}y'] = (1-x)^\alpha(1+x)^\beta z, \tag{1.4.9}$$

where z is a polynomial of degree n. Showing that z satisfies the orthogonality relation of $P_n^{(\alpha,\beta)}(x)$ establishes that $z = AP_n^{(\alpha,\beta)}(x)$, where A is a constant. Let $\pi(x)$ be an arbitrary polynomial of degree $< n$, then

$$\int_{-1}^{1}\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}y']\pi(x)dx$$
$$= (1-x)^{\alpha+1}(1+x)^{\beta+1}y'\pi(x)|_{-1}^{1} - \int_{-1}^{1}(1-x)^{\alpha+1}(1+x)^{\beta+1}y'\pi'(x)dx.$$

By definition $\alpha+1 > 0$ and $\beta+1 > 0$, so evaluating the first term at -1 and 1 causes it to disappear. This also happens integrating by parts a second time, giving

$$\int_{-1}^{1} y\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}\pi'(x)]dx.$$

If $\pi(x)$ has degree 0 this vanishes, otherwise, by the same reasoning as in the calculation of (1.4.9), this is equal to

$$\int_{-1}^{1} yp(x)(1-x)^\alpha(1+x)^\beta dx$$

where $p(x)$ has degree $< n$. As a result the integral is zero, so z satisfies the orthogonality condition and is equal to Ay, where A is a constant. Using the derivation of (1.4.9) the leading coefficient of z is

$$(-\alpha-1)nk_n + (-\beta-1)nk_n + (-n(n-1))k_n,$$

where k_n is the leading coefficient of y. This gives that $A = -n(n+\alpha+\beta+1)$, which proves that $P_n^{(\alpha,\beta)}(x)$ satisfies the differential equation. $\square$

It can be established that $P_n^{(\alpha,\beta)}(x)$ is the only polynomial solution of (1.4.2).

Theorem 1.4.7 (cf. [49], p.61)**.** *If y is a polynomial solution of*

$$(1-x^2)y'' + [\beta - \alpha - (\alpha+\beta+2)x]y' + n(n+\alpha+\beta+1)y = 0$$

then it is a constant multiple of $P_n^{(\alpha,\beta)}(x)$.

The proof of this result can be found in [49] on page 61.

From these results the previously given representation of the Jacobi polynomials as a hypergeometric series is established.

Definition 1.4.8 (cf. [42], p.187)**.** *The Hermite polynomials $\{H_n(x)\}_{n=0}^{\infty}$ are defined by the generating function*

$$e^{2xt-t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)t^n}{n!}.$$

Theorem 1.4.9 (cf. [42], p. 189)**.** *$H_n(x)$ satisfies the Rodrigues' formula*

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

Proof. Because $H_n(x)$ is the Taylor coefficient of t^n in the expansion of e^{2xt-t^2} as a Maclaurin series in t it follows that

$$H_n(x) = \left[\frac{d^n}{dt^n} e^{2xt-t^2}\right]_{t=0}.$$

The differentiation is with respect to t, so

$$e^{-x^2} H_n(x) = \left[\frac{d^n}{dt^n} e^{-(x-t)^2}\right]_{t=0}.$$

Let $x - t = w$ so that

$$e^{-x^2} H_n(x) = (-1)^n \left[\frac{d^n}{dw^n} e^{-w^2}\right]_{w=x}.$$

The result follows by multiplying both sides by e^{x^2}. $\square$

As a consequence of Theorem 1.4.9 the orthogonality relation for the Hermite polynomials can be derived.

Theorem 1.4.10. *The Hermite polynomials satisfy the orthogonality relation*

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}dx = h_n\delta_{mn}, \tag{1.4.10}$$

$h_n > 0$.

Proof. Using the Rodrigues' formula, the integral (1.4.10) reduces to

$$\begin{aligned}&\int_{-\infty}^{\infty}(-1)^{n+m}e^{x^2}\frac{d^n}{dx^n}e^{-x^2}\frac{d^m}{dx^m}e^{-x^2}dx\\ &= (-1)^n\int_{-\infty}^{\infty}\frac{d^n}{dx^n}e^{-x^2}H_m(x)dx. \end{aligned}\tag{1.4.11}$$

From

$$\frac{d^k}{dx^k}e^{-x^2} = P(x)e^{-x^2}$$

for some polynomial $P(x)$, it follows that

$$\left[\frac{d^k}{dx^k}e^{-x^2}\right]_{x=-\infty}^{x=\infty} = 0$$

for any natural number k. If $m < n$ then integration by parts can be used to eliminate $H_m(x)$ from the integral (1.4.11) so that the integral is equal to zero. If $m = n$ then because e^{-x^2} is positive and continuous and $H_n^2(x) \geq 0$ and not identically zero the integral is positive. □

Let $\{P_n^{(1)}(x)\}_{n=0}^{\infty}$ be the unique (up to normalisation) set of polynomials satisfying,

$$\int_0^{\infty} P_n^{(1)}(x)P_m^{(1)}(x)x^{-\frac{1}{2}}e^{-x}dx = h_n\delta_{nm},$$

and let $\{P_n^{(2)}(x)\}_{n=0}^{\infty}$ be the unique set of polynomials satsifying

$$\int_0^{\infty} P_n^{(2)}(x)P_m^{(2)}(x)x^{\frac{1}{2}}e^{-x}dx = h_n\delta_{nm}.$$

$\{H_n(x)\}_{n=0}^{\infty}$ is closely related to these sets.

Theorem 1.4.11 (cf. [41], p.88). *The relation* $H_{2n}(x) = AP_n^{(1)}(x^2)$ *and the relation* $H_{2n+1}(x) = BxP_n^{(2)}(x^2)$ *hold, where* A *and* B *are constants depending on normalisation.*

Proof. The orthogonality condition is established. For odd exponents of x less than $2n$

$$\int_{-\infty}^{\infty} P_n^{(1)}\left(x^2\right) x^{2k+1} e^{-x^2} dx = 0$$

because $P_n^{(1)}(x^2)$ and e^{-x^2} are even functions while x^{2k+1} is an odd function so the integrand is odd and disappears. For even exponents of x less than $2n$, the substitution $t = x^2$ gives $dt = 2xdx$, $x^{2k-1} = t^{k-\frac{1}{2}}2^{k-\frac{1}{2}}$. By the same reasoning as above the integrand is even, so equal to twice the integral from 0 to ∞. Carrying out the substitution gives

$$\int_{-\infty}^{\infty} P_n^{(1)}\left(x^2\right) x^{2k} e^{-x^2} dx = 2^{k-\frac{1}{2}} \int_{0}^{\infty} P_n^{(1)}(t) t^{k-\frac{1}{2}} e^{-t} dt = 0$$

by the orthogonality condition for $P_n^{(1)}(x)$, so the first part of the result is established. For even exponents of x less than $2n$

$$\int_{-\infty}^{\infty} xP_n^{(2)}\left(x^2\right) x^{2k} e^{-x^2} dx = 0,$$

because x is an odd function and the other functions in the integrand are even so the integrand is odd. For odd exponents of x less than $2n$, the same substitution as in the previous case results in,

$$\int_{-\infty}^{\infty} xP_n^{(2)}\left(x^2\right) x^{2k+1} e^{-x^2} dx = 2^{k+\frac{1}{2}} \int_{0}^{\infty} P_n^{(2)}(t) t^{k+\frac{1}{2}} e^{-t} dt = 0. \qquad \square$$

The polynomials $P^{(1)}(x)$ and $P^{(2)}(x)$ are specific examples of a more general class of polynomials.

Definition 1.4.12 (cf. [42], p.204). *The Laguerre polynomials* $\{L_n^{\alpha}(x)\}_{n=0}^{\infty}$ *are defined by the Rodrigues' formula*

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

Because the polynomials considered here have real coefficients it is stipulated that $\alpha \in \mathbb{R}$. *It is also stipulated that* $\alpha > -1$.

As with the Hermite polynomials the Rodrigues' formula can be used to derive the orthogonality relation for the Laguerre polynomials.

Theorem 1.4.13 (cf. [42], p.205)**.** *The Laguerre polynomials satisfy the orthogonality relation*

$$\int_0^\infty L_n^\alpha(x) L_m^\alpha(x) x^\alpha e^{-x} dx = h_n \delta_{nm}, \tag{1.4.12}$$

$h_n > 0$.

Proof. Using the Rodrigues' formula the integral (1.4.12) reduces to

$$\frac{1}{n!} \int_0^\infty \frac{d^n}{dx^n}(e^{-x} x^{n+\alpha}) L_m^\alpha(x) dx. \tag{1.4.13}$$

$\lim_{x\to\infty} e^{-x} = 0$ and $\lim_{x\to 0} x^{\alpha+n-k} = 0$ so

$$\left[\frac{d^k}{dx^k}(e^{-x} x^{n+\alpha})\right]_0^\infty = 0$$

for any natural number $k < n$. For $m < n$ this fact can be used to elimanate $L_m^\alpha(x)$ from (1.4.13) with integration by parts. For $m = n$, e^{-x} is positive and continuous over $[0, \infty)$ and $(L_n^\alpha(x))^2 \geq 0$ and not identically zero, so the integral is positive. $\square$

The discussion of the Laguerre polynomials makes it clear that the polynomials $\{P^{(1)}(x)\}_{n=0}^\infty$ and $\{P^{(2)}(x)\}_{n=0}^\infty$ which generate the Hermite polynomials are, up to normalisation, the Laguerre polynomials $\{L_n^{-\frac{1}{2}}(x)\}_{n=0}^\infty$ and $\{L_n^{\frac{1}{2}}(x)\}_{n=0}^\infty$ respectively.

Chapter 2

The moment problem

The moment problem requires the generalisation of the Riemann integral to the Riemann-Stieltjes integral. An important concept for Riemann-Stieltjes integration is the function of bounded variation which generates the integral. In particular the functions of bounded variation which are also distribution functions are important here. The first part of the chapter examines the rudiments of Riemann-Stieltjes integration. Subsequently results are developed for distribution functions. With the foundations laid the concept of a moment problem is introduced. It is shown how the Hamburger moment problem gives rise to a generalisation of the classical case of orthogonal polynomials and a proof of Theorem 1.1.8 is given. Several examples of moment problems are given and necessary and sufficient conditions are obtained for the existence of solutions. The most important result of the chapter is Favard's theorem which establishes the connection between the moment problem and orthogonal polynomials.

2.1 Distribution functions

Let $\{x_k\}_{k=0}^{\infty}$ be a sequence of arbitrary real numbers and $\{a_k\}_{k=0}^{\infty}$ be a sequence of positive real numbers. Assume that

$$\sum_{k=0}^{\infty} |P_n(x_k)a_k| < \infty$$

for an arbitrary polynomial $P_n(x)$ of degree n. Define a binary relation $\langle \cdot, \cdot \rangle$ on the space of real polynomials of a real variable by

$$\langle P_n(x), P_m(x) \rangle = \sum_{k=0}^{\infty} P_n(x_k)P_m(x_k)a_k.$$

This binary relation is well defined because of the Cauchy-Schwarz inequality. By definition

$$\langle P_n(x), P_m(x) \rangle = \langle P_m(x), P_n(x) \rangle,$$

and because of absolute convergence

$$\langle \alpha_1 P_l(x) + \alpha_2 P_m(x), P_n(x) \rangle = \alpha_1 \langle P_l(x), P_n(x) \rangle + \alpha_2 \langle P_m(x), P_n(x) \rangle.$$

Also, $\langle P_n(x), P_n(x) \rangle > 0$, because all of the terms in the sum are positive.

As a result of this $\langle \cdot, \cdot \rangle$ is an inner product on the space of polynomials (cf. [2], p.2), and the Gram-Schmidt process discussed in Chapter 1 can be applied to obtain a set of polynomials satisfying the orthogonality relation,

$$\sum_{k=0}^{\infty} P_n(x_k)P_m(x_k)a_k = h_n \delta_{nm}, \;\; h_n > 0.$$

An example of this phenomenon is the Charlier polynomials (cf. [15], p.4), which satisfy

$$\sum_{k=0}^{\infty} P_n(k)P_m(k)\frac{a^k}{k!} = \frac{e^a a^n}{n!}\delta_{nm},$$

for $a > 0$ a real constant.

This is a valid case of orthogonality but falls outside the scope of Riemann (or Lebesgue) integration in establishing an orthogonality condition.

To investigate cases such as this, the concept of a distribution function will be discussed.

An important concept in the characterisation of distribution functions is the variation of a function.

Definition 2.1.1 (cf. [44], p.10)**.** *Let $f(x)$ be an arbitrary real valued function defined on the interval (a, b). Let Π be the set of all partitions*

$$\pi = \{(x_i, x_{i+1}) | i = 1, 2, 3, \ldots, n, \ a = x_1 < x_2 < \ldots < x_n = b\}$$

of (a, b). The total variation of $f(x)$ on (a, b) is denoted by $T(a, b)$ and is defined to be

$$T(a, b) = \sup_{\pi \in \Pi} \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|.$$

Definition 2.1.2 (cf. [44], p.10)**.** *A function $f(x)$ is said to have bounded variation on an interval (a, b) if $T(a, b) < \infty$.*

Definition 2.1.3 (cf. [44], p.10)**.** *The indefinite total variation of a function $f(x)$ on an interval (a, b) denoted $T(x)$ is defined as $T(a, x)$.*

Non-decreasing functions can be used to totally characterise the class of functions of bounded variation.

Lemma 2.1.4 (cf.[44], p.10)**.** *If $T(x)$ is the indefinite total variation of a funtion of bounded variation $f(x)$ for the interval (a, b), $x,\ y \in (a, b)$ and $y > x$ then $T(y) = T(x) + T(x, y)$.*

Proof. Inserting a point into a partition of (a, y) can not decrease the sums

$$\sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|.$$

If the point is x then partitions of (a, x) and (x, y) can be considered seperately. This means that $T(a, y) = T(a, x) + T(x, y)$ which is what was required. $\square$

Theorem 2.1.5 (cf. [44], p.10). *Every function of bounded variation is the difference of two bounded non-decreasing functions.*

Proof. Let $f(x)$ be a function of bounded variation on the interval (a, b). Then $f(x)$ is bounded because it has bounded variation. $T(x)$ is bounded below by zero and above by $T(a, b)$. It follows that $T(x) - f(x)$ is also bounded.

$$f(x) = T(x) - \{T(x) - f(x)\},$$

is the required decomposition. From Lemma 2.1.4 $T(x)$ is non-decreasing. It remains to show that for $y \in (a, b)$ and $y > x$, $T(x) - f(x) \leq T(y) - f(y)$. After rearranging this is

$$f(y) - f(x) \leq T(y) - T(x).$$

$T(y) - T(x) = T(x, y)$ and by definition $|f(y) - f(x)| \leq T(x, y)$. □

Definition 2.1.6 (cf. [52], p.239). *Let $f(x)$ and $\alpha(x)$ be real valued functions defined on the interval (a, b). Let $a = x_1 < x_2 < \ldots < x_n = b$ be a partition of (a, b) and let $x_i \leq v_i \leq x_{i+1}$, $i = 1, 2, \ldots, n-1$. The Riemann-Stieltjes integral of $f(x)$ with respect to $\alpha(x)$ is defined by,*

$$\int_a^b f(x) d\alpha(x) = \lim_{\delta \to 0} \sum_{k=1}^{n-1} f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\}, \tag{2.1.1}$$

where $\delta = \max(x_{i+1} - x_i)$, $i = 1, 2, \ldots, n-1$. The Riemann-Stieltjes integral of a function exists if the same limit is obtained irrespective of how the partitions are taken.

The importance of functions of bounded variation in this theory is summed up by the following theorem.

Theorem 2.1.7 (cf. [52], p.241, [46], p.66). *If $f(x)$ is continuous and $\alpha(x)$ has bounded variation in the interval $[a, b]$, then the Riemann-Stieltjes integral of $f(x)$ with respect to $\alpha(x)$ exists.*

Proof. $f(x)$ is continuous on the closed interval $[a, b]$ so it is uniformly continuous. That is, for every $\epsilon > 0$ there is a δ such that if $\pi = \{(x_i, x_{i+1}) | a = x_1 < x_2 < \ldots < x_n = b\}$ is a partition of $[a, b]$ where the intervals $[x_i, x_{i+1}]$, $i = 1, 2, \ldots, n-1$, have maximum length δ, then $|f(x_{i+1}) - f(x_i)| < \epsilon$, $i = 1, 2, \ldots, n-1$. Such a partition will be called an ϵ-partition. Let $S(x_k, v_k)$ denote the sum

$$\sum_{k=1}^{n-1} f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\}$$

generated by some ϵ-partition π_k. Let π_{kl} be a new partition obtained by adding points to π_k such that $x_{k,l+1}$ denotes the l^{th} point of π_{kl} added to π_k between x_k and x_{k+1}, $x_{k,1} = x_k$. Let $m(k)$ denote the number of points added to the interval $[x_k, x_{k+1}]$ and set $x_{k,m(k)+2} = x_{k+1}$. The index l runs from 1 to $m(k)+2$. Let v_{kl} be chosen in the interval $[x_{k,l}, x_{k,l+1}]$. If $f(v_{kl}) - f(v_k) = \epsilon_{kl}$, it follows that $|\epsilon_{kl}| < \epsilon$. $S(x_{kl}, v_{kl})$ will be the sum generated by π_{kl}. Now

$$\begin{aligned}
&|S(x_k, v_k) - S(x_{kl}, v_{kl})| \\
&= \left| \sum_{k=1}^{n-1} \left[\sum_{l=1}^{m(k)+1} f(v_{kl})\{\alpha(x_{k,l+1}) - \alpha(x_{k,l})\} - f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\} \right] \right| \\
&= \left| \sum_{k=1}^{n-1} \left[\sum_{l=1}^{m(k)+1} \epsilon_{kl}\{\alpha(x_{k,l+1}) - \alpha(x_{k,l})\} \right] \right| \\
&< \epsilon \sum_{k=1}^{n-1} \sum_{l=1}^{m(k)+1} |\alpha(x_{k,l+1}) - \alpha(x_{k,l})| \le \epsilon T(a, b),
\end{aligned}$$

where $T(a, b)$ is the total variation of $f(x)$ on (a, b). Let $S(x, v)$ and $S(x', v')$ be sums generated by two $\frac{\epsilon}{2}$-partitions π and π'. Let $S(x'', v'')$ be generated by the partition π'' obtained by adding the points of π to the points of π'. Then using the above and the triangle inequality

$$|S(x, v) - S(x'', v'')| < \frac{\epsilon}{2} T(a, b) \text{ and } |S(x', v') - S(x'', v'')| < \frac{\epsilon}{2} T(a, b),$$

so

$$|S(x, v) - S(x', v')| < \epsilon T(a, b). \tag{2.1.2}$$

Let a decreasing sequence of numbers $\epsilon_p > 0$, $\lim_{p\to\infty} \epsilon_p = 0$, be given. For each p let $S(x^{(p)}, v^{(p)})$ be a sum generated by an ϵ_p-partition. For any $\eta > 0$, (2.1.2) gives an N such that

$$|S(x^{(p+n)}, v^{(p+n)}) - S(x^{(p)}, v^{(p)})| < \eta, \text{ for } p > N, \; n = 1, 2 \ldots.$$

This is a Cauchy sequence of real numbers so

$$\lim_{p\to\infty} S(x^{(p)}, v^{(p)}) = I < \infty$$

exists. If $S(x, v)$ is a sum corresponding to an arbitrary $\frac{\epsilon}{2}$-partition, then

$$\begin{aligned} |S(x, v) - I| &\leq |S(x, v) - S(x^{(p)}, v^{(p)})| + |S(x^{(p)}, v^{(p)}) - I| \\ &\leq \epsilon T(a, b) + \eta, \end{aligned}$$

so convergence doesn't depend on how partitions are taken. The Riemann-Stieltjes integral exists and is equal to I. $\square$

The Riemann-Stieltjes integral satisfies several useful properties some of which are analogous to properties of the Riemann integral. For instance, the Riemann-Stieltjes integral satisfies a formula for integration by parts.

Lemma 2.1.8 (cf. [52], p.240)**.** *If the Riemann-Stieltjes integral of $f(x)$ with respect to $\alpha(x)$ on the interval (a, b) exists then the Riemann-Stieltjes integral of $\alpha(x)$ with respect to $f(x)$ exists and,*

$$\int_a^b f(x) d\alpha(x) = f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b \alpha(x) df(x).$$

Proof. The partition $\pi_x = \{(x_i, x_{i+1}) | a = x_1 < x_2 < \ldots < x_n = b\}$ generates a dual partition choosing $v_1 = a$ and $v_{n-1} = b$ (which is permitted in Definition 2.1.6), $\pi_v = \{(v_i, v_{i+1}) | a = v_1 < v_2 < \ldots < v_{n-1} = b\}$. It can be seen that the coefficient of $\alpha(x_k)$ in (2.1.1) is $(-f(v_k) + f(v_{k-1}))$, except for the case $\alpha(x_1)$ which has coefficient $-f(v_1)$ and the case $\alpha(x_n)$ which has coefficient

$f(v_{n-1})$. Using these facts, the right hand side of (2.1.1) considered as a finite sum can be rewritten as

$$\alpha(b)f(b) - \alpha(a)f(a) + \sum_{k=2}^{n-1} \alpha(x_k)(-f(v_k) + f(v_{k-1}))$$

$$= \alpha(b)f(b) - \alpha(a)f(a) - \sum_{k=2}^{n-1} \alpha(x_k)(f(v_k) - f(v_{k-1})).$$

Subsequent partitions $\pi_v = \{(v_i, v_{i+1}) | a = v_1 < v_2 < \ldots < v_{m-1} = b\}$ can be chosen arbitrarily with new x_k chosen satisfying $v_{k-1} \leq x_k \leq v_k$, and the reduction of the length of the intervals (v_{k-1}, v_k) toward zero coincides with the reduction of the length of the intervals (x_k, x_{k+1}) toward zero. □

The following lemmas illustrate the role that the Stieltjes integral plays as a linear functional.

Lemma 2.1.9 (cf. [52], p.241)**.** *The Stieltjes integral is a linear functional on continuous functions,*

$$\int_a^b [f_1(x) + f_2(x)] d\alpha(x) = \int_a^b f_1(x) d\alpha(x) + \int_a^b f_2(x) d\alpha(x)$$

and

$$\int_a^b c f(x) d\alpha(x) = c \int_a^b f(x) d\alpha(x).$$

Proof. In the right hand side of (2.1.1) substitute $f_1(v_k) + f_2(v_k)$ for $f(v_k)$ so that it becomes

$$\lim_{\delta \to 0} \sum_{k=1}^{n-1} (f_1(v_k) + f_2(v_k))\{\alpha(x_{k+1}) - \alpha(x_k)\}$$

$$= \lim_{\delta \to 0} \sum_{k=1}^{n-1} f_1(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\} + \lim_{\delta \to 0} \sum_{k=1}^{n-1} f_2(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\}.$$

Similarly

$$\lim_{\delta \to 0} \sum_{k=1}^{n-1} c f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\}$$

$$= c \lim_{\delta \to 0} \sum_{k=1}^{n-1} f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\}.$$ □

In fact any continuous linear functional on the Banach space $C[a,b]$ of continuous functions on an interval $[a,b]$ is given by the Stieltjes integral of the function with respect to a function of bounded variation (cf. [44], p.110).

Lemma 2.1.10 (cf.[52], p.241)**.** *The Stieltjes integral is a linear functional on the function of bounded variation with respect to which the integration is carried out, i.e.*

$$\int_a^b f(x)d[\alpha_1(x)+\alpha_2(x)] = \int_a^b f(x)d\alpha_1(x) + \int_a^b f(x)d\alpha_2(x)$$

and

$$\int_a^b f(x)d[c\alpha(x)] = c\int_a^b f(x)d\alpha(x)$$

Proof. Again referring to (2.1.1),

$$\begin{aligned}
&\lim_{\delta\to 0}\sum_{k=1}^{n-1} f(v_k)\{(\alpha_1(x_{k+1})+\alpha_2(x_{k+1}))-(\alpha_1(x_k)+\alpha_2(x_k))\} \\
&= \lim_{\delta\to 0}\left\{\sum_{k=1}^{n-1} f(v_k)\{\alpha_1(x_{k+1})-\alpha_1(x_k)\} + \sum_{k=1}^{n-1} f(v_k)\{\alpha_2(x_{k+1})-\alpha_2(x_k)\}\right\} \\
&= \lim_{\delta\to 0}\sum_{k=1}^{n-1} f(v_k)\{\alpha_1(x_{k+1})-\alpha_1(x_k)\} + \lim_{\delta\to 0}\sum_{k=1}^{n-1} f(v_k)\{\alpha_2(x_{k+1})-\alpha_2(x_k)\}.
\end{aligned}$$

Finally

$$\lim_{\delta\to 0}\sum_{k=1}^{n-1} f(v_k)\{c\alpha(x_{k+1})-c\alpha(x_k)\} = c\lim_{\delta\to 0}\sum_{k=1}^{n-1} f(v_k)\{\alpha(x_{k+1})-\alpha(x_k)\}. \quad \square$$

In what follows the case of indefinite Riemann-Stieltjes integration will often be used.

Definition 2.1.11 (cf. [52], p.243)**.** *The indefinite Riemann-Stieltjes integral is defined in analogy with the case for the Riemann integral by,*

$$\int_{-\infty}^{\infty} f(x)d\alpha(x) = \lim_{a\to-\infty}\lim_{b\to\infty}\int_a^b f(x)d\alpha(x).$$

Certain Riemann-Stieltjes integrals can be reduced to Riemann integrals.

Theorem 2.1.12 (cf. [52], p.241)**.** *If $\alpha(x)$ has a continuous derivative $\alpha'(x)$ on (a,b) and $f(x)$ is Riemann-Stieltjes integrable with respect to $\alpha(x)$ on (a,b) then,*

$$\int_a^b f(x)d\alpha(x) = \int_a^b f(x)\alpha'(x)dx.$$

Proof. By the mean value property in each interval $[x_k, x_{k+1}]$ in (2.1.1) there is a point v_k such that

$$\alpha(x_{k+1}) - \alpha(x_k) = \alpha'(v_k)(x_{k+1} - x_k).$$

Choosing all v_k in (2.1.1) according to this rule gives the result. □

A bounded non-decreasing function is clearly a function of bounded variation. In fact the variation of such a function $\alpha(x)$ on an interval (a,b) is $\alpha(b) - \alpha(a)$.

Definition 2.1.13 (cf. [15], p.51)**.** *A bounded non-decreasing function $\alpha(x)$ satisfying*

$$\int_{-\infty}^{\infty} x^n d\alpha(x) < \infty$$

for $n = 0, 1, 2, \ldots$ is called a distribution function.

Returning to the example of the Charlier polynomials, if $\alpha(x)$ is taken as a step function with jumps of size $\dfrac{a^k}{k!}$ at the non-negative integers k then from (2.1.1),

$$\int_0^{\infty} P_n(x)P_m(x)d\alpha(x) = \lim_{\delta\to 0}\sum_{i=0}^{j-1} P_n(v_i)P_m(v_i)\{\alpha(x_{i+1}) - \alpha(x_i)\},$$

where the partitioning is undertaken in the same way as in (2.1.1). If (x_i, x_{i+1}) is an interval where this step function is constant then $\alpha(x_{i+1}) - \alpha(x_i) = 0$ and this interval contributes nothing to the sum. On the other hand if (x_i, x_{i+1}) is an interval where only one jump occurs (with fine enough partitions at most one can occur) then $\alpha(x_{i+1}) - \alpha(x_i) = \dfrac{a^k}{k!}$. As the partitions are taken more finely any interval containing a jump becomes smaller and, in the limit, $v_i \to k$

where k is some positive integer where a jump occurs. Because $P_n(x)P_m(x)$ is a continuous function the result of this limiting process is

$$\int_0^\infty P_n(x)P_m(x)d\alpha(x) = \sum_{k=0}^{\infty} P_n(k)P_m(k)\frac{a^k}{k!},$$

so that the step function generates a Stieltjes integral which corresponds to the functional introduced at the beginning of the section.

In a similar way the classical polynomials have continuous Riemann integrable weight functions $w(x) > 0$. The indefinite Riemann integral of one such function over the interval of orthogonality (a, b) gives a function $\alpha(x)$ which is non-decreasing (because $w(x) > 0$) and bounded (because $w(x)$ is Riemann integrable on (a, b)) and has a continuous derivative everywhere in (a, b) (by the fundamental theorem of calculus). Theorem 2.1.12 then gives that for any continuous function $f(x)$ defined on (a, b)

$$\int_a^b f(x)d\alpha(x) = \int_a^b f(x)\alpha'(x)dx = \int_a^b f(x)w(x)dx,$$

so that all classical cases of orthogonality can be represented by Stieltjes integrals.

2.2 Uniqueness of distributions

A distribution function was defined to be bounded and non-decreasing. Such a function can be discontinuous but its discontinuities satisfy restrictive conditions. For instance a distribution function can only have countably many discontinuities otherwise the sum of the jumps at these discontinuities would have to be infinite (and the function would be unbounded).

Lemma 2.2.1 (cf. [31], p.19). *If $\alpha(x)$ is a distribution function on (a, b) and $c \in (a, b)$ then*

$$\alpha(c-) = \lim_{x \to c^-} \alpha(x) = \sup_{x<c} \alpha(x)$$

and

$$\alpha(c+) = \lim_{x \to c^+} \alpha(x) = \inf_{x > c} \alpha(x),$$

exist. For $c = a$ only the second relation holds, and for $c = b$ only the first.

Proof. Because $\alpha(x)$ is a non-decreasing and bounded function the set

$$\{\alpha(x) | a \leq x < c\}$$

is bounded above and non-empty so $\sup_{x<c} \alpha(x)$ exists. Let $\epsilon > 0$, then as a property of the supremum there is an x_ϵ such that
$\sup_{x<c} \alpha(x) - \alpha(x_\epsilon) < \epsilon$, and because $\alpha(x)$ is non-decreasing, this difference can only get smaller for other values of x in $(c - x_\epsilon, c)$, giving

$$\sup_{x<c} \alpha(x) - \epsilon < \alpha(x) \leq \sup_{x<c} \alpha(x) < \sup_{x<c} \alpha(x) + \epsilon,$$

for $x \in (c - x_\epsilon, c)$. Letting ϵ go to zero gives the required limit. Because of the symmetry between the supremum and infimum, the same approach can be used for the cases $\alpha(c+)$, $\alpha(a+)$ and $\alpha(b-)$. $\square$

Lemma 2.2.2 (cf. [31], p.20)**.** *Let $\alpha(x)$ be a function which is bounded and non-decreasing on (a, b). Then $\alpha(x)$ has at most countably many discontinuities.*

Proof. Let a partition

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b$$

of (a, b) be given. Let $y_k \in (x_k, x_{k+1})$. Then $\alpha(x_k+) \leq \alpha(y_k)$ and $\alpha(y_{k-1}) \leq \alpha(x_k-)$ so

$$\alpha(x_k+) - \alpha(x_k-) \leq \alpha(y_k) - \alpha(y_{k-1}).$$

Because $\alpha(y_0) \geq \alpha(a+)$ and $\alpha(y_n) \leq \alpha(b-)$

$$\alpha(a+) - \alpha(a) \leq \alpha(y_0) - \alpha(a)$$

and

$$\alpha(b) - \alpha(b-) \leq \alpha(b) - \alpha(y_n).$$

Adding the inequalities over the index k the terms $\alpha(y_k)$ are telescoping and cancel leaving

$$\alpha(a+) - \alpha(a) + \sum_{k=1}^{n}[\alpha(x_k+) - \alpha(x_k-)] + \alpha(b) - \alpha(b-) \leq \alpha(b) - \alpha(a).$$

Let $A_n = \{x \in (a,b) | \alpha(x+) - \alpha(x-) > \frac{1}{n}\}$. If $\{x_i\}_{i=1}^{k} \subset A_n$ then

$$\alpha(b) - \alpha(a) \geq \sum_{i=1}^{k}[\alpha(x_i+) - \alpha(x_i-)] > \frac{k}{n},$$

so k has to be finite because $\alpha(b) - \alpha(a) < \infty$ and the union $\bigcup_{n=1}^{\infty} A_n$ is therefore countable. □

Theorem 2.2.3 (cf. [40], p.2)**.** *The complement of any countable set in an interval (a,b) is dense in that interval.*

Proof. Let $A = \bigcup_{n=1}^{\infty}\{a_n\}$ be a countable subset of (a,b), and I be an arbitrary subinterval of (a,b). Let I_1 be a closed subinterval of I such that $a_1 \notin I_1$. For $i > 1$ let I_i be a closed subinterval of I_{i-1} such that $a_i \notin I_i$. Then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$, so the complement of A has non-empty intersection with I. □

In (2.1.1) the Riemann-Stieltjes integral of a continuous function $f(x)$ with respect to a distribution function $\alpha(x)$ was defined as

$$\int_a^b f(x)d\alpha(x) = \lim_{\delta \to 0} \sum_{k=1}^{n-1} f(v_k)\{\alpha(x_{k+1}) - \alpha(x_k)\},$$

where $\delta = \max_{x_k}\{x_{k+1} - x_k\}$. If $\alpha(x)$ has a discontinuity at a point x_d then as the partitions get finer the contribution made by the point x_d to the integral is given by $f(x_d)d$, where d is the jump at x_d. Because only the jump matters and not the specific value of $\alpha(x_d)$ a distribution function can take on any value c in the interval

$$\sup_{x<x_d} \alpha(x) < c < \inf_{x>x_d} \alpha(x),$$

and still generate the same value for the integral of $f(x)$. It seems unreasonable that the same exception could hold for points of continuity. These considerations are dealt with in the following theorem.

Theorem 2.2.4 (cf. [52], p.243)**.** *For the relation*

$$\int_a^b f(x)d\alpha(x) = 0$$

to hold for all continuous functions $f(x)$ it is necessary and sufficient that $\alpha(x) = \alpha(a)$ for $x = b$ and for all x in (a, b) except countably many points where $\alpha(x)$ is discontinuous.

Proof. If the integral is zero for all continuous functions $f(x)$ then in particular it is zero for $f(x) = 1$. In this case the sums that make up the integral are telescoping for all partitions and the integral is equal to $\alpha(b) - \alpha(a)$, which gives $\alpha(b) = \alpha(a)$. Now let $f(x)$ be the continuous function

$$f(x) = \begin{cases} x & \text{if} \quad a \leq x \leq v \\ v & \text{if} \quad \quad x > v \end{cases}$$

then

$$0 = \int_a^v x d\alpha(x) + v \int_v^b d\alpha(x).$$

Using integration by parts gives

$$v\alpha(v) - a\alpha(a) - \int_a^v \alpha(x)dx + v\alpha(b) - v\alpha(v),$$

and since $\alpha(b) = \alpha(a)$ this is

$$(v - a)\alpha(a) - \int_a^v \alpha(x)dx.$$

If v is a point of continuity of $\alpha(x)$ then take the derivative with respect to v, which gives

$$\alpha(a) - \alpha(v) = 0.$$

Suppose that $\alpha(a) = \alpha(b) = \alpha(v)$ for v any point of continuity of $\alpha(x)$ in (a, b). The set of discontinuities is countable, so the set of points of continuity is dense. As a result the endpoints of intervals in the partitions can be chosen to miss discontinuities and so that the maximum length of intervals in the partitions go to zero. The limit exists and is unique because $\alpha(x)$ has bounded variation. □

Because of this theorem, if $\alpha_1(x)$ and $\alpha_2(x)$ are distributions whose difference is constant at a and b and at points of continuity then by Lemma 2.1.9,

$$\int_a^b f(x)d\alpha_1(x) - \int_a^b f(x)d\alpha_2(x) = \int_a^b f(x)d(\alpha_1(x) - \alpha_2(x)) = 0.$$

2.3 Measure and decomposition

It was seen that some distributions have continuous derivatives, and in this case the Stieltjes integral reduces to

$$\int_{-\infty}^{\infty} f(x)d\alpha(x) = \int_{-\infty}^{\infty} f(x)\alpha'(x)dx.$$

Some distributions are jump functions and the Stieltjes integral reduces to

$$\sum_{k=0}^{\infty} f(x_k)a_k.$$

These two cases can be generalised and a further case occurs that was not even treated previously, the case of a singular distribution.

Every distribution function (in fact every function of bounded variation) generates a measure via, for instance, the Daniell scheme (cf.[46]). An advantage to using measure is that if L^p_α denotes the collection of functions $f(x)$

such that

$$\int_{-\infty}^{\infty} |f(x)|^p d\alpha(x) < \infty,$$

where integration is carried out with respect to the measure generated by $\alpha(x)$, then $L^p_\alpha(x)$ is a Banach space with respect to the norm (cf. [1], p.34)

$$\|f(x)\| = \left\{\int_{-\infty}^{\infty} |f(x)|^p d\alpha(x)\right\}^{\frac{1}{p}}.$$

The question of when polynomials are dense in this Banach space has been comprehensively dealt with in [8]. Questions relating to Riemann-Stieltjes integrability and Lebesgue-Stieltjes integrability are dealt with in [29].

The discussion will still focus on the distribution functions that generate the associated Lebesgue-Stieltjes measures.

Characterisation of the different cases of distributions is closely related to concepts from Lebesgue measure and integration.

Definition 2.3.1 (cf. [44], p.5)**.** *A subset of the real line is said to have Lebesgue measure zero if it can be covered by countably many intervals, of any kind, of arbitrarily small total length.*

The following lemma will be used to prove Theorem 2.3.4.

Lemma 2.3.2 (cf. [44], p.6)**.** *Let $g(x)$ be a function defined in the interval $[a, b]$ such that $g(x+)$, $g(x-)$ and $g(x)$ exist and are finite for the interval (a, b). For a, $g(a+)$ must exist and be finite and for b, $g(b-)$ must exist and be finite. Let $G(x) = \max\{g(x-), g(x), g(x+)\}$ for $x \in (a, b)$, $G(a) = g(a+)$, $G(b) = g(b-)$. Let E be the set of points $x \in (a, b)$ such that there is a $\zeta > x$ and $g(\zeta) > G(x)$. Then E is either empty or it is a finite or countable union of open disjoint intervals (a_k, b_k), satisfying $g(a_k+) \leq G(b_k)$.*

Proof. Let x_0 and ζ be points satisfying $x_0 < \zeta$, $G(x_0) < g(\zeta)$. Because $G(x) \geq g(x+)$ there is an interval $[x_0, x_0 + \epsilon)$ such that $G(x) < g(\zeta)$ for $x \in [x_0, x_0 + \epsilon)$. Similarly there is an interval $(x_0 - \epsilon, x_0]$ where this holds.

Taking the union of the largest possible of these intervals for each such x_0 gives the decomposition of E into disjoint open intervals. Let (a_k, b_k) be an interval in the decomposition and let $x \in (a_k, b_k)$. Let x_1 be the largest number in $(x, b_k]$ such that $G(x) \leq G(x_1)$. If $x_1 < b_k$, then the ζ_1 corresponding to x_1 would be greater than b_k. Because b_k is not in E, $G(b_k) \geq g(\zeta_1)$, but $G(x_1) > G(b_k)$ and $G(x_1) < g(\zeta_1)$, so $G(x_1) < G(x_1)$, a contradiction. It follows that $x_1 = b_k$. Letting $x \to a_k$ gives the result. $\square$

Definition 2.3.3 (cf. [44], p.7)**.** *Let $h > 0$. The lower and upper right derived numbers λ_r and Λ_r are given by*

$$\Lambda_r = \limsup_{h \to 0} \frac{f(x+h) - f(x)}{h}, \;\; \lambda_r = \liminf_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

The lower and upper left derived numbers Λ_l and λ_l are defined analogously, with $f(x+h)$ replaced by $f(x-h)$.

A function $f(x)$ has a finite derivative at a point x if all of its derived numbers are finite and equal. If a set has Lebesgue measure zero it will be called a null set. A property that holds everywhere except on a null set will be said to hold almost everywhere.

Theorem 2.3.4 (cf. [44], p.11)**.** *Every function $h(x)$, of bounded variation, has a finite derivative $h'(x)$ almost everywhere.*

Proof. The result is first proved for a bounded non-decreasing function $f(x)$. It is sufficient to prove that

$$\Lambda_r < \infty \text{ and} \tag{2.3.1}$$

$$\Lambda_r \leq \lambda_l, \tag{2.3.2}$$

almost everywhere. Applying (2.3.2) to the function $-f(-x)$ gives $\Lambda_l \leq \lambda_r$ so that

$$\Lambda_r \leq \lambda_l \leq \Lambda_l \leq \lambda_r \leq \Lambda_r < \infty.$$

Denote by E_∞ the set where $\Lambda_r = \infty$ and $f(x)$ is continuous. This set is contained in the set E_C of points where $\Lambda_r > C$ and $f(x)$ is continuous. Let $x \in E_C$. Then there is a point ζ such that

$$\frac{f(\zeta) - f(x)}{\zeta - x} > C.$$

Setting $g(x) = f(x) - Cx$ and $G(x) = \max\{g(x-), g(x), g(x+)\}$ gives $G(x) < g(\zeta)$, by continuity at x. By Lemma 2.3.2 E_C is covered by countably many disjoint intervals (a_k, b_k) and $g(a_k+) \leq G(b_k)$ or

$$f(a_k+) - Ca_k \leq f(b_k+) - Cb_k, \text{ or } f(a_k+) - Ca_k \leq f(b-) - Cb_k, \text{ if } b_k = b,$$

which gives

$$C(b_k - a_k) \leq f(b_k+) - f(a_k+), \text{ or } C(b_k - a_k) \leq f(b-) - f(a_k+), \text{ if } b_k = b.$$

Summing the above yields

$$C \sum_k (b_k - a_k) \leq f(b) - f(a),$$

because $f(x)$ is non-decreasing. Because C can be made arbitrarily large, the total length of the covering intervals can be made to go zero. As a result, E_∞ has Lebesgue measure zero. Now, let $0 < c < C$ be two given numbers. Let E_c be the collection of points where $\lambda_l < c$ and $f(x)$ is continuous. If x is a point of continuity of $f(x)$ where $\lambda_l < c$ then $-x$ is a point of continuity of $f(-x)$ where $\lambda_r > c$, so for every point of E_c there is an h such that

$$\frac{f(-x+h) - f(-x)}{h} > c.$$

If $x - h = \zeta$ the condition reads

$$\frac{f(-\zeta) - f(-x)}{x - \zeta} > c. \tag{2.3.3}$$

Let $g_c(x) = f(-x) + cx$ and $G_c(x) = \max\{g_c(x-), g_c(x), g_c(x+)\}$. By Lemma 2.3.2, there are countably many disjoint intervals $(-b_k, -a_k)$ where (2.3.3) can hold and $g(-b_k+) \leq G_c(-a_k)$ so

$$f(-b_k+) + cb_k \leq f(-a_k-) + ca_k,$$

which gives

$$c(-a_k + b_k) \leq f(-a_k-) - f(-b_k+).$$

Reflecting these intervals around the origin gives

$$f(b_k-) - f(a_k+) \leq c(b_k - a_k).$$

The total length of the intervals (a_k, b_k) will be denoted by Σ_1. Let $g_C(x) = f(x) - Cx$ and $G_C(x) = \max\{g_C(x-), g_C(x), g_C(x+)\}$. For each interval (a_k, b_k) generate the set of points in (a_k, b_k) where $G_C(x)$ satisfies the conditions of Lemma 2.3.2. This set is covered by a collection of disjoint intervals (a_{kl}, b_{kl}) for each k. Σ_2 will denote the total length of the intervals (a_{kl}, b_{kl}). The following identity then holds

$$C\Sigma_2 \leq c\Sigma_1.$$

To verify this it suffices to recognise that if $b_{kl} = b_k$ for some k then $G_C(b_{kl}) = g_C(b_k-)$. Alternately applying Lemma 2.3.2 to the remaining intervals for the functions $g_c(x)$ and $g_C(x)$ generates a sequence of families of intervals whose lengths Σ_n satisfy

$$\Sigma_{2n} \leq \frac{c}{C}\Sigma_{2n-1} \leq \left(\frac{c}{C}\right)^n \Sigma_1 \to 0, \text{ as } n \to \infty.$$

The set E_{cC} where $f(x)$ is continuous and $\Lambda_r > C$ and $\lambda_l < c$ at the same time is contained in all of the intervals generated above so it has Lebesgue measure zero. The union of the countable family of sets E_{cC}, for c and C rational numbers, contains all points where $f(x)$ is continuous and $\Lambda_r > \lambda_l$. This is a countable union of sets of measure zero, so it also has measure zero. The points of discontinuity of $f(x)$ have not been considered, but according to Lemma 2.2.2 there are only countably many such points so the collection of them has Lebesgue measure zero. It follows that a bounded non-decreasing function $f(x)$ has a derivative almost everywhere. By Theorem 2.1.5, every function of bounded variation is the difference of two bounded non-decreasing functions. Using the linearity of the derivative on this decomposition gives the result. □

Absolutely continuous distributions constitute a very important type of distribution function.

Definition 2.3.5 (cf. [44], p.51). *Let $\{(\alpha_k, \beta_k)\}_{k=0}^{n}$ be a countable collection of disjoint real intervals where n is finite or infinite. A function $f(x)$ of bounded variation is said to be absolutely continuous if for any $\epsilon > 0$, there is a $\delta > 0$, such that*

$$\sum_{k=0}^{n}(\alpha_k - \beta_k) < \delta$$

implies that

$$\sum_{k=0}^{n}|f(\alpha_k) - f(\beta_k)| < \epsilon.$$

Choosing the trivial covering of a single interval and letting its length go to zero shows that absolutely continuous functions are necessarily continuous (cf.[31], p.155).

In this context absolutely continuous functions are a powerful generalisation of the weight functions that occur in the case of classical polynomials.

Theorem 2.3.6 (cf. [44], p.53). *A function $f(x)$, of bounded variation, is absolutely continuous if and only if it is the indefinite Lebesgue integral of its almost everywhere derivative $f'(x)$.*

The proof of this result can be found in [44] on page 50. Because of this result, if a distribution $\alpha(x)$ is absolutely continuous, then

$$\int f(x)d\alpha(x) = \int f(x)\alpha'(x)dx.$$

Results on absolutely continuous functions can be found in [51].

Definition 2.3.7 (cf. [44], pp.13, 14). *Let $\{u_n\}_{n=0}^{\infty}$ and $\{v_n\}_{n=0}^{\infty}$ be absolutely convergent series and $\{x_n\}_{n=0}^{\infty}$ be a countable sequence of points in the interval*

(a, b). *The function*

$$f(x) = \sum_{x_n \leq x} u_n + \sum_{x_n < x} v_n,$$

defined on (a, b) *is called a saltus function. It is continuous everywhere except the points* $\{x_n\}_{n=0}^{\infty}$ *and has jumps from the left and right at* x_n *equal to* u_n *and* v_n *respectively.*

Jump from the left at x refers to $|f(x) - f(x-)|$, and jump from the right at x refers to $|f(x+) - f(x)|$.

From the above definition a saltus function can be badly behaved. For instance the set $\{x_n\}_{n=0}^{\infty}$ can be chosen as the rational numbers in (a, b), so that $f(x)$ has discontinuities which are dense in (a, b).

Let the real numbers in $[0, 1]$ be given by their ternary expansions (expansions in base three); i.e. if $x \in [0, 1]$, then

$$x = 0.a_1a_2a_3\ldots,$$

where $0 \leq a_i \leq 2$ is a natural number for each i. The Cantor set is the set of real numbers with ternary expansions which contain no $1's$ (cf. [31], pp.27-29). A number whose ternary expansion ends with $2000\ldots$ can also be represented by an expansion which ends $1222\ldots$. In this instance the representation which ends $2000\ldots$ is chosen. A number whose ternary expansion ends $0222\ldots$ can also be represented by an expansion which ends $1000\ldots$. In this instance the representation which ends $0222\ldots$ is chosen. The Cantor set can be obtained constructively.

Divide $[0, 1]$ into three and remove the open middle third. Then remove the open middle thirds of the remaining outer intervals and at each step remove the open middle thirds of the remaining intervals. This algorithm works because the middle third interval has 0.1 beginning the ternary expansion of any number contained in it, and the closed outer intervals have 0.0 and 0.2 respectively. At the n^{th} step, numbers in the remaining intervals have ternary expansions with no $1's$ in the first n places, so that in the limit $n \to \infty$ the

Cantor set is obtained. The total length of the remaining set at each step is $\left(\frac{2}{3}\right)^n \to 0$ as $n \to \infty$. It follows that the Cantor set has Lebesgue measure zero.

The Cantor function $\omega(x)$ maps $x = 0.a_1a_2\ldots$ in the Cantor set to $\omega(x) = 0.b_1b_2\ldots$, where $b_i = \frac{a_i}{2}$ for each i and $f(x)$ is interpreted as the binary expansion of a real number in $[0,1]$ (cf. [31], pp.29,30). For x not in the Cantor set, it follows from the construction that x is in one of the open intervals removed at some step in the algorithm. The end points of such an interval are in the Cantor set and the image of the smaller endpoint under the Cantor function is the same as that of the larger endpoint (cf. [31], p.29). For x in such an interval let $\omega(x) = \omega(x')$ where x' is the smaller endpoint of the interval. Defined like this the Cantor function is continuous and non-decreasing (cf. [31], pp.29,30) but not absolutely continuous. On the one hand it maps a null set (the Cantor set) onto a set with positive measure (the interval $[0,1]$) and on the other hand it is constant almost everywhere (so $\omega'(x) = 0$ almost everywhere) and can't be represented as the indefinite integral of its almost everywhere derivative.

Definition 2.3.8 (cf. [44], p.53)**.** *A continuous function $f(x)$ of bounded variation satisfying $f'(x) = 0$ almost everywhere, is called a singularly continuous function.*

Examples of singularly continuous functions are given in [25]. Singularly continuous functions are often associated with fractal sets. Orthogonal polynomials associated with Julia sets have been used to solve problems related to singularly continuous distributions (cf. [9]).

The three cases above exhaust the characterisation of distributions and give rise to a canonical decomposition of distribution functions.

Theorem 2.3.9 (cf. [44], pp.15, 53)**.** *Every function $f(x)$ of bounded varia-*

tion can be decomposed into a sum

$$f(x) = j(x) + a(x) + s(x),$$

where $j(x)$ *is a pure jump saltus function,* $a(x)$ *is absolutely continuous and* $s(x)$ *is singularly continuous.*

$j(x)$ is constructed to have the same jumps and discontinuities as $f(x)$. Using Theorem 2.3.6,

$$a(x) = \int_a^x f'(x)dx - f(a),$$

where $f'(x)$ is the almost everywhere derivative of $f(x)$. The function $s(x)$ is then continuous and singularly continuous because $f'(x) = a'(x)$ almost everywhere and $j'(x) = 0$ almost everywhere (cf. [44], pp.11, 15, 52, 53).

2.4 Stieltjes' and Hausdorff's problems

The example of the Charlier polynomials where the function $\alpha(x)$ which generated the Riemann-Stieltjes integral was a step function, shows the importance of where the function $\alpha(x)$ is increasing in calculating the value of the integral. In particular intervals where $\alpha(x)$ is constant can be disregarded in the calculation.

Definition 2.4.1 (cf. [3], p.46)**.** x *is a point of increase (or point of growth) of a distribution function* $\alpha(x)$*, if*

$$\forall \epsilon > 0, \;\; \alpha(x+\epsilon) - \alpha(x-\epsilon) > 0.$$

Let $\alpha(x)$ be a bounded non-decreasing function satisfying

$$\int_{-\infty}^{\infty} x^n d\alpha(x) < \infty,$$

$n = 0, 1, 2, \ldots$ where the integral is a Riemann-Stieltjes integral. Furthermore suppose that $\alpha(x)$ is constant for all values of x less than zero (i.e. all points of increase of $\alpha(x)$ occur in $[0, \infty)$). Then the integral can be rewritten

$$\int_0^\infty x^n d\alpha(x).$$

Integrating each x^n generates a sequence of real constants $\{\mu_n\}_{n=0}^\infty$. A natural question that arises is when an arbitrary sequence of real constants can be represented by a Stieltjes integral like this.

Definition 2.4.2 (cf. [52], p.327)**.** *Given an arbitrary sequence of real constants $\{\mu_n\}_{n=0}^\infty$, the problem of finding a bounded non-decreasing function $\alpha(x)$ satisfying*

$$\int_0^\infty x^n d\alpha(x) = \mu_n,$$

is called the Stieltjes moment problem, and the constants $\{\mu_n\}_{n=0}^\infty$ are called moments.

Stieltjes solved this problem and invented the Stieltjes integral in the same famous paper (cf. [48]). Stieltjes' problem was a generalisation of a problem formulated by Chebyshev while studying what he called 'the limiting values of integrals' (cf. [11]).

Definition 2.4.3 (cf. [45], p.8)**.** *Let $\{\mu_n\}_{n=0}^\infty$ be a sequence of real constants. The problem of finding a bounded non-decreasing function $\alpha(x)$ satisfying*

$$\int_0^1 x^n d\alpha(x) = \mu_n,$$

is called the Hausdorff moment problem.

The Hausdorff moment problem is a specific instance of the Stieltjes moment problem; any solution to a Hausdorff moment problem is also the solution to a Stieljes moment problem. The solution of the Hausdorff moment problem is related to the theory of totally monotone sequences (cf. [53], [28]).

2.5 Hamburger's moment problem

Definition 2.5.1 (cf. [15], p.71)**.** *Given an arbitrary sequence of real numbers $\{\mu_n\}_{n=0}^{\infty}$, the problem of finding a distribution function $\alpha(x)$ satisfying,*

$$\int_{-\infty}^{\infty} x^n d\alpha(x) = \mu_n,$$

is called the Hamburger moment problem.

This generalisation of the Stieltjes moment problem was undertaken by Hamburger in [26]. An immmediate question is whether the solution to a moment problem is unique.

Definition 2.5.2 (cf. [45], p.9)**.** *A moment problem is called determined (or determinate) if any two solutions are equivalent.*

Definition 2.5.3 (cf. [45], p.52)**.** *A moment problem is called indeterminate if there exist solutions which are not equivalent.*

Any solution to the Stieltjes moment problem is also a solution to the Hamburger moment problem.

Because the interval $(-1, 1)$ can be linearly transformed into any other finite interval, the conditions for solving this problem are very similar to the conditions for solving the moment problem for another finite interval (cf. [4]). This problem is more general than Hausdorff's moment problem because a solution need not also be a solution to Stieltjes' moment problem.

Definition 2.5.4 (cf. [1], p.2)**.** *In analogy with the work in Chapter 1 a functional μ is defined on the space of polynomials of a real variable by*

$$\mu\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n} a_k \mu_k$$

where μ_k is the k^{th} element of a given sequence of moments.

By definition the functional μ is linear.

Definition 2.5.5 (cf. [45], p. xiii). *A linear functional μ on a space of functions M is non-negative if whenever $f(x) \geq 0$ and $f(x) \in M$,*

$$\mu(f(x)) \geq 0.$$

If $\alpha(x)$ is a distribution function it is obvious that

$$\int_a^b f(x)d\alpha(x) \geq 0$$

for any continuous $f(x)$ satisfying, $f(x) \geq 0$, $x \in (a, b)$.

Lemma 2.5.6 (cf. [52], p.244). *Let $P_n(x)$ be an arbitrary polynomial which is greater than or equal to zero over (a, b) and not identically zero. Then $\alpha(x)$ is a distribution function with infinitely many points of increase in an interval (a, b) if and only if*

$$\int_a^b P_n(x)d\alpha(x) > 0$$

Proof. The proof is similar to the proof of Lemma 1.1.1. No polynomial has infinitely many zeros, so there is an interval containing a point of increase of $\alpha(x)$, such that $P_n(x) > \epsilon_1 > 0$ on this interval. $\alpha(x)$ has positive variation over the interval equal to ϵ_2. The contribution of this interval to the integral is at least $\epsilon_1\epsilon_2 > 0$. For the converse if $\alpha(x)$ has finitely many points of increase then the polynomial $p_n(x)$ with double roots at each of these points and no other roots gives

$$\int_a^b p_n(x)d\alpha(x) = 0.$$

□

Definition 2.5.7 (cf. [1], p.2). *A linear functional μ on the space of polynomials is called positive if, whenever an arbitrary polynomial $P_n(x)$ satisfies $P_n(x) \geq 0$ and $P_n(x)$ is not identically zero then $\mu(P_n(x)) > 0$.*

So a necessary condition for the Hamburger moment problem to have a solution is that μ be a non-negative functional on the space of polynomials. In order that the solution have infinitely many points of increase it is necessary that μ be a positive functional on the space.

To solve the Hamburger moment problem it is necessary to establish which polynomials are non-zero on the entire axis $(-\infty, \infty)$.

Lemma 2.5.8 (cf. [41], p.77). *Any polynomial $P_n(x)$ which is non-negative on the entire real axis can be represented by,*

$$P_n(x) = q^2(x) + r^2(x)$$

where $q(x)$ and $r(x)$ are polynomials with real coefficients.

Proof. Any polynomial with real coefficients can be factorised into the product of linear factors and irreducible quadratic factors. For a polynomial to be non-negative on the entire real line the linear factors must have even multiplicity, and the irreducible quadratic factors must be of the form

$$(x - x_0)^2 + y_0^2, \; x_0, \; y_0 \in \mathbb{R}. \tag{2.5.1}$$

This follows by completing the square and recognising that the polynomial must be non-negative for $x = x_0$. Because y_0 is allowed to be 0 (2.5.1) accounts for pairs of linear factors as well. The identity

$$(p_1^2 + l_1^2)(p_2^2 + l_2^2) = (p_1 p_2 + l_1 l_2)^2 + (p_1 l_2 - p_2 l_1)^2 \tag{2.5.2}$$

can be verified by multiplying out both sides. Using (2.5.1) pairs of double linear factors and/or irreducible quadratic factors have the form of the left hand side of (2.5.2). The right hand side of (2.5.2) has the form of one of the factors on the left hand side. Repeatedly applying (2.5.2) reduces the original polynomial to an expression which has the form of the right hand side of (2.5.2), where $p_1, \; p_2, \; l_1$ and l_2 may be polynomials of any degree. This yields the result. □

Multiplying out the square on one of the polynomials above gives

$$q^2(x) = \sum_{i=0}^{n} \sum_{j=0}^{n} x^{i+j} a_i a_j,$$

where a_i and a_j are real numbers arising from the representation $q(x) = \sum_{k=0}^{n} a_k x^k$. Applying μ to this sum gives

$$\mu(q^2(x)) = \sum_{i=0}^{n} \sum_{j=0}^{n} \mu_{i+j} a_i a_j.$$

Because a_i and a_j can take on arbitrary real values, the necessary condition stated earlier, but now specific to the Hamburger moment problem is

$$\sum_{i=0}^{n} \sum_{j=0}^{n} \mu_{i+j} a_i a_j \geq 0,$$

for a_i, a_j arbitrary real numbers.

The non-negativity of these quadratic forms is equivalent to the non-negativity of determinants generated by the moment sequence. This idea arises in the theory of real symmetric matrices (cf. [6], pp.479-485).

Theorem 2.5.9 (cf. [45], p.5)**.** *For the Hamburger moment problem corresponding to the sequence $\{\mu_n\}_{n=0}^{\infty}$ to have a solution, it is necessary that the determinants*

$$D_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}$$

be non-negative. For the solution to have infinitely many points of increase it is necessary that these determinants be positive.

As a convention $D_{-1} = 1$ unless otherwise stipulated. Let μ be a positive functional defined on polynomials as above. Define $\langle \cdot, \cdot \rangle$ for real polynomials of a real variable by

$$\langle P_n(x), P_m(x) \rangle = \mu(P_n(x) P_m(x))$$

where $P_n(x)$ and $P_m(x)$ are arbitrary such polynomials. This definition is almost the same as that given in Chapter 1 but without appeal to an integral.

It is easy to establish that $\langle \cdot, \cdot \rangle$ is linear in both of its arguments, satisfies $\langle P_n(x), P_m(x) \rangle = \langle P_m(x), P_n(x) \rangle$ and because of positivity, $\mu(P_n(x)) > 0$ for any polynomial $P_n(x) \geq 0$ and not identically zero. As a result, a positive moment functional μ generates a canonical inner product on the space of polynomials, and the Gram-Schmidt algorithm can be applied to generate a set of polynomials orthogonal with respect to μ.

Using the moments associated with a set of orthogonal polynomials, it is a simple matter to investigate uniqueness. This furnishes a proof of Theorem 1.1.8. The coefficients of an orthogonal polynomial satisfy the equations (cf. [4], p.5)

$$\begin{aligned}
&a_0\mu_0 + a_1\mu_1 + \ldots + a_k\mu_k = 0\\
&a_0\mu_1 + a_1\mu_2 + \ldots + a_k\mu_{k+1} = 0\\
&\ \vdots \qquad\ \vdots \qquad\ \vdots \qquad\ \vdots\\
&a_0\mu_{k-1} + a_1\mu_k + \ldots + a_k\mu_{2k-1} = 0.
\end{aligned}$$

The system has rank k because $D_{k-1} > 0$ so the solution space is one dimensional.

Lemma 2.5.10 (cf [1], p.4)**.** *The n^{th} orthogonal polynomial associated with a positive functional is given by the formula*

$$\begin{vmatrix}
\mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n\\
\mu_1 & \mu_2 & \cdots & \mu_n & \mu_{n+1}\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
\mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1}\\
1 & x & \ldots & x^{n-1} & x^n
\end{vmatrix}$$

Proof. This can be seen by multiplying the last row by x^m, $m \leq n$ and applying μ. For all $m < n$ the determinant will have linearly dependent rows and for $m = n$ it will be greater than zero. $\square$

By Lemma 1.1.12 $\mu(P_n^2(x)) = k_n\mu(P_n(x)x^n)$ where k_n is the leading coefficient of $P_n(x)$. $\mu(P_n(x)x^n) = D_n$ and expanding along the bottom row in the determinant representation above gives $k_n = D_{n-1}$. These facts give the determinant representation of the orthonormal polynomials (cf. [1], p.3)

$$\frac{1}{\sqrt{D_{n-1}D_n}}\begin{vmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_n & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-2} & \mu_{2n-1} \\ 1 & x & \dots & x^{n-1} & x^n \end{vmatrix}. \tag{2.5.3}$$

The proofs of identities such as the three-term recurrence relation, Christoffel-Darboux formula and properties of zeros given in the first chapter only used the fact that the positive weight function $w(x)$ generated a positive functional on polynomials, by means of integration. As a result all of these identities are established for general orthogonal polynomials orthogonal with respect to some positive functional.

Consider the three-term recurrence relation

$$P_{n+1}(x) = (a_n x + b_n)P_n(x) - c_n P_{n-1}(x),$$

where for k_n the leading coefficient of $P_n(x)$, $a_n = \dfrac{k_{n+1}}{k_n}$ and $c_n = \dfrac{a_n}{a_{n-1}}\dfrac{h_n}{h_{n-1}}$. If $\{P_n(x)\}_{n=0}^{\infty}$ is an orthonormal set of polynomials, then $h_n = 1$ and dividing the recurrence by a_n results in

$$xP_n(x) = \frac{1}{a_n}P_{n+1}(x) - \frac{b_n}{a_n}P_n(x) + \frac{1}{a_{n-1}}P_{n-1}(x).$$

So the recurrence can be expressed in terms of two sequences $\{d_n\}_{n=0}^{\infty}$ and $\{e_n\}_{n=0}^{\infty}$ where $d_n = \dfrac{1}{a_n}$ and $e_n = -\dfrac{b_n}{a_n}$ for each n. Using the determinant form of the orthonormal polynomials gives $k_n = \dfrac{\sqrt{D_{n-1}}}{\sqrt{D_n}} > 0$ so that $d_n > 0$ for each n (cf. [4], p214).

These two sequences can be used to construct an infinite Jacobi matrix

$$\begin{bmatrix} e_0 & d_0 & 0 & 0 & \dots \\ d_0 & e_1 & d_1 & 0 & \dots \\ 0 & d_1 & e_2 & d_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

which plays an important role in the connection between orthogonal polynomials and distribution functions. This matrix has also been used to connect the Hamburger moment problem with spectral analysis of Jacobi matrices (cf. [24]).

Unlike the monic orthogonal polynomials, the orthonormal polynomial of the n^{th} degree is not uniquely determined. If $P_n(x)$ is an orthonormal polynomial of degree n then so is $-P_n(x)$. By choosing the leading coefficients of the orthonormal polynomials so that they alternate in sign a set of orthonormal polynomials can be constructed such that $d_n < 0$ for each n in the recurrence relation and Jacobi matrix. However, if it is assumed that $k_n > 0$ for each n, then $d_n > 0$.

2.6 Existence of solutions

Throughout the preceeding discussion an important idea has been the functional generated first by a weight function, then by a distribution function and finally by a moment sequence. The result that a distribution function generates a non-negative functional on its Riemann-Stieltjes integrable functions demonstrated the necessity that the functional associated with a moment sequence also be non-negative, because the integral is always an extension of a moment functional.

If a specific interval is under consideration a slight modification to the earlier definition of non-negativity is necessary. The linear functional μ will be called non-negative if for $f(x) \geq 0 \ \ \forall x \in (a,b)$, $\mu(f(x)) \geq 0$. If μ is a non-negative functional and $f(x) \geq g(x)$ on (a,b) then $f(x) - g(x) \geq 0$ on (a,b) and $\mu(f(x)) - \mu(g(x)) \geq 0$, so $\mu(f(x)) \geq \mu(g(x))$.

Theorem 2.6.1 (cf. [4], p.127)**.** *Let P be the space of real polynomials of a real variable, (a,b) be a given interval of real numbers (the case of a and b infinite included) and $\{x_k\}_{k=0}^{\infty}$ be the rational numbers contained in (a,b). A* *non-negative linear functional μ defined on P can be extended to the set of functions*

$$w_{x_k}(x) = \begin{cases} 1 & if \quad a \leq x \leq x_k \\ 0 & if \quad x_k < x \leq b \end{cases}$$

preserving its non-negativity.

Proof. Start with w_{x_0}. Let $P_n(x)$ denote an arbitrary polynomial satisfying $P_n(x) \leq w_{x_0}(x)$ and $P_m(x)$ denote an arbitrary polynomial satisfying $w_{x_0}(x) \leq P_m(x)$, in particular if $P_n(x)$ is 0 and $P_m(x)$ is 1 this holds. As a result there is at least one number ζ satisfying

$$\sup_{P_n(x)} \mu(P_n(x)) \leq \zeta \leq \inf_{P_m(x)} \mu(P_m(x)).$$

Choose an arbitrary such number ζ and define $\mu(w_{x_0}(x)) = \zeta$. Then it is shown below that for any polynomial $P_r(x)$ in P and real constant c satisfying,

$$P_r(x) + cw_{x_0}(x) \geq 0,$$

$$\mu(P_r(x) + cw_{x_0}(x)) = \mu(P_r(x)) + c\mu(w_{x_0}(x)) \geq 0,$$

where $x \in (a, b)$. If $c > 0$ then $-\frac{1}{c}P_r(x)$ is a polynomial less than or equal to $w_{x_0}(x)$ on (a, b) and by the above

$$-\frac{1}{c}\mu(P_r(x)) \leq \mu(w_{x_0}(x)),$$

which gives the required result. If $c < 0$ then $-\frac{1}{c}P_r(x)$ is a polynomial greater than or equal to $w_{x_0}(x)$ on (a, b) and by the above

$$-\frac{1}{c}\mu(P_r(x)) \geq \mu(w_{x_0}(x)),$$

which similarly gives the required result. This process is repeated for $w_{x_1}(x)$ except that now the supremum and infimum can be take over linear combinations of polynomials and $w_{x_0}(x)$. In this way the process is continued for $w_{x_k}(x)$, $k = 2, 3, \ldots$ and because the set of functions $\{w_{x_k}(x)\}_{k=0}^{\infty}$ is countable and well-ordered each function in the set is included at some step in the algorithm. $\square$

Having extended the functional to this set of functions it is now possible to construct a solution to the moment problem. The specific problem considered will be the Hamburger problem.

Theorem 2.6.2 (cf. [1], p.71, [4], p.126)**.** *For the Hamburger moment problem to have a solution it is necessary and sufficient that the moment functional generated by it be non-negative.*

Proof. Necessity has been deomonstrated on page 54. For sufficiency assume that μ is a non-negative functional generated by a Hamburger moment problem and also assume that μ has been extended to the set of functions $\{w_{x_k}(x)\}_{k=0}^{\infty}$ discussed above. Define $\alpha(x)$ for $x \in \{x_k\}_{k=0}^{\infty} = \mathbb{Q}$ by,

$$\alpha(x_k) = \mu(w_{x_k}(x)).$$

If $x_k \geq x_j$, then by defintion $w_{x_k}(x) \geq w_{x_j}(x)$, so

$$\alpha(x_k) \geq \alpha(x_j). \tag{2.6.1}$$

The rational numbers are dense in $\mathbb{R}$ so $\alpha(x)$ can be extended to $\mathbb{R}$ by

$$\alpha(x) = \sup_{x_k < x} \alpha(x_k), \tag{2.6.2}$$

where x_k is a rational number. The function $\alpha(x)$ is non-decreasing and bounded. The non-decreasing property follows from (2.6.1) and (2.6.2). The boundedness is ensured because $\sup\limits_{x} \alpha(x) = \mu(1) = \mu_0$ and $\inf\limits_{x} \alpha(x) = \mu(0) = 0$. Choose points

$$-B = \tau_0 < \tau_1 < \ldots < \tau_N = B,$$

where $\tau_j \in \mathbb{Q}$ and $B > 1$ for every j, such that in each interval $[\tau_i, \tau_{i+1}]$, $\max\limits_{x \in [\tau_i, \tau_{i+1}]} x^k - \min\limits_{x \in [\tau_i, \tau_{i+1}]} x^k < \epsilon$. Construct the functions

$$F_N^k(x) = \sum_{j=0}^{n-1} \tau_j^k \{w_{\tau_{j+1}}(x) - w_{\tau_j}(x)\}.$$

If $|x| > B$ then the $w_{\tau_j}(x)$ are equal to the same number (1 or 0) for each j so $F_N^k(x) = 0$. If $x \in (\tau_i, \tau_{i+1}]$ then $w_{\tau_j}(x) = 0$ for $j \le i$ and for $j > i$, $w_{\tau_{j+1}}(x) = w_{\tau_j}(x) = 1$ so that the only non-zero term in the sum is $\tau_i^k w_{\tau_{i+1}}(x)$ i.e. in this case $F_N^k(x) = \tau_i^k$. Because of the condition in choosing the τ_i, $i = 0, 1, \ldots, N$, if $-B \le x \le B$ and k is odd then $0 \le x^k - F_N^k(x) < \epsilon$ whereas if k is even and $-B \le x < 0$ then $-\epsilon < x^k - F_N^k(x) \le 0$ and if $0 \le x \le B$ then $0 \le x^k - F_N^k(x) < \epsilon$. So for $-B \le x \le B$

$$|x^k - F_N^k(x)| < \epsilon.$$

Let $n = 2k$. For $|x| > B$, $x^k - F_N^k(x) = x^k$ and because $B > 1$ for k even $0 < x^k < \dfrac{x^n}{B}$ and for k odd $|x^k| < \dfrac{x^n}{B}$. This combined with the other inequality gives in the entire interval $(-\infty, \infty)$,

$$|x^k - F_N^k(x)| < \epsilon + \frac{x^n}{B},$$

or

$$-\frac{x^n}{B} - \epsilon < x^k - F_N^k(x) < \epsilon + \frac{x^n}{B}.$$

Apply μ to both sides to get

$$-\frac{\mu_n}{B} - \mu_0\epsilon < \mu_k - \sum_{j=0}^{N-1} \tau_j^k \{\alpha(\tau_{j+1}) - \alpha(\tau_j)\} < \mu_0\epsilon + \frac{\mu_n}{B}.$$

Letting $\delta \to 0$ where $\delta = \max_i\{\tau_{i+1} - \tau_i\}$ causes ϵ to go to zero and generates a Stieltjes integral, which exists because x^k is continuous and $\alpha(x)$ has bounded variation. The inequality now reads

$$-\frac{\mu_n}{B} < \mu_k - \int_{-B}^{B} x^k d\alpha(x) < \frac{\mu_n}{B}.$$

Letting B tend to infinity gives the result. $\square$

Theorem 2.6.3. *A Hamburger moment problem $\{\mu_n\}_{n=0}^{\infty}$ has a solution with infinitely many points of increase if and only if the associated functional μ is positive.*

Proof. By the above theorem a Hamburger moment problem has a solution if and only if the associated functional is non-negative. By Lemma 2.5.6 and Definition 2.5.7 the solution has infinitely many points of increase if and only if the associated functional is positive. □

The strength of this approach is that the same proof with appropriate, but not substantial, modifications can be used to show that for an arbitrary interval (a, b) on the real line, a necessary and sufficient condition for a solution to exist is that the functional μ generated by the problem be non-negative on polynomials relative to the interval (cf. [45]).

Lemma 2.6.4 (cf. [41], p.78)**.** *Any polynomial which is non-negative in* $[0, \infty)$ *can be represented as*

$$q^2(x) + r^2(x) + x\{s^2(x) + t^2(x)\},$$

where $q(x)$*,* $r(x)$*,* $s(x)$ *and* $t(x)$ *are polynomials with real coefficients.*

Proof. All roots of odd multiplicity of a polynomial $P(x)$ which is non-negative for $x \in [0, \infty)$ are non-positive. $P(x)$ can be factorised with factors

$$(x - x_0)^2 + y_0^2,\ x_0,\ y_0 \text{ real, and } x + x_1,\ x_1 \geq 0.$$

Either of these factors can be represented by the expression

$$p_1^2 + q_1^2 + x(r_1^2 + s_1^2).$$

The right hand side of the identity

$$\begin{aligned}
&[p_1^2 + q_1^2 + x(r_1^2 + s_1^2)][p_2^2 + q_2^2 + x(r_2^2 + s_2^2)] \\
&= [(p_1^2 + q_2^2)(p_2^2 + q_2^2) + x^2(r_1^2 + s_1^2)(r_2^2 + s_2^2)] \\
&+ x[(p_1^2 + q_1^2)(r_2^2 + s_2^2) + (r_1^2 + s_1^2)(p_2^2 + q_2^2)]
\end{aligned}$$

has two terms which are non-negative on the whole real line. Applying Lemma 2.5.8 then gives the result. □

Theorem 2.6.5 (cf. [45], p.5). *A necessary and sufficient condition for the Stieltjes moment problem to have a solution is that*

$$\sum_{i=0}^{n}\sum_{j=0}^{n}\mu_{i+j}a_i a_j \geq 0 \tag{2.6.3}$$

and

$$\sum_{i=0}^{n}\sum_{j=0}^{n}\mu_{i+j+1}a_i a_j \geq 0, \tag{2.6.4}$$

for a_i, a_j arbitrary real numbers.

Proof. From the above a necessary and sufficient condition for a solution to exist is that the functional μ generated by the problem be non-negative on $[0,\infty)$. $\mu(q^2(x))$ gives (2.6.3) and $\mu(xs^2(x)) = \mu\left(\sum_{i=0}^{n}\sum_{j=0}^{n} xx^{i+j}a_i a_j\right)$ which gives (2.6.4). □

Using the criteria for existence of a solution to the Hamburger moment problem the fundamental theorem connecting orthogonal polynomials and the moment problem can be established. This result establishes that any Jacobi matrix

$$\begin{bmatrix} e_0 & d_0 & 0 & 0 & \dots \\ d_0 & e_1 & d_1 & 0 & \dots \\ 0 & d_1 & e_2 & d_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

with $e_n \in \mathbb{R}$ and $d_n > 0$ for all n corresponds to a distribution function $\alpha(x)$. The polynomials $\{P_n(x)\}_{n=0}^{\infty}$ which satisfy a three-term recurrence relation

$$xP_n(x) = d_n P_{n+1}(x) + e_n P_n(x) + d_{n-1}P_{n-1}(x),$$

with parameters $\{e_n\}_{n=0}^{\infty}$ and $\{d_n\}_{n=0}^{\infty}$ are orthonormal with respect to $\alpha(x)$. This result is known as Favard's theorem.

Theorem 2.6.6 (**Favard's Theorem** cf. [4], p.216). *Let sequences $\{d_n\}_{n=0}^{\infty}$ and $\{e_n\}_{n=0}^{\infty}$ be given such that $e_n \in \mathbb{R}$ and $d_n > 0$. Then there is a distribution function $\alpha(x)$ such that the polynomials $P_n(x)$ generated by the recurrence* *relation*

$$xP_n(x) = d_n P_{n+1}(x) + e_n P_n(x) + d_{n-1} P_{n-1}(x),$$

$P_{-1}(x) = 0$, $P_0(x) = \dfrac{1}{\sqrt{\mu_0}}$, *where* $\mu_0 > 0$ *is arbitrary, are orthonormal with respect to* $\alpha(x)$.

Proof. Because $\{P_n(x)\}_{n=0}^{\infty}$ is a simple set, two arbitrary real polynomials $G(x)$, $H(x)$ each of degree n, can be expanded as a linear combination of polynomials $P_k(x)$, $0 \leq k \leq n$ such that

$$G(x) = \sum_{k=0}^{n} \zeta_k P_k(x)$$

$$H(x) = \sum_{k=0}^{n} \eta_k P_k(x).$$

Define $\langle \cdot, \cdot \rangle$ by,

$$\langle G(x), H(x) \rangle = \sum_{k=0}^{n} \eta_k \zeta_k.$$

It follows that

1. $\langle G(x), H(x) \rangle = \langle H(x), G(x) \rangle$

2. $\langle G_1(x) + G_2(x), H(x) \rangle = \langle G_1(x), H(x) \rangle + \langle G_2(x), H(x) \rangle$

3. $\langle \alpha G(x), H(x) \rangle = \alpha \langle G(x), H(x) \rangle$

4. $\langle G(x), G(x) \rangle > 0$, if $G(x)$ is not identically 0.

It is natural to define a functional μ on polynomials $R(x)$ by decomposing $R(x)$ into factors $S(x)$ and $T(x)$ so that

$$\mu(R(x)) = \langle S(x), T(x) \rangle,$$

but for this definition to hold it is must be shown that μ takes the same value irrespective of the decomposition of $R(x)$ into factors. To this end note first that the three-term recurrence relation gives

$$xP_i(x) = d_i P_{i+1}(x) + e_i P_i(x) + d_{i-1} P_{i-1}(x)$$

and the trivial sum $P_i(x) = \sum_{k=0}^{n} \delta_{ik} P_k(x)$ holds for $P_i(x)$, where $i = 0, 1, 2 \ldots, n$ and $n = 0, 1, 2 \ldots$. These facts give

$$\langle xP_i(x), P_k(x) \rangle = \langle P_i(x), xP_k(x) \rangle. \tag{2.6.5}$$

If $G(x)$ and $H(x)$ are arbitrary polynomials where $xG(x) = x \sum_{k=0}^{n} \zeta_k P_k(x)$, $xH(x) = x \sum_{k=0}^{n} \eta_k P_k(x)$ then (2.6.5) and linearity of $\langle \cdot, \cdot \rangle$ give

$$\langle xG(x), H(x) \rangle = \langle G(x), xH(x) \rangle. \tag{2.6.6}$$

Finally if $F(x)$ is some polynomial of degree not greater than n then $F(x) = \sum_{k=0}^{n} a_k x^k$ and (2.6.6) leads to the conclusion

$$\langle F(x)G(x), H(x) \rangle = \langle G(x), F(x)H(x) \rangle,$$

which establishes that μ is well-defined. A moment sequence can be defined by

$$\mu_{i+j} = \mu(x^{i+j}), \; i = 0, 1, 2, \ldots, j = 0, 1, 2, \ldots.$$

By definition $\mu(P_n(x)P_m(x)) = \langle P_n(x), P_m(x) \rangle = \delta_{nm}$. For this sequence to generate a distribution function it is sufficient that the quadratic forms,

$$\sum_{i=0}^{n} \sum_{j=0}^{n} \mu_{i+j} a_i a_j, \tag{2.6.7}$$

where a_i, a_j are real numbers be non-negative. Equation (2.6.7) is $\mu(q^2(x)) = \langle q(x), q(x) \rangle > 0$ for some polynomial with real coefficients $q(x)$. The condition is satisfied and a distribution function $\alpha(x)$ can be constructed so that

$$\int_{-\infty}^{\infty} P_n(x) P_m(x) d\alpha(x) = \delta_{nm},$$

because the integral is an extension of the functional μ. $\square$

Conditions have been found which relate the determinacy of a Hamburger moment problem to the coefficients in the three-term recurrence relation of a set of orthogonal polynomials (cf. [13]).

If $\alpha_1(x)$ is a distribution generated by the Jacobi matrix with the choice $\mu_0 = 1$ then $\alpha_2(x) = c\alpha_1(x)$ is a distribution generated by the Jacobi matrix with the choice $\mu_0 = c$. Without loss of generality the choice $\mu_0 = 1$ will be used throughout the rest of the discussion (cf. [1], p.3).

Under this choice of μ_0, k_0 the coefficient of the 0^{th} orthonormal polynomial is $\dfrac{1}{\sqrt{\mu_0}} = 1$. So $d_n = \dfrac{k_n}{k_{n+1}}$ gives

$$\prod_{i=0}^{n-1} \frac{1}{d_i} = \prod_{i=0}^{n-1} \frac{k_{i+1}}{k_i} = \frac{k_n}{k_0} = k_n.$$

As a result if $p_n(x)$ is the n^{th} monic orthogonal polynomial and $P_n(x)$ is the n^{th} orthonormal polynomial then

$$P_n(x) = k_n p_n(x) = \left\{ \prod_{i=0}^{n-1} \frac{1}{d_i} \right\} p_n(x). \tag{2.6.8}$$

2.7 The true interval of orthogonality

Theorem 2.7.1 (cf. [15], p.59). *If $\alpha(x)$ is a distribution function with infinitely many points of increase and $\{P_n(x)\}_{n=0}^{\infty}$ is the set of orthogonal polynomials generated by $\alpha(x)$ on the interval (a, b), then between any two zeros of $P_n(x)$ there is a point of increase of $\alpha(x)$.*

Proof. Suppose that there are two zeros x_1 and x_2 of $P_n(x)$ such that $\alpha(x)$ has no point of increase between them. Then the polynomial

$$Q(x) = P_n^2(x) \frac{1}{(x - x_1)(x - x_2)}$$

is non-negative outside the interval (x_1, x_2), and because $\alpha(x)$ has no points of increase in (x_1, x_2) this interval contributes nothing to the integral so that

$$\int_a^b Q(x) d\alpha(x) > 0.$$

This contradicts the orthogonality condition because $Q(x)$ is the product of an orthogonal polynomial and a polynomial of lower degree. □

This establishes that wherever there are zeros of orthogonal polynomials there are also points of increase of the associated distribution. It can be shown (cf. [45], pp.106-113) that a solution of the associated moment problem exists with all of its points of increase contained in the smallest closed interval containing the roots of all of the orthogonal polynomials $\{P_n(x)\}_{n=0}^{\infty}$. This interval is also the smallest interval for which there exists a solution with all points of increase contained in it.

2.8 The trigonometric moment problem

Instead of the functions $\{x^n\}_{n=0}^{\infty}$ the trigonometric moment problem examines the functions $\{e^{inx}\}_{n=-\infty}^{\infty}$, together with a sequence of constants $\{\nu_n\}_{n=-\infty}^{\infty}$, $\nu_n = \overline{\nu}_{-n}$. In analogy with the moment problems considered so far, a linear functional ν is defined on linear combinations of these functions (cf. [4], p.1)

$$\nu\left(\sum_{k=-n}^{n} a_k e^{ikx}\right) = \sum_{k=-n}^{n} a_k \nu_k.$$

Definition 2.8.1 (cf. [23], p.742)**.** *Given a sequence of constants* $\{\nu_n\}_{n=-\infty}^{\infty}$, $\nu_n = \overline{\nu}_{-n}$ *the problem of finding a bounded non-decreasing function* $\sigma(\theta)$ *such that*

$$\frac{1}{2\pi}\int_0^{2\pi} e^{inx} d\sigma(x) = \nu_n,$$

is called the trigonometric moment problem.

Using a generalisation of the approach for the Hamburger moment problem it can be established that the trigonometric moment problem is solvable if and only if the functional ν is non-negative relative to the interval $[0, 2\pi]$ (cf. [4], p.180). In the event that a solution $\sigma(\theta)$, with infinitely many points of increase, exists an inner product can be constructed on the set of complex polynomials of a complex variable by setting

$$\langle P_n(z), P_m(z)\rangle = \int_0^{2\pi} P_n(e^{i\theta})\overline{P_m(e^{i\theta})} d\sigma(\theta).$$

To verify this it suffices to establish the properties (cf. [2], p.2)

(a.) $\langle P_n(z), P_m(z)\rangle = \overline{\langle P_m(z), P_n(z)\rangle}$,

(b.) $\langle \alpha_1 P_l(z) + \alpha_2 P_m(z), P_n(z)\rangle = \alpha_1\langle P_l(z), P_n(z)\rangle + \alpha_2\langle P_m(z), P_n(z)\rangle$,

(c.) $\langle P_n(z), P_n(z)\rangle > 0$ for $P_n(z)$ not identically zero.

The first property can be established directly from the definition by considering the real and imaginary parts of the integrand. The second property follows from the linearity of the integral. The third property is a consequence of the fact that $P_n(z)\overline{P_n(z)} = |P_n(z)|^2 > 0$ and $\sigma(z)$ is non-decreasing and has infinitely many points of increase.

Using the Gram-Schmidt process a set of orthogonal polynomials given by $\{P_n(z)\}_{n=0}^{\infty}$ can be constructed. Because the integration is carried out on the unit circle of the complex plane these polynomials are often called unit circle orthogonal polynomials (or polynomials orthogonal relative to a circle (cf. [1], p.182)).

If the moments of a solvable trigonometric moment problem are real then the resulting distribution can be transformed into a distribution which is the solution of an ordinary moment problem on the interval $(-1,1)$. Conversely a distribution function which solves a moment problem on the interval $(-1,1)$ can transformed into a distribution which solves a trigonometric moment problem with real moments (cf. [22], p.169, [23], pp.757-760). The connection between orthogonal polynomials on the real line and on the unit circle has been used to transfer known facts about distributions on the real line to distributions on the unit circle (cf.[37]). Analogies between the recurrence relations for unit circle orthogonal polynomials and polynomials orthogonal on the real line have been explored (cf. [10], [21]). The trigonometric moment problem can also be solved using continued fractions. This establishes a connection between the problem and Schur's algorithm for bounded analytic functions in the unit circle (cf. [36]).

Chapter 3

Continued fractions

The earliest investigations of the moment problem were undertaken with contintued fractions (cf. [48]). They continue to be an important avenue for research in orthogonal polynomials. The chapter begins with an overview of essential theorems from the theory of continued fractions. Next, Jacobi continued fractions are introduced. Jacobi continued fractions have an essential connection with orthogonal polynomials and the classical moment problem. This connection is exhibited using asymptotic series. A truncated Jacobi continued fraction is used to present the limit circle and limit point cases which arise for indeterminate and determinate moment problems respectively. Finally the Nevanlinna parametrisation of solutions to an indeterminate moment problem is presented.

3.1 Basic theory

Definition 3.1.1 (cf. [15], p.77, [19], p.58). *Let* $\{a_n\}_{n=1}^{\infty}$ *and* $\{b_n\}_{n=0}^{\infty}$ *be infinite sequences of complex numbers. A continued fraction is defined as the* *formal expression*

$$b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{\ddots + \cfrac{a_n}{b_n + \ddots}}}.$$

If the sequences are infinite the continued fraction is called an infinite continued fraction, otherwise it is called a finite continued fraction.

This expression in some instances converges to a complex number. In order to study convergence of a continued fraction, the fraction is truncated and the behaviour of the truncated fractions is studied, in analogy with the partial sums of an infinite series.

Definition 3.1.2 (cf. [15], p.77)**.** *Let the sequences above be truncated at the n^{th} term, leaving $\{a_k\}_{k=1}^{n}$, $\{b_k\}_{k=0}^{n}$. Then the finite continued fraction*

$$C_n = b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{\ddots + \cfrac{a_n}{b_n}}},$$

is called the n^{th} convergent of the continued fraction generated by the sequences $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$.

It is difficult to gauge the behaviour of the convergents just by looking at the continued fraction. It can be shown by back substitution that an arbitrary finite continued fraction reduces to a ratio of two complex numbers. A recurrence relation exists to calculate the n^{th} convergent of an arbitrary continued fraction by providing an expression for the numerator and for the denominator.

Theorem 3.1.3 (cf. [19], p.59)**.** *Let the n^{th} convergent C_n of the continued fraction generated by the sequences $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ be given by the ratio*

$$C_n = \frac{A_n}{B_n},$$

Then A_n and B_n satisfy the recurrence relations

$$A_{n+1} = b_{n+1}A_n + a_{n+1}A_{n-1},$$

$$B_{n+1} = b_{n+1}B_n + a_{n+1}B_{n-1},$$

where $A_{-1} = 1$, $A_0 = b_0$, $B_{-1} = 0$ and $B_0 = 1$.

Proof. $C_0 = b_0$ and $C_1 = b_0 + \frac{a_1}{b_1}$. Since $C_0 = \frac{A_0}{B_0} = \frac{b_0}{1}$, and $C_1 = \frac{A_1}{B_1} = \frac{b_1 b_0 + a_1}{b_1 1 + 0} = \frac{b_1 A_0 + a_1 A_{-1}}{b_1 B_0 + a_1 B_{-1}}$ the hypothesis holds for C_1. Assume the hypothesis holds for any n^{th} convergent of an arbitrary continued fraction. C_{n+1} is the same as C_n except that $b_n + \frac{a_{n+1}}{b_{n+1}}$ is substituted for b_n, so $C'_n = C_{n+1}$ is an n^{th} convergent for some continued fraction and the hypothesis can used to calculate its value,

$$C'_n = \frac{(b_n + \frac{a_{n+1}}{b_{n+1}})A_{n-1} + a_n A_{n-2}}{(b_n + \frac{a_{n+1}}{b_{n+1}})B_{n-1} + a_n B_{n-2}}.$$

A_{n-1}, A_{n-2}, B_{n-1} and B_{n-2} are not affected by the new choice of b_n (because of the recurrence relation) so,

$$\begin{aligned} C_{n+1} &= \frac{(b_n A_{n-1} + a_n A_{n-2}) + \frac{a_{n+1}}{b_{n+1}} A_{n-1}}{(b_n B_{n-1} + a_n B_{n-2}) + \frac{a_{n+1}}{b_{n+1}} B_{n-1}} \\ &= \frac{A_n + \frac{a_{n+1}}{b_{n+1}} A_{n-1}}{B_n + \frac{a_{n+1}}{b_{n+1}} B_{n-1}}. \end{aligned}$$

Multiplying the last fraction by $\frac{b_{n+1}}{b_{n+1}}$ gives the result. $\square$

Lemma 3.1.4 (cf. [35], p.12). *Let*

$$\zeta = b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{\ddots + \cfrac{a_n}{b_n + u}}}.$$

Then

$$\zeta = \frac{A_n + uA_{n-1}}{B_n + uB_{n-1}}.$$

Proof. In Theorem 3.1.3 simply replace $\dfrac{a_{n+1}}{b_{n+1}}$ with u to obtain this result. $\square$

Lemma 3.1.5 (cf. [35], pp.12,14)**.** *If A_n, B_n, a_n and b_n are as above for each n then*

$$A_nB_{n-1} - A_{n-1}B_n = (-1)^{n-1}a_1a_2\ldots a_n,$$

and

$$\frac{A_n}{B_n} - \frac{A_{n-1}}{B_{n-1}} = \frac{(-1)^{n-1}a_1a_2\ldots a_n}{B_nB_{n-1}}, \; n \geq 1. \tag{3.1.1}$$

Proof. $A_1B_0 - A_0B_1 = (b_1b_0 + a_1)1 - (b_11)b_0 = a_1$. Assuming the hypothesis for n and using Theorem 3.1.3,

$$\begin{aligned} A_{n+1}B_n - A_nB_{n+1} &= (b_{n+1}A_n + a_{n+1}A_{n-1})B_n \\ &\quad - A_n(b_{n+1}B_n + a_{n+1}B_{n-1}) \\ &= -a_{n+1}(A_nB_{n-1} - A_{n-1}B_n) \\ &= (-1)^n a_1a_2\ldots a_na_{n+1}. \end{aligned}$$

To get the second part of the result divide both sides by B_nB_{n-1}. $\square$

It is possible to transform the parameters of a continued fraction while maintaining the same value for the convergents.

Theorem 3.1.6 (cf. [52], p.19)**.** *Let A_n and B_n be the approximants of the continued fraction*

$$b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \ddots}}}.$$

Then the continued fraction

$$b_0 + \cfrac{c_1 a_1}{c_1 b_1 + \cfrac{c_1 c_2 a_2}{c_2 b_2 + \cfrac{c_2 c_3 a_3}{c_3 b_3 + \ddots}}},$$

has approximants $A'_p = c_0 c_1 c_2 \dots c_p A_p$ *and* $B'_p = c_0 c_1 c_2 \dots c_p B_p$*, where* $c_0 = 1$*. This means that the approximants of the transformed continued fraction are the same as the original continued fraction.*

Proof. $A_{-1} = 1$, $A_0 = b_0$, $B_{-1} = 0$, $B_0 = 1$ are unaffected by the transformation so $A'_0 = c_0 A_0$ and $B'_0 = c_0 B_0$. Suppose that the hypothesis holds for k, then using the formula for computing the approximants A_p gives

$$\begin{aligned} A'_{k+1} &= c_{k+1} b_{k+1} A'_k + c_{k+1} c_k a_{k+1} A'_{k-1} \\ &= b_{k+1} c_{k+1} c_k c_{k-1} \dots c_0 A_k + a_{k+1} c_{k+1} c_k c_{k-1} \dots c_0 A_{k-1} \\ &= c_{k+1} c_k c_{k-1} \dots c_0 A_{k+1}. \end{aligned}$$

The same reasoning establishes the result for B_{k+1}. □

If $a_k \neq 0$ for each k then the parameters c_k can be determined so that $c_k c_{k-1} a_k = 1$ and all of the transformed numerators are 1.

Every rational number can be expanded as a finite continued fraction using the Euclidean division algorithm.

Theorem 3.1.7 (cf. [35], p.1)**.** *Let a rational number be given by* $\frac{x_0}{x_1}$*,* $x_0 > x_1 > 0$ *then there is a finite sequence of parameters* $\{b_k\}_{k=0}^n$ *which are positive integers such that*

$$\frac{x_0}{x_1} = b_0 + \cfrac{1}{b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_n}}}.$$

Proof. $x_0 = b_0 x_1 + x_2$ where $x_2 < x_1$ is the remainder after dividing x_0 by x_1. Similarly

$$\begin{aligned}
x_1 &= b_1 x_2 + x_3 \\
x_2 &= b_2 x_3 + x_4 \\
x_3 &= b_3 x_4 + x_5 \\
&\vdots \\
x_{n-1} &= b_{n-1} x_n + x_{n+1} \\
x_n &= b_n x_{n+1}.
\end{aligned}$$

This descending sequence of natural numbers must terminate. For each k $\dfrac{x_{k-1}}{x_k} = b_{k-1} + \dfrac{1}{\frac{x_k}{x_{k+1}}}$ and the continued fraction expansion follows. □

3.2 Jacobi continued fractions

The parameters of an orthonormal set of polynomials $\{d_n\}_{n=0}^{\infty}$ and $\{e_n\}_{n=0}^{\infty}$, $d_n \in \mathbb{R}$ and $e_n > 0$ for each n, generate a Jacobi matrix

$$\begin{bmatrix} e_0 & d_0 & 0 & 0 & \dots \\ d_0 & e_1 & d_1 & 0 & \dots \\ 0 & d_1 & e_2 & d_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{3.2.1}$$

Denote by k_n the leading coefficient of the n^{th} orthonormal polynomial. d_n was defined to be $\dfrac{1}{a_n} = \dfrac{k_n}{k_{n+1}}$. Denote by $p_n(x)$ the polynomials satisfying

$$p_{n+1}(x) = (x - e_n)p_n(x) - d_{n-1}^2 p_{n-1}(x), \tag{3.2.2}$$

$p_0(x) = 1$, $p_{-1}(x) = 0$. Let $\dfrac{1}{k_n}P_n(x) = p_n(x)$ then equation (3.2.2) gives

$$\frac{1}{k_{n+1}}P_n(x) = (x - e_n)\frac{1}{k_n}P_n(x) - d_{n-1}\frac{k_{n-1}}{k_n}\frac{1}{k_{n-1}}P_{n-1}(x).$$

Multiply by k_n and use $d_n = \dfrac{k_n}{k_{n+1}}$ to get

$$d_n P_{n+1}(x) = (x - e_n)P_n(x) - d_{n-1}P_{n-1}(x),$$

$P_0(x) = k_0$, $P_{-1}(x) = 0$. This is the three-term recurrence relation for the orthonormal polynomials so $p_n(x)$ is the n^{th} monic orthogonal polynomial.

Definition 3.2.1 (cf. [52], pp.64,103)**.** *A continued fraction*

$$\cfrac{1}{x - e_0 - \cfrac{d_0^2}{\ddots - \cfrac{d_n^2}{x - e_{n+1} - \ddots}}},$$

where $e_n \in \mathbb{R}$ and $d_n > 0$ are sequences associated with a real infinite Jacobi matrix, is called a Jacobi continued fraction.

The above definition is narrower than the definition found elsewhere (cf. [54]), but it is sufficient to encompass all cases of this continued fraction occurring here.

Using Theorem 3.1.3 to compute the approximants A_n and B_n gives

$$A_{n+1} = (x - e_n)A_n - d_{n-1}^2 A_{n-1},$$

$$B_{n+1} = (x - e_n)B_n - d_{n-1}^2 B_{n-1},$$

$A_0 = 0$, $A_1 = 1$, $B_{-1} = 0$, $B_0 = 1$. It is immediate from the preceding calculations and these formulae that B_n is the n^{th} monic orthogonal polynomial generated by the Jacobi matrix (3.2.1), and A_n is the $(n-1)^{th}$ monic orthogonal polynomial generated by the Jacobi matrix

$$\begin{bmatrix} e_1 & d_1 & 0 & 0 & \dots \\ d_1 & e_2 & d_2 & 0 & \dots \\ 0 & d_2 & e_3 & d_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{3.2.3}$$

The Euclidean division algorithm which was used to generate the continued fraction expansion of a rational number can also be used on polynomials as the familiar polynomial long division. In a completely analogous way this gives rise to a continued fraction expansion for a rational function. This analogy is thoroughly developed in [34].

Theorem 3.2.2 (cf. [35], p.248)**.** *Let* $\frac{f_1}{f_0}$*, be a rational function, i.e. a ratio of two polynomials* f_1 *and* f_0*. Furthermore let the degree of* f_1 *be less than the degree of* f_0*. Then there exist polynomials* $\{b_k\}_{k=0}^{n}$ *such that*

$$\frac{f_1}{f_0} = \cfrac{1}{b_0 + \cfrac{1}{b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_n}}}}.$$

Proof. $f_0 = b_0 f_1 + f_2$ where f_2 is the remainder after dividing f_0 by f_1 and f_2 has degree less than f_1. Similarly

$$\begin{aligned} f_1 &= b_1 f_2 + f_3 \\ f_2 &= b_2 f_3 + f_4 \\ &\vdots \\ f_{n-1} &= b_{n-1} f_n + f_{n+1} \\ f_n &= b_n f_{n+1}. \end{aligned}$$

Consequently $\frac{f_1}{f_0} = \frac{1}{\frac{f_0}{f_1}} = \cfrac{1}{b_0 + \cfrac{1}{\frac{f_1}{f_2}}}$, and in general $\frac{f_{k-1}}{f_k} = b_{k-1} + \frac{f_{k+1}}{f_k}$. The required expansion follows. □

The convergents of a Jacobi continued fraction are rational functions and an equivalent continued fraction representation for the convergents can be obtained from this algorithm. It is important to determine when an arbitrary rational function can be written as a Jacobi continued fraction.

Theorem 3.2.3 (cf. [52], pp.165-167). *Let* $\frac{f_1}{f_0}$ *be a rational function where the polynomials* f_1 *and* f_0 *are given by*

$$f_0 = a_{00}z^n + a_{01}z^{n-1} + \ldots + a_{0n},$$

$$f_1 = a_{11}z^{n-1} + a_{12}z^{n-2} + \ldots + a_{1n},$$

then the continued fraction expansion of $\frac{f_1}{f_0}$ *is a Jacobi continued fraction if the determinants*

$$\Delta_0 = a_{00}$$

$$\Delta_1 = a_{11}$$

$$\Delta_2 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{00} & a_{01} & a_{02} \\ 0 & a_{11} & a_{12} \end{vmatrix}$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{00} & a_{01} & a_{02} & a_{03} & a_{04} \\ 0 & a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{00} & a_{01} & a_{02} & a_{03} \\ 0 & 0 & a_{11} & a_{12} & a_{13} \end{vmatrix}$$

$$\vdots$$

where $a_{0p} = a_{1p} = 0$ *if* $p > n$, *are all greater than zero.*

Details of the proof can be found in [52] on pages 165-167.

3.3 Asymptotic expansions

Definition 3.3.1 (cf. [39], p.4). *Let* $f(z)$ *and* $g(z)$ *be functions of a complex variable* z. *If* $\frac{|f(z)|}{|g(z)|}$ *is bounded as* $z \to \infty$ *then*

$$f(z) = O(g(z)),$$

or $f(z)$ *is* $O(g(z))$.

Definition 3.3.2 (cf. [5], p.611). *Let $f(z)$ be a function of a complex variable in an unbounded region D of the complex plane and $\sum_{k=0}^{\infty} a_k z^{-k}$ be a formal power series. If*

$$f(z) = \sum_{k=0}^{n-1} a_k z^{-k} + R_n(z),$$

and $R_n(z) = O(z^{-n})$ as $z \to \infty$ in D, then the formal series $\sum_{k=0}^{\infty} a_k z^{-k}$ is called an asymptotic expansion for $f(z)$.

Lemma 3.3.3 (cf. [39], p.19). *If $f(z)$ has an asymptotic expansion $\sum_{k=0}^{\infty} f_k z^{-k}$ and $g(z)$ has an asymptotic expansion $\sum_{k=0}^{\infty} g_k z^{-k}$ then $h(z) = f(z)g(z)$ has an asymptotic expansion $\sum_{k=0}^{\infty} h_k z^{-k}$ where $h_k = \sum_{i=0}^{k} f_i g_{k-i}$.*

The proof can be found in [39] on page 19.

If $\sum_{k=0}^{\infty} a_k z^{-k}$ is an asymptotic expansion for $\frac{1}{G(z)}$, where $G(z)$ is a polynomial of degree n, then the first n coefficients of the expansion must be zero because $\frac{1}{G(z)}$ is $O(z^{-n})$.

Let $\frac{f_1}{f_0}$ be a rational function which can be expanded as a Jacobi continued fraction. Let the n^{th} approximant of the Jacobi continued fraction be denoted by $\frac{q_n(z)}{p_n(z)}$. Referring to (3.1.1) the approximants to the continued fraction satisfy

$$\begin{aligned} \frac{q_{n+1}(z)}{p_{n+1}(z)} - \frac{q_n(z)}{p_n(z)} &= \frac{(-1)^n(-d_0^2)(-d_1^2)(-d_2^2)(-d_3^2)\dots(-d_{n-1}^2)}{p_{n+1}(z)p_n(z)} \\ &= \frac{d_0^2 d_1^2 d_2^2 d_3^2 \dots d_{n-1}^2}{p_{n+1}(z)p_n(z)} \end{aligned} \tag{3.3.1}$$

If a rational function can be expanded as a Jacobi continued fraction then because it is a finite continued fraction (as can be seen from the expansion algorithm in Theorem 3.2.2), it is the n^{th} (and final) convergent of the continued fraction, $\frac{f_1}{f_0} = \frac{q_n(z)}{p_n(z)}$. As a result

$$\begin{aligned}\frac{f_1}{f_0} - \frac{q_i(z)}{p_i(z)} &= \sum_{k=i}^{n-1} \left(\frac{q_{k+1}(z)}{p_{k+1}(z)} - \frac{q_k(z)}{p_k(z)} \right) \\ &= \sum_{k=i}^{n-1} \frac{d_0^2 d_1^2 d_2^2 d_3^2 \dots d_{k-1}^2}{p_{k+1}(z) p_k(z)} \end{aligned} \tag{3.3.2}$$

This difference is $O(z^{-2i-1})$, so the asymptotic expansion of $\frac{f_1}{f_0}$ coincides with the asymptotic expansion of $\frac{q_i(z)}{p_i(z)}$ up to the coefficient of $\frac{1}{z^{2i}}$ (cf. [52], p.167).

Suppose an asymptotic expansion $\sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}}$ is given, and that the determinants

$$D_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} \tag{3.3.3}$$

are positive. Construct a rational function (cf. [52], p.168)

$$\frac{\mu_0 z^{2n} + \mu_1 z^{2n-1} + \mu_2 z^{2n-2} + \ldots + \mu_{2n-1} z + t}{z^{2n+1}} = \sum_{k=0}^{2n-1} \frac{\mu_k}{z^{k+1}} + \frac{t}{z^{2n+1}}. \tag{3.3.4}$$

The numerator and denominator polynomials are

$$f_1 = \mu_0 z^{2n} + \mu_1 z^{2n-1} + \mu_2 z^{2n-2} + \ldots + \mu_{2n-1} z + t$$

$$f_0 = 1 z^{2n+1}.$$

Using the determinant condition it is sufficient for this rational function to have a Jacobi fraction expansion that

$$\Delta_0 = 1 > 0$$

$$\Delta_1 = \mu_0 > 0$$

$$\Delta_2 = \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 \\ 1 & 0 & 0 \\ 0 & \mu_0 & \mu_1 \end{vmatrix} > 0$$

$$\Delta_3 = \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \mu_0 & \mu_1 & \mu_2 & \mu_3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \mu_0 & \mu_1 & \mu_2 \end{vmatrix} > 0$$

$$\vdots$$

$$\Delta_{n+1} = \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \dots & t \\ 1 & 0 & 0 & \dots & 0 \\ 0 & \mu_0 & \mu_1 & \dots & \mu_{2n-1} \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \mu_0 & \dots & \mu_n \end{vmatrix} > 0$$

Expanding Δ_2 down the first column gives

$$\Delta_2 = (-1) \begin{vmatrix} \mu_1 & \mu_2 \\ \mu_0 & \mu_1 \end{vmatrix} = D_1.$$

Similarly

$$\Delta_3 = (-1) \begin{vmatrix} \mu_2 & \mu_3 & \mu_4 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_0 & \mu_1 & \mu_2 \end{vmatrix} = D_2.$$

In general for $k < n+1$, $\Delta_k = D_{k-1} > 0$. $\Delta_{n+1} = tD_{n-1} + c$ where c is a real constant. For large enough t, $\Delta_{n+1} > 0$. The rational function $\frac{f_1}{f_0}$ satisfies the conditions for a Jacobi continued fraction expansion and by (3.3.4) the asymptotic expansion of $\frac{f_1}{f_0}$ coincides with the aymptotic expansion $\sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}}$ up to the coefficient of $\frac{1}{z^{2n}}$. Suppose a rational function $\frac{f_1}{f_0}$ coincides with the asymptotic expansion up to the coefficient of $\frac{1}{z^{2l}}$ and a rational function $\frac{f_1'}{f_0'}$ coincides with the asymptotic expansion up to the coefficient of $\frac{1}{z^{2m}}$ where $m > l$. Let $\frac{q_l(z)}{p_l(z)}$ and $\frac{q_l'(z)}{p_l'(z)}$ be the l^{th} convergents of $\frac{f_1}{f_0}$ and $\frac{f_1'}{f_0'}$. Then since the asymptotic expansion of

$$\frac{q_l(z)}{p_l(z)} - \frac{q_l'(z)}{p_l'(z)},$$

starts at $\frac{a_{2l}}{z^{2l+1}}$, the asymptotic expansion of

$$q_l(z)p_l'(z) - q_l'(z)p_l(z),$$

must start at $\frac{b_0}{z}$. So

$$q_l(z)p_l(z)' - q_l'p_l(z) = 0 + R_1(z) \tag{3.3.5}$$

or $q_l(z)p_l'(z) - q_l'(z)p_l(z) = O(z^{-1})$. Because $q_l(z)p_l'(z) - q_l'(z)p_l(z)$ is a polynomial it must be zero to be $O(z^{-1})$(cf [52],p.169). This shows that there is a unique formal infinite Jacobi continued fraction associated with an asymptotic expansion satisfying (3.3.3). Let $\{\mu_n\}_{n=0}^{\infty}$ be a sequence of moments satisfying (3.3.3) then this sequence generates a positive moment functional with associated orthogonal polynomials. Let $p_n(z)$ be the n^{th} monic orthogonal polynomial associated with the moment functional μ generated by the moment sequence. The coefficient of $\frac{1}{z^k}$ in the product $p_n(z)\sum_{k=0}^{\infty}\frac{\mu_k}{z^{k+1}}$ is $\sum_{i=0}^{n}\mu_{k-1+i}c_i$ where c_i is the coefficient of z^i in $p_n(z)$. Because $\mu(z^{k-1}p_n(z))= \sum_{i=0}^{n}\mu_{k-1+i}c_i$, the coefficient of $\frac{1}{z^k}$ for $k = 0, 1, \ldots, n$ in the above product is 0 (cf. [52], p.196). The product is a polynomial $q_n(z)$ added to an asymptotic expansion starting with the term with $\frac{1}{z^{n+1}}$, so $p_n(z)\sum_{k=0}^{\infty}\frac{\mu_k}{z^{k+1}} - q_n(z) = \sum_{k=n}^{\infty}\frac{a_k}{z^{k+1}}$.

The rational function $\frac{q_n(z)}{p_n(z)}$ coincides with $\sum_{k=0}^{\infty}\frac{\mu_k}{z^{k+1}}$ up to the term with $\frac{1}{z^{2n}}$. Recalling (3.3.2), the n^{th} approximant of the Jacobi continued fraction generated by this asymptotic expansion also conincides with the expansion up to the term with $\frac{1}{z^{2n}}$. The difference between $\frac{q_n(z)}{p_n(z)}$ and this n^{th} approximant is $O(z^{-2n-1})$. This gives rise to the identity (3.3.5) with l replaced with n and $q_l'(z)$ and $p_l'(z)$ replaced with the numerator and denominator of the

n^{th} approximant. It follows that the function $\dfrac{q_n(z)}{p_n(z)}$ is identical with the n^{th} approximant generated by the Jacobi continued fraction. It is immediate that the Jacobi continued fraction generated by the asymptotic expansion is the same as the Jacobi continued fraction generated by the three-term recurrence relation for the orthogonal polynomials associated with the moment sequence (cf. [52], pp.165-167,197).

Theorem 3.3.4 (cf. [52], p.247)**.** *If $\alpha(x)$ is a distribution function then the integral*

$$F(z) = \int_{-\infty}^{\infty} \frac{d\alpha(x)}{z - x}, \tag{3.3.6}$$

represents a function which is analytic for z in the upper half plane.

There is a canonical inversion formula for retrieving the distibution function $\alpha(x)$ from the function $F(x)$.

Theorem 3.3.5 (cf. [52], p.250)**.**

$$\frac{1}{\pi} \lim_{y \to 0} \int_s^t \Im[F(x + iy)]dx = \frac{\alpha(s-0) + \alpha(s+0)}{2} - \frac{\alpha(t-0) + \alpha(t+0)}{2}.$$

where $\alpha(s+0) = \lim_{x \to 0} \alpha(s+x)$, $\alpha(s-0) = \lim_{x \to 0} \alpha(s-x)$ and similarly for t.

The integrand of (3.3.6) can be expanded using

$$\frac{1}{z-x} = \frac{1}{z} + \frac{x}{z(z-x)},$$

so in general

$$\frac{1}{z-x} = \frac{1}{z} + \frac{x}{z^2} + \ldots + \frac{x^{n-1}}{z^n} + \frac{x^n}{z^n(z-x)}.$$

This leads to the expansion

$$\int_{-\infty}^{\infty} \frac{d\alpha(x)}{z-x} = \frac{\mu_0}{z} + \frac{\mu_1}{z^2} + \ldots + \frac{\mu_{n-1}}{z^n} + \int_{-\infty}^{\infty} \frac{x^n d\alpha(x)}{z^n(z-x)}. \tag{3.3.7}$$

It can be shown (cf. [52], pp.322-324) that (3.3.7) generates an asymptotic expansion for $\displaystyle\int_{-\infty}^{\infty} \frac{d\alpha(x)}{z-x}$ in any half-plane $\Im(z) \geq \delta > 0$. Any convergent subsequence of the convergents of the Jacobi continued fraction generated by this asymptotic expansion will have the same asymptotic expansion (cf. [52], p.316). It is known from prior results that the condition that the determinants $D_p > 0$ is necessary and sufficient for a solution to the moment problem to exist.

Theorem 3.3.6 (cf. [52], pp.231,324)**.** *If $\alpha(x)$ is a solution of a moment problem with moments $\{\mu_n\}_{n=0}^{\infty}$ it can be recovered from a function*

$$F(z)=\int_{-\infty}^{\infty} \frac{d\alpha(x)}{z-x} \tag{3.3.8}$$

and the limit of each convergent subsequence of the convergents of the Jacobi continued fraction generated by the asymptotic expansion of (3.3.8) is one of the functions $F(z)$.

This leaves open the question of whether there are other functions $F(z)$ and how to obtain them.

3.4 Limit circle and limit point

Throughout this section it will be assumed that $\Im(z) \neq 0$.

Let the Jacobi continued fraction

$$\cfrac{1}{z-e_0-\cfrac{d_0^2}{\ddots - \cfrac{d_n^2}{z-e_{n+1}-\ddots}}}, \tag{3.4.1}$$

be given.

Denote by $q_n(z)$ and $p_n(z)$ the n^{th} numerator and denominator of the convergents of this continued fraction. Then $q_n(z)$ and $p_n(z)$ are the monic orthogonal polynomials generated by the Jacobi matrices (3.2.1) and (3.2.3).

Definition 3.4.1 (cf. [45], p.33)**.** *The truncated continued fraction*

$$\cfrac{1}{z - e_0 - \cfrac{d_0^2}{\ddots - \cfrac{d_n^2}{z - e_{n+1} - \cfrac{1}{\tau}}}}, \tag{3.4.2}$$

is called the generalized approximant of the continued fraction (3.4.1).

By Lemma 3.1.4, the truncated continued fraction in (3.4.2) reduces to

$$\frac{q_n(z) - \frac{1}{\tau}q_{n-1}(z)}{p_n(z) - \frac{1}{\tau}p_{n-1}(z)} = \frac{q_n(z)\tau - q_{n-1}(z)}{p_n(z)\tau - p_{n-1}(z)}.$$

In the equivalence transformation of Theorem 3.1.6 let $c_1 = 1$, $c_2 = \dfrac{1}{d_0}$, $c_3 = \dfrac{1}{d_1} \ldots$, then the transformed continued fraction of (3.4.1) has $\dfrac{q_n(z)}{d_0 d_1 d_2 \ldots d_{n-1}}$ as its n^{th} numerator and $\dfrac{p_n(z)}{d_0 d_1 d_2 \ldots d_{n-1}}$ as its n^{th} denominator. Referring to (2.6.8) the denominator is the n^{th} orthonormal polynomial $P_n(z)$ generated by the Jacobi matrix associated with the continued fraction. The transformed n^{th} numerator will be denoted by $Q_n(z)$. The corresponding truncated continued fraction (3.4.2) under the equivalence transformation reduces to

$$\frac{Q_n(z)\tau - Q_{n-1}(z)}{P_n(z)\tau - P_{n-1}(z)}.$$

Substituting $P_n(z)$ for $p_n(z)$ and $Q_n(z)$ for $q_n(z)$ in (3.3.1) gives

$$\frac{Q_n(z)}{P_n(z)} - \frac{Q_{n-1}(z)}{P_{n-1}(z)} = \frac{1}{d_{n-1}P_n(z)P_{n-1}(z)},$$

or

$$Q_n(z)P_{n-1}(z) - Q_{n-1}(z)P_n(z) = \frac{1}{d_{n-1}}. \tag{3.4.3}$$

The three-term recurrence relation for the orthonormal polynomials is

$$zP_n(z) = d_nP_{n+1}(z) + e_nP_n(z) + d_{n-1}P_{n-1}(z).$$

This is a particular case of the recurrence relation

$$\mu z_k = d_k z_{k+1} + e_k z_k + d_{k-1} z_{k-1}. \tag{3.4.4}$$

This recurrence relation admits a formula analogous to the Christoffel-Darboux formula.

Lemma 3.4.2 (cf. [1], p.9)**.** *Let z_k be a solution of (3.4.4) with the parameter μ and y_k be a solution of (3.4.4) with parameter λ, then*

$$(\mu - \lambda) \sum_{k=m}^{n-1} y_k z_k = d_{n-1}(y_{n-1} z_n - y_n z_{n-1}) - d_{m-1}(y_{m-1} z_m - y_m z_{m-1}), \tag{3.4.5}$$

holds.

Proof. Multiply (3.4.4) by y_k to obtain

$$\mu z_k y_k = d_k z_{k+1} y_k + e_k z_k y_k + d_{k-1} z_{k-1} y_k. \tag{3.4.6}$$

Similarly multiply the relation

$$\lambda y_k = d_k y_{k+1} + e_k y_k + d_{k-1} y_{k-1}$$

by z_k to obtain

$$\lambda y_k z_k = d_k y_{k+1} z_k + e_k y_k z_k + d_{k-1} y_{k-1} z_k. \tag{3.4.7}$$

Subtracting (3.4.7) from (3.4.6) gives

$$(\mu - \lambda) z_k y_k = d_k (y_k z_{k+1} - y_{k+1} z_k) - d_{k-1}(y_{k-1} z_k - y_k z_{k-1}). \tag{3.4.8}$$

The sum is telescoping and the result is obtained. □

Recall that it is assumed that $\mu_0 = 1$. Under this assumption the initial conditions for the orthonormal polynomials are given by

$$P_{-1}(z) = 0, P_0(z) = 1, P_1(z) = \frac{z - e_0}{d_0}.$$

The parameters for the difference equation (3.4.4) are real so if $P_n(z)$ is a solution of (3.4.4) then $\overline{P_n(z)} = P_n(\overline{z})$ is also a solution. For $m = 1$ and substituting $P_k(z)$ for z_k, $\overline{P_k(z)}$ for y_k, z for μ and $\overline{z}$ for λ in (3.4.5) gives

$$(z-\overline{z})\sum_{k=1}^{n-1}|P_k(z)|^2 = d_{n-1}(P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)})$$

$$-d_0(P_1(z)\overline{P_0(z)} - P_0(z)\overline{P_1(z)}).$$

The last term on the right is

$$-d_0\left(\frac{z-e_0}{d_0} - \frac{\overline{z}-e_0}{d_0}\right) = (-1)(z-\overline{z}) = -(z-\overline{z})|P_0(z)|^2.$$

So that

$$(z-\overline{z})\sum_{k=0}^{n-1}|P_k(z)|^2 = d_{n-1}(P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)}). \tag{3.4.9}$$

Theorem 3.4.3 (cf. [1], p.11)**.** *For z fixed the function*

$$w_n(z,\tau) = \frac{Q_n(z)\tau - Q_{n-1}(z)}{P_n(z)\tau - P_{n-1}(z)},\ n = 0,1,2,3\ldots$$

maps the real line onto a circle in the complex plane with centre

$$\frac{Q_n(z)\overline{P_{n-1}(z)} - Q_{n-1}(z)\overline{P_n(z)}}{P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)}}$$

and radius

$$\frac{1}{|z-\overline{z}|\sum_{k=0}^{n-1}|P_k(z)|^2}$$

Proof. For fixed z, $w_n(z,\tau)$ is a Möbius transformation so the image of the real axis is a generalised circle in the complex plane. It suffices to compute the radius and centre.

$$\begin{aligned}
&(Q_n(z)\overline{P_{n-1}(z)} - Q_{n-1}(z)\overline{P_n(z)})(P_n(z)\tau - P_{n-1}(z))\\
&-(Q_n(z)P_{n-1}(z) - Q_{n-1}(z)P_n(z))(\overline{P_n(z)}\tau - \overline{P_{n-1}(z)})\\
=&Q_n(z)\overline{P_{n-1}(z)}P_n(z)\tau - Q_n(z)\overline{P_{n-1}(z)}P_{n-1}(z)
\end{aligned}$$

$$
\begin{aligned}
&- Q_{n-1}(z)P_n(z)\overline{P_n(z)}\tau + Q_{n-1}(z)\overline{P_n(z)}P_{n-1}(z)\\
&- Q_n(z)P_{n-1}(z)\overline{P_n(z)}\tau + Q_n(z)P_{n-1}(z)\overline{P_{n-1}(z)}\\
&+ Q_{n-1}(z)P_n(z)\overline{P_n(z)}\tau - Q_{n-1}(z)P_n(z)\overline{P_{n-1}(z)}\\
=&(Q_n(z)\tau - Q_{n-1}(z))(P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)})
\end{aligned}
$$

The above calculations show that

$$
\begin{aligned}
\frac{Q_n(z)\tau - Q_{n-1}(z)}{P_n(z)\tau - P_{n-1}(z)} =& \frac{Q_n(z)\overline{P_{n-1}(z)} - Q_{n-1}(z)\overline{P_n(z)}}{P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)}}\\
&- \frac{Q_n(z)P_{n-1}(z) - Q_{n-1}(z)P_n(z)}{P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)}}\,\frac{\overline{P_n(z)}\tau - \overline{P_{n-1}(z)}}{P_n(z)\tau - P_{n-1}(z)}.
\end{aligned}
$$

Notice that

$$
\left|\frac{\overline{P_n(z)}\tau - \overline{P_{n-1}(z)}}{P_n(z)\tau - P_{n-1}(z)}\right| = 1
$$

and the argument of the last term on the right is a function of τ only, so the centre is as stipulated and the radius is given by

$$
\left|\frac{Q_n(z)P_{n-1}(z) - Q_{n-1}(z)P_n(z)}{P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)}}\right|.
$$

It has been shown in (3.4.3) that $Q_n(z)P_{n-1}(z) - Q_{n-1}(z)P_n(z) = \dfrac{1}{d_{n-1}}$. By (3.4.9)

$$
\frac{1}{d_{n-1}(P_n(z)\overline{P_{n-1}(z)} - P_{n-1}(z)\overline{P_n(z)})} = \frac{1}{(z-\overline{z})\sum_{k=0}^{n-1}|P_k(z)|^2},
$$

so that the radius is as required. □

The circle $w_{n+1}(z,\tau)$ is contained in the circle $w_n(z,\tau)$ (cf. [1], p.13). They have a common point because (cf. [1], p.14)

$$
w_{n+1}(z,0) = w_n(z,\infty) = \frac{Q_n(z)}{P_n(z)}.
$$

$K_n(z)$ will be used to denote the circumference and interior of the circle $w_n(z,\tau)$.

From the formula for the radius for fixed z, if

$$\sum_{k=0}^{\infty} |P_k(z)|^2 = \infty,$$

then the circles reduce to a point in the limit, while if

$$\sum_{k=0}^{\infty} |P_k(z)|^2 < \infty$$

then the limit of the circles is a circle. These two cases will be referred to as the limit point case and the limit circle case. The limit circle corresponding to z will be denoted by $K_\infty(z)$.

Theorem 3.4.4 (cf. [1], pp.34, 41)**.** *The limit circle case corresponds to an indeterminate moment problem and the limit point case to a determinate moment problem.*

3.5 The Nevanlinna parametrisation

Theorem 3.5.1 (cf. [1], pp.16-19)**.** *If the series* $\sum_{k=0}^{\infty} |P_k(z)|^2$ *converges for any* z *then it converges uniformly on compact subsets of the complex plane. The same holds for the series* $\sum_{k=0}^{\infty} |Q_k(z)|^2$*.*

As a consequence of this theorem if the limit circle case holds for a single point then it holds for all points z with $\Im(z) \neq 0$. In this section it will be assumed that the limit circle case holds.

Define four polynomials $A_n(z)$, $B_n(z)$, $C_n(z)$ and $D_n(z)$ by

$$\begin{aligned}
A_n(z) &= d_{n-1}\{Q_{n-1}(0)Q_n(z) - Q_n(0)Q_{n-1}(z)\}, \\
B_n(z) &= d_{n-1}\{Q_{n-1}(0)P_n(z) - Q_n(0)P_{n-1}(z)\}, \\
C_n(z) &= d_{n-1}\{P_{n-1}(0)Q_n(z) - P_n(0)Q_{n-1}(z)\}, \\
D_n(z) &= d_{n-1}\{P_{n-1}(0)P_n(z) - P_n(0)P_{n-1}(z)\}.
\end{aligned}$$

Lemma 3.5.2. *The function $w_n(0,u)$ where u is any complex number defines a Möbius transformation which maps the half plane $\Im(u) \geq 0$ onto the half plane $\Im(w_n(0,u)) \leq 0$.*

Proof. $P_n(0)$, $P_{n-1}(0)$, $Q_n(0)$ and $Q_{n-1}(0)$ are real numbers so the Möbius transformation

$$\frac{Q_n(0)u - Q_{n-1}(0)}{P_n(0)u - P_{n-1}(0)}$$

leaves the real line invariant. It is sufficient to show that an arbitrary point in the half plane $\Im(u) > 0$ is mapped to the half plane $\Im(w_n(0,u)) < 0$.

$$\begin{aligned} w_n(0,i) &= \frac{Q_n(0)i - Q_{n-1}(0)}{P_n(0)i - P_{n-1}(0)} \\ &= \frac{(Q_n(0)P_n(0) + Q_{n-1}(0)P_{n-1}(0)) - i(Q_n(0)P_{n-1}(0) - P_n(0)Q_{n-1}(0))}{(P_n^2(0) + P_{n-1}^2(0))} \\ &= \frac{(Q_n(0)P_n(0) + Q_{n-1}(0)P_{n-1}(0)) - i(\frac{1}{d_{n-1}})}{(P_n^2(0) + P_{n-1}^2(0))}. \end{aligned}$$

So $\Im(w_n(0,i)) = -\dfrac{1}{(P_n^2(0) + P_{n-1}^2(0))d_{n-1}} < 0.$ □

Lemma 3.5.3 (cf. [1], p.15)**.** *The function,*

$$w_n(z,\tau) = \frac{Q_n(z)\tau - Q_{n-1}(z)}{P_n(z)\tau - P_{n-1}(z)}$$

can be expressed as

$$w_n(z,\tau) = \frac{C_n(z)t - A_n(z)}{D_n(z)t - B_n(z)},$$

where t is a real number depending on τ.

Proof.

$$-Q_{n-1}(0)P_n(0) + Q_n(0)P_{n-1}(0) = \frac{1}{d_{n-1}} \neq 0,$$

so $A_n(z)$ and $C_n(z)$ are linearly independent combinations of $Q_n(z)$ and $Q_{n-1}(z)$. As a result any linear combination of $Q_n(z)$ and $Q_{n-1}(z)$ can be obtained using a linear combination of $A_n(z)$ and $C_n(z)$. In particular the linear combination $Q_n(z)\tau - Q_{n-1}(z)$ can be obtained in this way, i.e. $a_1C_n(z) - a_2A_n(z) = Q_n(z)\tau - Q_{n-1}(z)$. The same holds for $B_n(z)$ and $D_n(z)$

in relation to $P_n(z)\tau - P_{n-1}(z)$. Furthermore because $A_n(z)$ is the same linear combination of $Q_n(z)$ and $Q_{n-1}(z)$ (in terms of coefficients) as $B_n(z)$ is of $P_n(z)$ and $P_{n-1}(z)$ and this relationship exists between $C_n(z)$ and $D_n(z)$ also, $a_1 D_n(z) - a_2 B_n(z) = P_n(z)\tau - P_{n-1}(z)$. Let $t = \dfrac{a_1}{a_2}$, then

$$\begin{aligned}\frac{Q_n(z)\tau - Q_{n-1}(z)}{P_n(z)\tau - P_{n-1}(z)} &= \frac{a_1 C_n(z) - a_2 A_n(z)}{a_1 D_n(z) - a_2 B_n(z)} \\ &= \frac{C_n(z)t - A_n(z)}{D_n(z)t - B_n(z)}.\end{aligned} \qquad \square$$

The real number t can be calculated explicitly as follows. It is sufficient to equate the coefficients of $Q_k(z)$ and $Q_{k-1}(z)$ in $C_k(z)t - A_k(z)$ and $c(Q_k(z)\tau - Q_{k-1}(z))$.

$$\begin{aligned}c(Q_k(z)\tau - Q_{k-1}(z)) &= t d_{n-1}(P_{k-1}(0)Q_k(z) \\ &\qquad - P_k(0)Q_{k-1}(z)) - d_{n-1}(Q_{k-1}(0)Q_k(z) + Q_k(0)Q_{k-1}(z)) \\ c\tau &= d_{n-1}(tP_{k-1}(0) - Q_{k-1}(0)) \\ c &= d_{n-1}(tP_k(0) - Q_k(0)) \\ (tP_k(0) - Q_k(0))\tau &= tP_{k-1}(0) - Q_{k-1}(0) \\ t(P_k(0)\tau - P_{k-1}(0)) &= Q_k(0)\tau - Q_{k-1}(0) \\ t &= \frac{Q_k(0)\tau - Q_{k-1}(0)}{P_k(0)\tau - P_{k-1}(0)} \\ &= w_n(0,\tau).\end{aligned}$$

Using the three-term recurrence formula for the orthonormal polynomials gives (cf. [1], p.14)

$$\begin{aligned}A_{n+1}(z) &= d_n\{Q_n(0)Q_{n+1}(z) - Q_{n+1}(0)Q_n(z)\} \\ &= Q_n(0)\{(z - e_n)Q_n(z) - d_{n-1}Q_{n-1}(z)\} \\ &\quad - Q_n(z)\{(-e_n)Q_n(0) - d_{n-1}Q_{n-1}(0)\} \\ &= d_{n-1}\{Q_{n-1}(0)Q_n(z) + Q_n(0)Q_{n-1}(z)\} + zQ_n(0)Q_n(z) \\ &= A_n(z) + zQ_n(0)Q_n(z).\end{aligned}$$

So $A_{n+1}(z) - A_n(z) = zQ_n(0)Q_n(z)$. Similar calculations show that

$$B_{n+1}(z) - B_n(z) = zQ_n(0)P_n(z)$$

$$C_{n+1}(z) - C_n(z) = zP_n(0)Q_n(z)$$

$$D_{n+1}(z) - D_n(z) = zP_n(0)P_n(z).$$

Because $Q_0(z) = 0$, $A_1(z) = 0$, and $zQ_0(0)Q_0(z) = 0$,

$$A_n(z) = \sum_{k=1}^{n-1} A_{k+1}(z) - A_k(z) = z\sum_{k=0}^{n-1} Q_k(0)Q_k(z).$$

$Q_1(z) = \dfrac{q_1(z)}{d_0} = \dfrac{1}{d_0}$, so $B_1(z) = -1$, and

$$B_n(z) = -1 + z\sum_{k=0}^{n-1} Q_k(0)P_k(z).$$

$C_1(z) = 1$ so

$$C_n(z) = 1 + z\sum_{k=0}^{n-1} P_k(0)Q_k(z).$$

$D_1(z) = 0$ so

$$D_n(z) = z\sum_{k=0}^{n-1} P_k(0)P_k(z).$$

$A(z)$ will be used to denote $\lim_{n\to\infty} A_n(z)$ and similarly for $B(z)$, $C(z)$, $D(z)$.

Theorem 3.5.4 (cf. [52], p.101)**.** *The functions $A(z)$, $B(z)$, $C(z)$, $D(z)$ are entire.*

Proof.

$$A(z) = \lim_{n\to\infty} A_n(z) = z\sum_{k=0}^{\infty} Q_k(0)Q_k(z).$$

By uniform convergence of $\sum_{k=0}^{\infty} |Q_k(z)|^2$ on compact subsets of the complex plane, for $|z| < M$ for an arbitrary $0 < \epsilon < 1$ and for $i = 0, 1, 2, \ldots$ there is a natural number N such that for $m > N$

$$\sum_{k=m}^{m+i} |Q_k(z)|^2 < \frac{\epsilon}{M}.$$

Under these conditions $\sum_{k=m}^{m+i} |Q_k(0)|^2 < \frac{\epsilon}{M}$. For $|z| < M$, Schwartz's inequality gives

$$|z| \sum_{k=m}^{m+i} |Q_k(0) Q_k(z)| \leq M \sqrt{\sum_{k=m}^{m+i} |Q_k(0)|^2 \sum_{k=m}^{m+i} |Q_k(z)|^2} < \epsilon.$$

Similar reasoning for $B(z)$, $C(z)$ and $D(z)$ completes the proof. □

For z fixed the Möbius transformation $w_n(z, u)$ maps the half plane $\Im(u) \geq 0$ into the interior of the circle $K_n(z)$ (cf. [52], pp.66, 71). Recall

$$w_n(z, u) = \frac{C_n(z) w_n(0, u) - A_n(z)}{D_n(z) w_n(0, u) - B_n(z)}.$$

Because $w_n(0, u)$ maps the half plane $\Im(u) \geq 0$ onto the half plane $\Im(w_n(0, u)) \leq 0$, if $w_n(0, u)$ is replaced by any function $\theta(z)$ analytic for $\Im(z) > 0$ and for $\Im(z) > 0$ satisfying $\Im(\theta(z)) \leq 0$ then the function

$$w_n(z) = \frac{C_n(z)\theta(z) - A_n(z)}{D_n(z)\theta(z) - B_n(z)}$$

will have its values in $K_n(z)$ for $\Im(z) > 0$ and for each n (cf. [52], p.320). Let

$$F(z) = \lim_{n \to \infty} w_n(z).$$

Then

$$F(z) = \frac{C(z)\theta(z) - A(z)}{D(z)\theta(z) - B(z)} \tag{3.5.1}$$

(cf. [52], p.320).

In the limit circle case, a function is called an equivalent function of a Jacobi continued fraction if it is analytic for $\Im(z) > 0$ and for $\Im(z) > 0$ takes all of its values inside the limit circles $K_\infty(z)$ (cf. [52], p.231). Every equivalent function of a Jacobi fraction can be represented by the formula (3.5.1). Furthermore in the limit circle case every solution of the moment problem corresponding to a Jacobi continued fraction corresponds to an equivalent function of the Jacobi continued fraction (cf. [52], pp.324, 326).

Theorem 3.5.5 (**The Nevanlinna Parametrisation** cf. [52], pp.321, 326)**.** *If $\theta(z)$ is analytic in the upper half-plane and satisfies $\Im(\theta(z)) \leq 0$ for $\Im(z) > 0$, or if $\theta(z) \equiv \infty$, then the function*

$$F(z) = \frac{C(z)\theta(z) - A(z)}{D(z)\theta(z) - B(z)},$$

is one of the functions $F(z)$ mentioned in Theorem 3.3.6. The function corresponds to the Jacobi continued fraction generated by the moment sequence $\{\mu_n\}_{n=0}^{\infty}$ via an asymptotic series. Each of these functions can be represented by the formula

$$F(z) = \int_{-\infty}^{\infty} \frac{d\alpha(x)}{z - x},$$

where $\alpha(x)$ is a solution of the moment problem with moments $\{\mu_n\}_{n=0}^{\infty}$. The solution $\alpha(x)$ corresponding to the function $F(z)$ can be recovered with the Stieltjes inversion formula. Every solution to an indeterminate moment problem corresponds to one of these functions $F(z)$.

The Nevanlinna parametrisation gives a complete characterisation of solutions to an indeterminate Hamburger moment problem.

Chapter 4

Symmetric moment problems

A special case of the moment problem arises when the odd moments are zero. In the classical cases this is a consequence of the weight function being even. In general this situation is called a symmetric moment problem. Every symmetric Hamburger moment problem generates a Stieltjes moment problem and every Stieltjes moment problem generates a symmetric Hamburger moment problem. This connection is presented and used to obtain a connection between the orthogonal polynomials generated by the Hamburger moment problem and the Stieltjes moment problem respectively. Symmetric moment problems give rise to chain sequences. These are special numerical sequences generated by the parameters of the three-term recurrence relation. To conclude the chapter basic results on chain sequences are given.

4.1 Symmetric distributions

Definition 4.1.1 (cf. [14], p.332)**.** *A Hamburger moment problem* $\{\mu_n\}_{n=0}^{\infty}$ *such that the odd moments* μ_{2n+1} *are zero is called a symmetric moment problem.*

It will be assumed that all symmetric moment problems considered here have at least one solution.

Definition 4.1.2 (cf. [14], p.332)**.** *A solution $\alpha(x)$ of a moment problem is called symmetric if $\alpha(-x) + \alpha(x) = C$, C a real constant.*

Lemma 4.1.3. *Let $\alpha(x)$ be a distribution, $a, b > 0$ and let $f(x)$ be a continuous function, then*

$$\int_{-a}^{b} f(-x) d\alpha(-x) = -\int_{-b}^{a} f(x) d\alpha(x). \tag{4.1.1}$$

Proof. Let $-b < t_1 < t_2 < \ldots < t_n = a$, be a partition of $(-b, a)$. Let $-a < x_1 < x_2 < \ldots < x_n = b$ be a partition of $(-a, b)$ such that $x_k = -t_{n+1-k}$ for each k. For v_k chosen in $[x_k, x_{k+1}]$, $-v_k$ is in $[t_{n-k}, t_{n+1-k}]$. Let $u_{n-k} = -v_k$.

$$\begin{aligned}\sum_{k=1}^{n-1} f(-v_k)\{\alpha(-x_{k+1}) - \alpha(-x_k)\} &= \sum_{k=1}^{n-1} f(u_{n-k})\{\alpha(t_{n-k}) - \alpha(t_{n+1-k})\} \\ &= (-1)\sum_{k=1}^{n-1} f(u_{n-k})\{\alpha(t_{n+1-k}) - \alpha(t_{n-k})\} \\ &= (-1)\sum_{k=1}^{n-1} f(u_k)\{\alpha(t_{k+1}) - \alpha(t_k)\},\end{aligned}$$

where the order of summation was reversed for the last step. Constructing all partitions in this way and taking the limit as the partitions become finer gives the required identity. □

Theorem 4.1.4 (cf. [14], p.332)**.** *Suppose that $\alpha_1(x)$ is a distribution function with moments $\{\mu_n\}_{n=0}^{\infty}$, $\mu_{2n+1} = 0$ for each n. Then*

$$\alpha_2(x) = \frac{1}{2}[\alpha_1(x) - \alpha_1(-x)]$$

is a distribution function with the same moments as $\alpha_1(x)$ and is a symmetric solution of the moment problem.

Proof. In (4.1.1) let $a = b$, then

$$\begin{aligned}\int_{-b}^{b} (-x)^n d\alpha_1(-x) &= -\int_{-b}^{b} x^n d\alpha_1(x) \\ (-1)^{n+1}\int_{-b}^{b} x^n d\alpha_1(-x) &= \int_{-b}^{b} x^n d\alpha_1(x)\end{aligned}$$

As $b \to \infty$

$$(-1)^{n+1} \int_{-\infty}^{\infty} x^n d\alpha_1(-x) = \int_{-\infty}^{\infty} x^n d\alpha_1(x).$$

Using Lemma 2.1.10 the even moments of $-\alpha_1(-x)$ are the same as the even moments of $\alpha_1(x)$. By hypothesis the odd moments are zero, so in fact all moments are equal. Again using Lemma 2.1.10 and the assumption about the odd moments, the moments of

$$\alpha_2(x) = \frac{1}{2}[\alpha_1(x) - \alpha_1(-x)]$$

are the same as the moments of $\alpha_1(x)$ and $\alpha_2(x) + \alpha_2(-x) = 0$. $\square$

Let $\alpha(x)$ be a symmetric solution of a symmetric moment problem, then it generates a set of orthogonal polynomials which will be denoted by $S_n(x)$.

Lemma 4.1.5 (cf. [14], p. 332)**.** *The polynomials $\{S_n(x)\}_{n=0}^{\infty}$ satisfy $S_n(x) = (-1)^n S_n(-x)$.*

Proof. By assumption $\alpha(x) + \alpha(-x) = C$, $C \in \mathbb{R}$. Let $\alpha_2(x) = \alpha(x) + \alpha(-x) \equiv C$, and k_n be the leading coefficient of $S_n(x)$. Because $\alpha_2(x)$ is a constant function, the integral of any continuous function with respect to $\alpha_2(x)$ is zero. In particular,

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} S_n(-x)(-x)^m d\alpha_2(x) \\ &= \int_{-\infty}^{\infty} S_n(-x)(-x)^m d\alpha(x) + \int_{-\infty}^{\infty} S_n(-x)(-x)^m d\alpha(-x) \\ &= \int_{-\infty}^{\infty} S_n(-x)(-x)^m d\alpha(x) - \int_{-\infty}^{\infty} S_n(x)(x)^m d\alpha(x). \end{aligned}$$

Thus

$$\int_{-\infty}^{\infty} (-1)^m S_n(-x) x^m d\alpha(x) = \int_{-\infty}^{\infty} S_n(x) x^m d\alpha(x) = 0 \quad m < n.$$

It follows that $\{S_n(-x)\}_{n=0}^{\infty}$ is a set of orthogonal polynomials corresponding to $\alpha(x)$. Such a set is determined up to constant multiples, so to establish the result it is sufficient to note that the leading coefficient of $S_n(-x)$ is $(-1)^n k_n$. $\square$

This result implies that $S_{2n}(x)$ is a linear combination of even powers of x and $S_{2n+1}(x)$ is a linear combination of odd powers of x.

Throughout the rest of this chapter it will be assumed that the polynomials $S_n(x)$ are monic, i.e. $k_n = 1$. The function $\alpha(x)$ is non-decreasing and bounded on $(-\infty, \infty)$, so in particular it is non-decreasing and bounded on $[0, \infty)$. $\sqrt{x}$ is non-decreasing on $[0, \infty)$, so $\alpha(\sqrt{x})$ will be non-decreasing on $[0, \infty)$, bounded above by $\alpha(\infty) < \infty$ and below by $\alpha(0) > -\infty$.

Lemma 4.1.6. *Let $f(x)$ be a continuous function on $[0, b]$, $b > 0$. Let $\alpha(x)$ and $\psi(x)$ be distributions such that $\psi(x) = \alpha(\sqrt{x})$.*

$$\int_0^{b^{\frac{1}{2}}} f(x^2)d\alpha(x) = \int_0^b f(x)d\psi(x). \tag{4.1.2}$$

Proof. Let $0 = x_1 < x_2 < \ldots < x_n = b^{\frac{1}{2}}$ be a partition of $(0, b^{\frac{1}{2}})$. Let $0 = t_1 < t_2 < \ldots < t_n = b$ be a partition of $(0, b)$ such that $t_k = x_k^2$. For v_k in $[x_k, x_{k+1}]$ v_k^2 is in $[t_k, t_{k+1}]$. Let $u_k = v_k^2$.

$$\begin{aligned}
\sum_{k=1}^{n-1} f(v_k^2)\{\alpha(x_{k+1}) - \alpha(x_k)\} &= \sum_{k=1}^{n-1} f(u_k)\{\alpha(\sqrt{t_{k+1}}) - \alpha(\sqrt{t_k})\} \\
&= \sum_{k=1}^{n-1} f(u_k)\{\psi(t_{k+1}) - \psi(t_k)\}.
\end{aligned}$$

If all partitions are constructed in this way, then in the limit the required identity is obtained. □

Lemma 4.1.7. *The moments of $\alpha(x)$ and $\psi(x)$ are related by*

$$\int_{-\infty}^{\infty} x^{2n} d\alpha(x) = 2\int_0^{\infty} x^n d\psi(x).$$

Proof. If the distribution $\alpha(x)$ has a jump at the point 0 then it contributes nothing to the calculation of the moments of $\alpha(x)$ because x^n has a zero at 0 for each n.

$$\int_{-\infty}^{\infty} x^{2n} d\alpha(x) = \int_{-\infty}^{0} x^{2n} d\alpha(x) + \int_0^{\infty} x^{2n} d\alpha(x)$$

$$\begin{aligned}
&= -\int_0^\infty (-x)^{2n} d\alpha(-x) + \int_0^\infty x^{2n} d\alpha(x) \\
&= -\int_0^\infty x^{2n} d[C - \alpha(x)] + \int_0^\infty x^{2n} d\alpha(x) \\
&= 2\int_0^\infty x^{2n} d\alpha(x).
\end{aligned}$$

Letting $b \to \infty$ in (4.1.2) gives

$$\int_{-\infty}^\infty x^{2n} d\alpha(x) = 2\int_0^\infty x^n d\psi(x). \qquad \square$$

This shows that $\psi(x)$ is a distribution. Let $\{P_n(x)\}_{n=0}^\infty$ be the monic polynomials orthogonal with respect to $\psi(x)$.

Theorem 4.1.8 (cf. [14], p.332)**.** *$P_n(x^2) = S_{2n}(x)$ for each n.*

Proof. Using the linearity of the integral,

$$\begin{aligned}
\int_{-\infty}^\infty P_n(x^2)x^{2m} d\alpha(x) &= 2\int_0^\infty P_n(x)x^m d\psi(x) \\
&= 0, \qquad \text{for} \qquad 2m < 2n.
\end{aligned}$$

$$\int_{-\infty}^\infty P_n(x^2)x^{2m+1} d\alpha(x) = 0, \qquad \text{for} \qquad 2m+1 < 2n,$$

because $P_n(x^2)x^{2m+1}$ is a polynomial with only odd powers of x. $\{P_n(x^2)\}$ is a set of monic orthogonal polynomials of degree $2n$ with respect to $\alpha(x)$ so $P_n(x^2)$ must be identical to $S_{2n}(x)$ for each n. $\square$

Because $x > 0$ for $x \in [0, \infty)$, the function $\omega(x)$ given by

$$\omega(x) = \int_0^x x d\psi(x),$$

is non-decreasing. Furthermore, $\omega(x)$ is bounded above by

$$\int_0^\infty x d\psi(x) = \frac{1}{2}\mu_2,$$

and below by 0. Using (4.1.2)

$$\int_{-\infty}^\infty x^{2n+2} d\alpha(x) = 2\int_0^\infty x^n x d\psi(x)$$

$$= 2\int_0^\infty x^n d\omega(x),$$

so $\omega(x)$ is a distribtion. Let $\{K_n(x)\}_{n=0}^\infty$ be the monic polynomials orthogonal with respect to $\omega(x)$.

Theorem 4.1.9 (cf. [14], p. 332)**.** $xK_n(x^2) = S_{2n+1}(x)$.

Proof. Using linearity

$$\int_{-\infty}^\infty xK_n(x^2)x^{2m+1}d\alpha(x) = 2\int_0^\infty K_n(x)x^m d\omega(x)$$
$$= 0, \qquad \text{for} \qquad 2m+1 < 2n+1$$

$$\int_{-\infty}^\infty xK_n(x^2)x^{2m}d\alpha(x) = 0, \qquad \text{for} \qquad 2m < 2n+1,$$

because $x^{2m+1}K_n(x^2)$ is a polynomial with only odd powers of x. $\{xK_n(x^2)\}_{n=0}^\infty$ is a set of monic orthogonal polynomials with respect to $\alpha(x)$ such that $xK_n(x^2)$ has degree $2n+1$ for each n. Consequently $xK_n(x^2) = S_{2n+1}(x)$ for each n. $\square$

The process above can be undertaken in the reverse direction. That is, given a distribution with points of increase contained in $[0,\infty)$, a symmetric distribution with points of increase in $(-\infty,\infty)$ can be constructed. The polynomials $P_n(x)$, $K_n(x)$ and $S_n(x)$ can also be constructed with the same relationship existing between them (cf. [12], pp.1-3).

Let $\psi(x)$ be a distribution with points of increase in $[0,\infty)$. Then x^2 is non-decreasing on $[0,\infty)$ so $\psi(x^2)$ is non-decreasing on $[0,\infty)$. Also, x^2 is non-increasing on $(-\infty,0)$, so $\psi(x^2)$ will be non-increasing on $(-\infty,0)$ and $-\psi(x^2)$ will be non-decreasing on $(-\infty,0)$. It follows that $\alpha(x) = \text{sgn}(x)\psi(x^2)$ is non-decreasing on $(-\infty,\infty)$. $\alpha(x)$ is bounded above by $\psi(\infty) < \infty$ and below by $-\psi(\infty) > -\infty$. $\alpha(x)+\alpha(-x) = 0$ so using the approach in Lemma (4.1.7)

$$\int_{-\infty}^\infty x^{2n}d\alpha(x) = 2\int_0^\infty x^n d\psi(x)$$
$$\int_{-\infty}^\infty x^{2n+1}d\alpha(x) = 0.$$

So $\alpha(x)$ is a symmetric distribution associated with a symmetric moment problem.

4.2 Chain sequences

Definition 4.2.1 (cf. [15], p. 91)**.** *A sequence of real numbers $\{a_n\}_{n=0}^{\infty}$ is a chain sequence if there is another sequence $\{g_n\}_{n=-1}^{\infty}$, such that*

$$0 \le g_{-1} < 1, \quad 0 < g_n < 1, \qquad n \ge 0$$
$$a_n = (1 - g_{n-1})g_n, \qquad n = 0, 1, 2, 3 \dots .$$

The sequence $\{g_n\}_{n=-1}^{\infty}$ is called the parameter sequence corresponding to the given chain sequence.

In some references g_n is allowed to be zero for $n \ge 0$ (cf. [52], p.79). This case does not arise in the present discussion and is excluded from the definition.

It will be shown that there is a connection between Stieltjes moment problems and chain sequences. The first step is the following theorem.

Theorem 4.2.2 (cf. [52], p.67)**.** *Let $\{d_n\}_{n=0}^{\infty}$ and $\{e_n\}_{n=0}^{\infty}$ be real sequences. There exists a real sequence $\{g_n\}_{n=-1}^{\infty}$ such that*

$$d_n^2 = e_n e_{n+1}(1 - g_{n-1})g_n, \quad 0 \le g_n \le 1, \quad n = 0, 1, 2, \dots \tag{4.2.1}$$

if and only if $e_n \ge 0$ for all n and

$$\sum_{i=0}^{n} e_i \zeta_i^2 + 2\sum_{i=0}^{n-1} d_i \zeta_i \zeta_{i+1} \ge 0,$$

where ζ_i is an arbitrary real number for all i.

From this result the following can be deduced.

Lemma 4.2.3. *The ratio*

$$\frac{d_n^2}{e_n e_{n+1}}$$

is a chain sequence if and only if

$$\sum_{i=0}^{n} e_i \zeta_i^2 + 2\sum_{i=0}^{n-1} d_i \zeta_i \zeta_{i+1} \geq 0,$$

where ζ_i is an arbitrary real number for all i, $d_n \neq 0$ and $e_n > 0$ for all n.

Proof. If

$$\frac{d_n^2}{e_n e_{n+1}}$$

is a chain sequence, then $d_n \neq 0$, $e_n \geq 0$ for all n and

$$\sum_{i=0}^{n} e_i \zeta_i^2 + 2\sum_{i=0}^{n-1} d_i \zeta_i \zeta_{i+1} \geq 0,$$

where ζ_i is an arbitrary real number for all i. All that remains is to observe that $e_n \neq 0$ for each n. If

$$\sum_{i=0}^{n} e_i \zeta_i^2 + 2\sum_{i=0}^{n-1} d_i \zeta_i \zeta_{i+1} \geq 0,$$

where ζ_i is an arbitrary real number for all i, $d_n \neq 0$ and $e_n > 0$ for all n, then (4.2.1) holds, $g_n \neq 0$ for all n except possibly $n = 0$ and $g_n \neq 1$ for all n. $\square$

Let $\alpha(x)$ be a solution of a Stieltjes moment problem $\{\mu_n\}_{n=0}^{\infty}$ with infinitely many points of increase and with orthonormal polynomials $\{Q_n(x)\}_{n=0}^{\infty}$. Let the real parameters of the orthonormal three-term recurrence relation be $\{e_n\}_{n=0}^{\infty}$ and $\{d_n\}_{n=0}^{\infty}$. $\alpha(x)$ satisfies these properties if and only if $d_n \neq 0$ for all n, and

$$\mu\left(x\left\{\sum_{i=0}^{n} a_i x^i\right\}^2\right) = \sum_{i=0}^{n}\sum_{j=0}^{n} \mu_{i+j+1} a_i a_j \geq 0,$$

where a_i is an arbitrary real number for $0 \leq i \leq n$. Because the polynomials $\{Q_n(x)\}_{n=0}^{\infty}$ are a simple set, (cf. [1], p. 7)

$$\mu\left(x\left\{\sum_{i=0}^{n} a_i x^i\right\}^2\right) = \mu\left(x\left\{\sum_{i=0}^{n} \zeta_i Q_i(x)\right\}^2\right)$$

$$= \sum_{i=0}^{n} \sum_{j=0}^{n} \zeta_i \zeta_j \mu(xQ_i(x)Q_j(x))$$
$$= \sum_{i=0}^{n} \sum_{j=0}^{n} \zeta_i \zeta_j (d_i \delta_{i+1,j} + e_i \delta_{ij} + d_{i-1}\delta_{i-1,j})$$
$$= \sum_{i=0}^{n} e_i \zeta_i^2 + 2\sum_{i=0}^{n-1} d_i \zeta_i \zeta_{i+1} \geq 0.$$

Setting $\zeta_k = \delta_{ik}$ in the above gives $e_k \geq 0$, $k = 0, 1, 2, \ldots$.Because $d_n \neq 0$ for each n referring to (4.2.1) gives $|e_n| > 0$ for each n, so $e_n > 0$. Consequently

$$\frac{d_n^2}{e_n e_{n+1}}$$

is a chain sequence if and only if $\{d_n\}_{n=0}^{\infty}$ and $\{e_n\}_{n=0}^{\infty}$ are the parameters from the orthonormal recurrence for a set of polynomials generated by a solution $\alpha(x)$ of a Stieltjes moment problem with infinitely many points of increase.

The polynomials $\{S_n(x)\}_{n=0}^{\infty}$ satisfy the three-term recurrence relation

$$S_{n+1}(x) = xS_n(x) - \gamma_n S_{n-1}(x),$$

$\gamma_n > 0$ $n \geq 1$, $S_{-1}(x) = 0$, $S_0(x) = 1$. The absence of a constant coefficient of $S_n(x)$ is due to the fact that $S_n(x)$ has either only odd powers of x or only even powers of x. It is stipulated here that $\gamma_0 = 0$ (cf. [12], p.2).

It is known that the monic polynomials $\{P_n(x)\}_{n=0}^{\infty}$ and $\{K_n(x)\}_{n=0}^{\infty}$ satisfy three-term recurrence relations

$$P_{n+1}(x) = (x - c_n)P_n(x) - \lambda_n P_{n-1}(x)$$

$$K_{n+1}(x) = (x - u_n)K_n(x) - v_n K_{n-1}(x),$$

such that c_n, u_n are real, λ_n, $v_n > 0$ and $P_{-1}(x) = K_{-1}(x) = 0$, $P_0(x) = K_0(x) = 1$.

Considering the recurrence relation for $S_{2n}(x)$ and the connection between $S_n(x)$, $P_n(x)$ and $K_n(x)$ gives (cf. [12], p.2)

$$P_n(x^2) = x^2 K_{n-1}(x^2) - \gamma_{2n-1} P_{n-1}(x^2)$$

so that when x replaces x^2,

$$P_{n+1}(x) = xK_n(x) - \gamma_{2n+1}P_n(x). \tag{4.2.2}$$

The recurrence for $S_{2n+1}(x)$ gives,

$$xK_n(x^2) = xP_n(x^2) - \gamma_{2n}xK_{n-1}(x^2),$$

which after first dividing by x and then replacing x^2 with x gives

$$K_n(x) = P_n(x) - \gamma_{2n}K_{n-1}(x). \tag{4.2.3}$$

Solving for $P_n(x)$ in (4.2.3) and substituting into (4.2.2) gives for $K_n(x)$

$$\begin{aligned} K_{n+1}(x) + \gamma_{2n+2}K_n(x) &= xK_n(x) - \gamma_{2n+1}K_n(x) - \gamma_{2n+1}\gamma_{2n}K_{n-1}(x) \\ K_{n+1}(x) &= (x - \gamma_{2n+2} - \gamma_{2n+1})K_n(x) - \gamma_{2n+1}\gamma_{2n}K_n(x). \end{aligned}$$

Solving for $K_n(x)$ in (4.2.2) and substituting into (4.2.3) gives for $P_n(x)$

$$\begin{aligned} P_{n+1}(x) + \gamma_{2n+1}P_n(x) &= xP_n(x) - \gamma_{2n}P_n(x) - \gamma_{2n}\gamma_{2n-1}P_{n-1}(x) \\ P_{n+1}(x) &= (x - \gamma_{2n+1} - \gamma_{2n})P_n(x) - \gamma_{2n}\gamma_{2n-1}P_{n-1}(x). \end{aligned}$$

It follows that

$$\begin{aligned} c_n &= \gamma_{2n+1} + \gamma_{2n}, \quad \lambda_n = \gamma_{2n}\gamma_{2n-1}, \\ u_n &= \gamma_{2n+2} + \gamma_{2n+1}, \quad v_n = \gamma_{2n+1}\gamma_{2n}, \quad n \geq 1. \end{aligned} \tag{4.2.4}$$

These relations establish that

$$\frac{\lambda_{n+1}}{c_n c_{n+1}}$$

and

$$\frac{v_{n+1}}{u_n u_{n+1}}$$

are chain sequences. This is because (cf. [12], p.4) $0 < \gamma_{2n} < c_n$, so $\gamma_{2n} = g_{n-1}c_n$ and $\gamma_{2n+1} = (1 - g_{n-1})c_n$, giving

$$\frac{\lambda_{n+1}}{c_n c_{n+1}} = \frac{\gamma_{2n+2}\gamma_{2n+1}}{c_n c_{n+1}} \tag{4.2.5}$$

$$= (1 - g_{n-1})g_n,$$

in this case $g_{-1} = 0$ because $\gamma_0 = 0$. The same reasoning applies for

$$\frac{v_{n+1}}{u_n u_{n+1}}, \tag{4.2.6}$$

except that $g_{-1} > 0$ has to hold because $\gamma_1 > 0$.

Lemma 4.2.4 (cf. [15], p. 92)**.** *Let $\{a_n\}_{n=0}^{\infty}$ be a chain sequence. If $\{g_n\}_{n=-1}^{\infty}$ and $\{h_n\}_{n=-1}^{\infty}$ are parameter sequences for $\{a_n\}_{n=0}^{\infty}$ then $g_k < h_k$ if and only if $g_{-1} < h_{-1}$.*

Proof.

$$(1 - g_{n-1})g_n = a_n = (1 - h_{n-1})h_n$$

so

$$\frac{g_k}{h_k} = \frac{1 - h_{k-1}}{1 - g_{k-1}}.$$

The left hand side is less than 1 if and only if the right side side is less than 1; this happens if and only if

$$g_{k-1} < h_{k-1}.$$

□

Definition 4.2.5 (cf. [15], p. 93)**.** *If $\{a_n\}_{n=0}^{\infty}$ is a chain sequence and $\{m_n\}_{n=-1}^{\infty}$ is a parameter sequence corresponding to $\{a_n\}_{n=0}^{\infty}$, then $\{m_n\}_{n=-1}^{\infty}$ is called a minimal parameter sequence if $m_{-1} = 0$.*

By the above lemma this definition amounts to the fact that the parameters $\{m_n\}_{n=-1}^{\infty}$ are less than any other parameters of the sequence.

Definition 4.2.6 (cf. [15], p. 94)**.** *Let $\{a_n\}_{n=0}^{\infty}$ be chain sequence, then $\{M_n\}_{n=-1}^{\infty}$ is called a maximal parameter sequence if $M_k > g_k$, $k \geq -1$, for any other parameter sequence $\{g_n\}_{n=-1}^{\infty}$.*

Theorem 4.2.7 (cf. [15], p. 93). *Let $\{a_n\}_{n=0}^{\infty}$ be a chain sequence. If $\{g_n\}_{n=-1}^{\infty}$ is a parameter sequence for $\{a_n\}$ and $g_{-1} > 0$ then for every number h_{-1}, $0 \le h_{-1} < g_{-1}$, there is a corresponding parameter sequence $\{h_n\}_{n=-1}^{\infty}$ for the chain sequence $\{a_n\}_{n=0}^{\infty}$.*

In particular the preceding theorem implies the existence of minimal parameters for a given chain sequence.

Theorem 4.2.8 (cf. [15], p. 94). *If $\{a_n\}_{n=0}^{\infty}$ is a chain sequence then it has a maximal parameter sequence $\{M_n\}_{n=-1}^{\infty}$.*

Theorem 4.2.9 (cf. [15], p. 101). *A parameter sequence $\{M_n\}_{n=-1}^{\infty}$ is a maximal parameter sequence for a chain sequence if and only if*

$$\sum_{n=0}^{\infty} \frac{M_0 M_1 \dots M_n}{(1 - M_0)(1 - M_1) \dots (1 - M_n)} = \infty.$$

If the maximal parameter sequence $\{M_n\}_{n=-1}^{\infty}$ corresponding to the chain sequence $\{a_n\}_{n=0}^{\infty}$ satisfies $M_{-1} = 0$ then it is also a minimal parameter sequence and as a result there is only one parameter sequence corresponding to $\{a_n\}_{n=0}^{\infty}$.

The parameters of the polynomials $P_n(x)$ can correspond to a chain sequence with unique parameters because the given parameter sequence for (4.2.5) satisfied $g_{-1} = 0$. However, the parameters of the polynomials $K_n(x)$ necessarily correspond to a chain sequence with non-unique parameters, because the given parameter sequence for (4.2.6) satisfied $g_{-1} > 0$.

From the prior discussion, the chain sequence (cf. [14], p. 334)

$$\frac{\lambda_{n+1}}{c_n c_{n+1}}$$

has minimal parameters $m_{-1} = 0$

$$m_{n-1} = \frac{\gamma_{2n}}{c_n}.$$

The chain sequence

$$\frac{v_{n+1}}{u_n u_{n+1}}$$

has non-minimal parameters

$$g_n = \frac{\gamma_{2n+1}}{u_n}.$$

By choosing different parameters $0 \le h_n < M_n$, $h_n \ne g_n$ for each n a new choice for γ_n is arrived at via $\gamma_0^h = 0$,

$$\gamma_{2n+1}^h = h_n u_n$$
$$\gamma_{2n+2}^h = u_n - \gamma_{2n+1}.$$

The same polynomial set $K_n(x)$ with recurrence sequences $\{u_n\}_{n=0}^{\infty}$, $\{v_n\}_{n=0}^{\infty}$ corresponds to families of polynomials $\{S_n^h(x)\}_{n=0}^{\infty}$ and $\{P_n^h(x)\}_{n=0}^{\infty}$ (cf. [12], p. 5).

Lemma 4.2.10 (cf. [52], p.79)**.** *A constant term sequence* $\{a\}_{n=0}^{\infty}$ *is a chain sequence if and only if*

$$0 < a \le \frac{1}{4}.$$

Proof. Let $\{g_n\}_{n=-1}^{\infty}$ denote a parameter sequence for $\{a\}_{n=0}^{\infty}$. If $a > \frac{1}{4}$, then

$$(1 - g_{n-1})g_n > \frac{1}{4},$$

so

$$(\sqrt{(1 - g_{n-1})} - \sqrt{g_n})^2 \ge 0$$
$$\frac{(1 - g_{n-1}) + g_n}{2} \ge \sqrt{(1 - g_{n-1})g_n} > \frac{1}{2}$$
$$g_n > g_{n-1}.$$

The sequence of parameters is increasing and bounded above so it converges to a limit g. $(1 - g)g \ge \frac{1}{4}$ but if $(1 - g)g > \frac{1}{4}$ then by the above $g > g$, so $(1 - g)g = \frac{1}{4}$. This means that that the terms of the chain sequence must converge to $\frac{1}{4}$ i.e. the chain sequence is not constant. If $0 < a \le \frac{1}{4}$, then solving the quadratic equation

$$g(1 - g) = a$$

$$g^2 - g + a = 0$$

leads to

$$g = \frac{1 + \sqrt{1 - 4a}}{2},$$

which gives a constant parameter sequence $\{g\}_{n=-1}^{\infty}$ for the given chain sequence. $\square$

A discussion of results on chain sequences can be found in [18].

Chapter 5

Moment problems of classical polynomials

The classical orthogonal polynomials are discussed in light of the previous work. The Chebyshev polynomials are associated with a symmetric moment problem and this connection is used to explore the theory introduced in the previous chapter. Jacobi matrices of the Chebyshev polynomials are given. Moments are computed for the Chebyshev, Legendre, Hermite and Laguerre polynomials and it is shown that the moment problems associated with these orthogonal polynomial sets are determinate.

5.1 Chebyshev polynomials moment problem

The recurrence relation for the Chebyshev polynomials $\{T_n(x)\}_{n=0}^{\infty}$ was given in Theorem 1.3.10 by

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

with initial conditions

$$T_0(x) = 1, \quad T_1(x) = x.$$

From this it can be deduced that $T_{2n}(x)$ has only even powers of x and $T_{2n+1}(x)$ has only odd powers of x. Because $\{T_n(x)\}_{n=0}^{\infty}$ is a simple set,

$$x^{2n+1} = \sum_{k=0}^{2n+1} a_n T_n(x).$$

However, because $T_{2k}(x)$ has only even powers of x, $a_{2k} = 0$ for $0 < k < n$.

$$\int_{-1}^{1} x^{2n+1} \frac{1}{\sqrt{1-x^2}} dx = \sum_{k=0}^{n} a_{2k+1} \int_{-1}^{1} T_{2k+1}(x) \frac{1}{\sqrt{1-x^2}} dx = 0,$$

because

$$\int_{-1}^{1} T_{2n+1}(x) \frac{1}{\sqrt{1-x^2}} dx = 0$$

$n = 0, 1, 2 \ldots$, by the orthogonality condition. It is concluded that the moment problem associated with $\{T_n(x)\}_{n=0}^{\infty}$ is symmetric. For the second kind polynomials $U_n(x)$ the same recurrence relation is satisfied (see Theorem 1.3.11) and the initial conditions are

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

Similar reasoning to the case of $\{T_n(x)\}_{n=0}^{\infty}$ shows that $\{U_n(x)\}_{n=0}^{\infty}$ are associated with a symmetric moment problem.

Recall that from the orthogonality relation for the Chebyshev polynomials

$$\int_{-1}^{1} T_n^2(x) \frac{1}{\sqrt{1-x^2}} dx = \begin{cases} \frac{\pi}{2}, & n \geq 1 \\ \pi, & n = 0, \end{cases}$$

so the orthonormal Chebyshev polynomials are given by

$$T_0'(x) = \frac{T_0(x)}{\sqrt{\pi}}$$

and

$$T_n'(x) = \sqrt{\frac{2}{\pi}} T_n(x), \quad n \geq 1.$$

For the time being let $\{P_n(x)\}_{n=0}^{\infty}$ denote the orthonormal Chebyshev polynomials. Then

$$P_2(x) = 2xP_1(x) - \sqrt{2}P_0(x)$$

$$P_{n+1}(x) = 2xP_n - P_{n-1}(x), \quad n \geq 2.$$

Dividing the above recurrences by 2 and examining the result gives the Jacbobi matrix associated with the Chebyshev polynomials

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \dots \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & 0 & \dots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where $e_n = 0$ for each n, $d_0 = \dfrac{1}{\sqrt{2}}$, and $d_n = \dfrac{1}{2}$, $n \geq 1$ are the sequences $\{e_n\}_{n=0}^{\infty}$ and $\{d_n\}_{n=0}^{\infty}$ from the orthonormal recurrence relation.

For the rest of this section let $T_n(x)$ denote the monic Chebyshev polynomial of degree n. Then

$$\begin{aligned} T_2(x) &= xT_1(x) - \frac{1}{2}T_0(x) \\ T_{n+1}(x) &= xT_n(x) - \frac{1}{4}T_{n-1}(x), \quad n \geq 2. \end{aligned}$$

Let the sequence $\{\gamma_n\}_{n=0}^{\infty}$ be given by $\gamma_0 = 0$, $\gamma_1 = \dfrac{1}{2}$, $\gamma_n = \dfrac{1}{4}$, $n \geq 1$. The sequences $\{c_n\}_{n=0}^{\infty}$ and $\{\lambda_n\}_{n=0}^{\infty}$, can then be constructed using the formula (4.2.4). It follows that

$$\frac{\lambda_1}{c_0 c_1} = \frac{1}{2} \text{ and } \frac{\lambda_{n+1}}{c_n c_{n+1}} = \frac{1}{4},$$

is a chain sequence. The minimal parameters of this chain sequence are

$$m_{-1} = 0 \text{ and } m_n = \frac{1}{2} \quad n \geq 0.$$

These are also the maximal parameters since referring to Theorem 4.2.9

$$\sum_{n=0}^{\infty} \frac{m_0 m_1 \dots m_n}{(1-m_0)(1-m_1)\dots(1-m_n)} = \sum_{n=0}^{\infty} \frac{\frac{1}{2}^{n+1}}{\frac{1}{2}^{n+1}} = \infty.$$

So this chain sequence determines its parameters uniquely.

Recalling that

$$\int_{-1}^{1} U_n^2(x)\sqrt{1-x^2}dx = \pi, \quad \forall n.$$

The Jacobi matrix for the Chebyshev polynomials of the second kind is given by

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & \cdots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Let $U_n(x)$ now denote the monic Chebyshev polynomial of the second kind. Then

$$U_{n+1}(x) = xU_n(x) - \frac{1}{4}U_{n-1}(x).$$

Let $\gamma_0 = 0$, $\gamma_n = \frac{1}{4}$, $n \geq 1$. Then the chain sequence

$$\frac{\lambda_{n+1}}{c_n c_{n+1}}$$

is identical to that obtained from the Chebyshev polynomials of the first kind. Consider the chain sequence

$$\frac{v_{n+1}}{u_n u_{n+1}} = \frac{1}{4}, \quad n \geq 0.$$

This is the maximal constant chain sequence. It can be seen from the calculation for the Chebyshev polynomials of the first kind that the parameters $\left\{\frac{1}{2}\right\}_{n=-1}^{\infty}$ are the maximal parameters of this sequence (cf. [52], p. 80). The minimal parameters can be calculated by setting $m_{-1} = 0$ and noting that inductively (cf. [52], p. 80)

$$\begin{aligned} m_0 &= \frac{1}{4} = \frac{1}{2}\left(1 - \frac{1}{2}\right) \\ \text{if} \quad m_{n-1} &= \frac{1}{2}\left(1 - \frac{1}{n+1}\right) \quad \text{then} \\ m_n &= \frac{1}{4}\frac{1}{(1 - \frac{1}{2}(1 - \frac{1}{n+1}))} \\ &= \frac{1}{2}\frac{n+1}{n+2} \\ &= \frac{1}{2}\left(1 - \frac{1}{n+2}\right) \end{aligned}$$

From (4.2.4),

$$u_n = \frac{1}{2}$$

and the parameters generated by the original choice of $\{\gamma_n\}_{n=0}^{\infty}$ are

$$\frac{\gamma_{2n+1}}{u_n} = \frac{1}{2}.$$

So the original recurrences generate the maximal parameter sequence associated with the chain sequence.

Theorem 5.1.1. *The even moments $\{\mu_{2n}\}_{n=0}^{\infty}$ of the Chebyshev polynomials of the first kind $\{T_n(x)\}_{n=0}^{\infty}$ are given by*

$$\mu_{2n} = \frac{(2n)!}{2^{2n}(n!)^2}\pi.$$

Proof.

$$\begin{aligned}\int_0^{\pi} \cos^{2n}(\theta)d\theta &= (\sin(\theta)\cos^{2n-1}(\theta))_0^{\pi} + (2n-1)\int_0^{\pi} \sin^2(\theta)\cos^{2n-2}(\theta)d\theta \\ &= (2n-1)\int_0^{\pi}(1-\cos^2(\theta))\cos^{2n-2}(\theta)d\theta \\ \int_0^{\pi} \cos^{2n}(\theta)d\theta &= \frac{2n-1}{2n}\int_0^{\pi}\cos^{2n-2}(\theta)d\theta.\end{aligned}$$

Iterating this relation with the initial condition

$$\int_0^{\pi}\cos^0(\theta)d\theta = \pi,$$

gives

$$\frac{2n-1}{2n}\frac{2n-3}{2n-2}\frac{2n-5}{2n-4}\cdots\frac{1}{2}\pi, \tag{5.1.1}$$

which is the required result. □

The even moments associated with the polynomials of the second kind $\{U_n(x)\}_{n=0}^{\infty}$ are given by

$$\begin{aligned}\int_0^{\pi}\cos^{2n}(\theta)\sin^2(\theta)d\theta &= \int_0^{\pi}(1-\cos^2(\theta))\cos^{2n}(\theta)d\theta \\ &= \int_0^{\pi}\cos^{2n}(\theta)d\theta - \int_0^{\pi}\cos^{2n+2}(\theta)d\theta\end{aligned}$$

$$= \mu_{2n} - \mu_{2n+2}, \tag{5.1.2}$$

where μ_{2n} is the $2n^{th}$ moment associated with the polynomials of the first kind.

From relation (5.1.1) the moments of the polynomials of the first kind are all less than 1. Furthermore it is seen from the fact that the even moments of $\{U_n(x)\}_{n=0}^{\infty}$ are positive and (5.1.2) that the even moments of $\{T_n(x)\}_{n=0}^{\infty}$ are a decreasing sequence.

A condition that guarantees determinacy of a moment problem is if the moments of a distribution grow slowly enough.

Lemma 5.1.2 (cf. [47], p88)**.** *Let $\{\mu_n\}_{n=0}^{\infty}$ be the moments of a solvable Hamburger moment problem. Suppose there are positive constants C and R such that*

$$|\mu_n| \leq CR^n n!,$$

then the moment problem is determinate.

It can be seen from the above discussion and the lemma that the moment problems associated with $\{T_n(x)\}_{n=0}^{\infty}$ and $\{U_n(x)\}_{n=0}^{\infty}$ are determinate.

5.2 Legendre moment problem

The moments $\{\mu_n\}_{n=0}^{\infty}$ of the Legendre polynomials are given by

$$\begin{aligned}\mu_n &= \int_{-1}^{1} x^n dx = \frac{1}{n+1} - \frac{(-1)^{n+1}}{n+1} \\ &= 0, \quad n \text{ odd} \\ &= \frac{2}{n+1}, \quad n \text{ even.}\end{aligned}$$

The moments are all less than 1 and the even moments form a decreasing sequence. By Lemma 5.1.2 the Legendre moment problem is determinate. The Legendre moment problem is also a symmetric problem.

5.3 Ultraspherical polynomials

Definition 5.3.1 (cf. [15], p. 44). *Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ which satisfy $\alpha = \beta$ are called ultrashperical polynomials.*

It was mentioned earlier that the Chebyshev polynomials of the first and second kind and the Legendre polynomials are examples of ultraspherical polynomials. It is known that the weight function for Jacobi polynomials is

$$(1-x)^{\alpha}(1+x)^{\beta}.$$

If $\alpha = \beta$ this becomes

$$(1-x^2)^{\alpha},$$

which is an even function. Because of this, all moment problems associated with ultraspherical polynomials are symmetric moment problems. The approach used to prove that $T_n(x)$ and $U_n(x)$ are associated with symmetric moment problems can be used to prove that a moment problem is symmetric if and only if the associated monic orthogonal polynomials $\{P_n(x)\}_{n=0}^{\infty}$ satisfy

$$P_{n+1}(x) = xP_n(x) - \gamma_n P_{n-1}(x),$$

$$P_1(x) = x, \quad P_0(x) = 1.$$

On the one hand a symmetric orthogonal polynomial of odd degree is a sum of odd powers of x and a symmetric orthogonal polynomial of even degree is a sum of even powers of x so the recurrence and initial conditions hold in this case. With the recurrence and initial conditions given, the proof for $T_n(x)$ reworded suffices to show that the moment problem is symmetric. A class of orthogonal polynomials that have a similar recurrence relation to the symmetric case (except that $P_1(0) \neq 0$) has been studied (cf. [16]).

5.4 Hermite and Laguerre moment problems

Define $\alpha(x)$ by

$$\alpha(x) = \int_{-\infty}^{x} e^{-t^2} dt.$$

The integrand is positive, so $\alpha(x)$ is non-decreasing. $\alpha(x)$ is bounded above by

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Using integration by parts and the fact that $x^n e^{-x^2} \to 0$ as $x \to \infty$ or $x \to -\infty$ then gives that $\alpha(x)$ has finite moments, so it is a distribution. Because e^{-x^2} is an even function $x^{2n+1}e^{-x^2}$ is odd so all of the odd moments of $\alpha(x)$ vanish; i.e. $\alpha(x)$ is associated with a symmetric moment problem. $\alpha(x) + \alpha(-x) = \sqrt{\pi}$ so $\alpha(x)$ is a symmetric distribution. The polynomials which are orthogonal with respect to $\alpha(x)$ are the Hermite polynomials $\{H_n(x)\}_{n=0}^{\infty}$. Because orthogonal polynomials are essentially the same up to normalisation it is stipulated that the polynomials $H_n(x)$ are monic.

Following the discussion in Section 4.1 let $\psi(x)$ be the distribution defined by $\psi(x) = \alpha(\sqrt{x})$ for $x \in [0, \infty)$. Then

$$d\psi(x) = \frac{d\alpha(\sqrt{x})}{dx} dx = \frac{1}{2} x^{-\frac{1}{2}} e^{-x} dx.$$

Let $\{P_n(x)\}_{n=0}^{\infty}$ be the monic polynomials orthogonal with respect to $\psi(x)$. Then $P_n(x^2) = H_{2n}(x)$.

Let $\omega(x)$ be the distribution defined by

$$\omega(x) = \int_{0}^{x} t d\psi(t),$$

and $\{K_n(x)\}_{n=0}^{\infty}$ be the monic polynomials orthogonal with respect to $\omega(x)$. Then

$$d\omega(x) = x d\psi(x) = \frac{1}{2} x^{\frac{1}{2}} e^{-x} dx,$$

and $xK_n(x^2) = H_{2n+1}(x)$.

Let the function $\nu^\alpha(x)$ be defined by,

$$\nu^\alpha(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \displaystyle\int_0^x t^\alpha e^{-t} dt & \text{if } x > 0 \end{cases}$$

where $\alpha > -1$ is a real number. This function is non-decreasing, because $t^\alpha e^{-t}$ is positive for $t \in [0, \infty)$. The integral converges for all $x \in [0, \infty)$ and represents the incomplete gamma function $\gamma(\alpha+1, x)$ (cf. [5], p197). In particular the total variation of $\nu^\alpha(x)$ is given by

$$\int_0^\infty t^\alpha e^{-t} dt = \Gamma(\alpha+1),$$

which is finite for all real $\alpha > -1$. The discussion in the previous chapter indicates that it is sufficient to consider the weight function which generates $\nu^\alpha(x)$. Because of it's relation with the gamma function it is easy to represent all of the moments associated with $\nu^\alpha(x)$

$$\begin{aligned} &\int_0^\infty x^n d\nu^\alpha(x) \\ &= \int_0^\infty x^n x^\alpha e^{-x} dx \\ &= \Gamma(\alpha+n+1) < \infty, \end{aligned}$$

so $\nu^\alpha(x)$ is a distribution function.

The polynomials orthogonal with respect to $\nu^\alpha(x)$ are the Laguerre polynomials $\{L_n^\alpha(x)\}_{n=0}^\infty$. Let $\alpha = -\dfrac{1}{2}$. Then from the orthogonality relations

$$\int_0^\infty L_n^{-\frac{1}{2}}(x) L_m^{-\frac{1}{2}}(x) x^{-\frac{1}{2}} e^{-x} dx = h_n \delta_{mn} \quad h_n > 0$$

$$\int_0^\infty P_n(x) P_m(x) x^{-\frac{1}{2}} e^{-x} dx = g_n \delta_{mn} \quad g_n > 0,$$

it follows, that $L_n^{-\frac{1}{2}}(x^2) = CP_n(x^2) = CH_{2n}(x)$, C a real constant, which is a result from Chapter 1.

Similarly from the orthogonality relations

$$\int_0^\infty L_n^{\frac{1}{2}}(x) L_m^{\frac{1}{2}}(x) x^{\frac{1}{2}} e^{-x} dx = h_n \delta_{mn} \quad h_n > 0$$

$$\int_0^\infty K_n(x)K_m(x)x^{\frac{1}{2}}e^{-x}dx = g_n\delta_{mn} \quad g_n > 0,$$

it follows, that $xL_n^{\frac{1}{2}}(x^2) = AxK(x^2) = AH_{2n+1}(x)$, A a real constant.

As in the previous chapter, the moments of $\alpha(x)$ can be related to the moments of $\psi(x)$ by

$$\int_{-\infty}^\infty x^{2n}d\alpha(x) = 2\int_0^\infty x^n d\psi(x).$$

The moments of $\psi(x)$ can then be related to the moments of $\nu^{-\frac{1}{2}}(x)$:

$$\begin{aligned}\int_0^\infty x^n d\psi(x) &= \frac{1}{2}\int_0^\infty x^n x^{-\frac{1}{2}}e^{-x}dx \\ &= \frac{1}{2}\int_0^\infty x^n d\nu^{-\frac{1}{2}}(x).\end{aligned}$$

It follows that

$$\int_{-\infty}^\infty x^{2n}d\alpha(x) = \Gamma\left(n+\frac{1}{2}\right).$$

Lemma 5.1.2 can be used to establish the determinacy of the moment problem associated with the Laguerre polynomials.

The moments of $\nu^\alpha(x)$ are given by

$$\mu_n = \Gamma(\alpha+n+1),$$

where $\alpha > -1$ is real. Let m be the smallest integer greater than α then because $\Gamma(x)$ is an increasing real-valued function for $x > 2$ real (this can be seen from the relation $\Gamma(x+1) = x\Gamma(x)$),

$$\Gamma(\alpha+n+1) < C_1\Gamma(n+m+1),$$

for all n and a suitably chosen real $C_1 > 0$ (because there can only be two moments before $n+\alpha+1 > 2$). Also, $\Gamma(n+m+1) = (n+m)!$, and

$$(n+m)! = \prod_{k=1}^m (n+k)n!$$

$$= P_m(n)n! < e^n n!$$

for large enough n. Choose $C_2 > 0$ such that for all n

$$\Gamma(\alpha + n + 1) < C_1 C_2 e^n n!.$$

Then if $C = C_1 C_2$ and $R = e$ the conditions of Lemma 5.1.2 are established.

Because the odd moments of the Hermite polynomials are zero and the even moments are moments of Laguerre polynomials, the same constant C that was used for the laguerre polynomials satisfies the conditions of the lemma and for R the choice $R = e$ is sufficient. This shows that the moment problem associated with the Hermite polynomials is also determined.

Chapter 6

q-Extensions

q-Extensions are a current field of active research in orthogonal polynomials and special functions. To begin the chapter the basic theory is provided. The theory presented is applied to the q-Laguerre orthogonal polynomials and the associated q-Laguerre moment problem.

6.1 Basic hypergeometric series

By assigning a parameter q to a class of special functions, and in particular to the hypergeometric series, analogues of the functions have been discovered which preserve many important properties. As $q \to 1$ the original functions are recovered. These analogues are called q-analogues or q-extensions.

Hypergeometric series are built from Pochammer symbols so it is natural that in extending the hypergeometric series an extension of the Pochammer symbol is obtained. This is given by the q-shifted factorial.

Definition 6.1.1 (cf. [20], p.3). *The symbol $(a;q)_n$ is called the q-shifted factorial and is given by*

$$(a;q)_n = (1-a)(1-aq)\dots(1-aq^{n-1}), \qquad n = 1, 2, 3, \dots$$
$$(a;q)_\infty = \prod_{k=0}^{\infty}(1-aq^k), \qquad |q| < 1$$

$(a;q)_0 = 1.$

$$\lim_{q\to 1} \frac{1-q^a}{1-q} = a$$

so

$$\lim_{q\to 1} \frac{(q;q)_n}{(1-q)^n} = n!$$

and in general

$$\lim_{q\to 1} \frac{(q^a;q)_n}{(1-q)^n} = (a)_n$$

where $(a)_n$ is Pochammer's symbol (1.2.1).

Some identities with q-shifted factorials will be frequently used in basic hypergeometric series so they are given here explicitly.

Lemma 6.1.2 (cf. [20], p.6)**.**

$$(a;q)_n = \frac{(a;q)_\infty}{(aq^n;q)_\infty}.$$

Proof.

$$\begin{aligned} \frac{(a;q)_\infty}{(aq^n;q)_\infty} &= \frac{\prod_{k=0}^{n-1}(1-aq^k)\prod_{k=n}^{\infty}(1-aq^k)}{(aq^n;q)_\infty} \\ &= \frac{(a;q)_n(aq^n;q)_\infty}{(aq^n;q)_\infty} \\ &= (a;q)_n. \end{aligned}$$

□

Definition 6.1.3 (cf. [20], p. 4)**.** *Here the symbol* $\binom{n}{2}$ *denotes* $\frac{n(n-1)}{2}$.

Lemma 6.1.4 (cf. [20], p.6)**.**

$$(a^{-1}q^{1-n};q)_n = (a;q)_n(-a^{-1})^n q^{-\binom{n}{2}}. \tag{6.1.1}$$

Proof. First note that $\binom{n}{2} = \frac{n(n-1)}{2} = \sum_{k=0}^{n-1} k$. Then

$$\begin{aligned}
(a;q)_n(-a^{-1})^n q^{-\binom{n}{2}} &= \prod_{k=0}^{n-1}(q^k - a^{-1})q^{-\binom{n}{2}} \\
&= \prod_{k=0}^{n-1}(1 - a^{-1}q^{-k}) \prod_{k=0}^{n-1} q^k q^{-\binom{n}{2}} \\
&= \prod_{k=0}^{n-1}(1 - a^{-1}q^{1-n}q^k) \\
&= (a^{-1}q^{1-n};q)_n,
\end{aligned}$$

where the order of the product was reversed at the third step to get the result. □

Lemma 6.1.5 (cf. [20], p.6)**.**

$$(q^{-n};q)_n = (q;q)_n(-1)^n q^{-\binom{n+1}{2}}.$$

Proof.

$$\begin{aligned}
(q;q)_n(-1)^n q^{-\binom{n+1}{2}} &= \prod_{k=1}^{n}(1 - q^{-k}) \prod_{k=1}^{n} q^k q^{-\binom{n}{2}-n} \\
&= \prod_{k=0}^{n-1}(1 - q^{-k-1}) \\
&= (q^{-n};q)_n,
\end{aligned}$$

where the order of the product was reversed for the last step. □

Products of q-shifted factorials are common and a compact notation is used for them.

Definition 6.1.6 (cf. [20], p.6)**.**

$$(a_1, a_2, \ldots, a_m; q)_n = (a_1;q)_n(a_2;q)_n \ldots (a_m;q)_n$$

$$(a_1, a_2, \ldots, a_m; q)_\infty = (a_1;q)_\infty(a_2;q)_\infty \ldots (a_m;q)_\infty.$$

Letting $q \to 1$ in the series (cf. [20], p.3)

$$\phi(q^a, q^b; q^c; q, z) = \sum_{n=0}^{\infty} \frac{(q^a;q)_n (q^b;q)_n}{(q;q)_n (q^c;q)_n} z^n,$$

gives the hypergeometric series

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

For this reason the series $\phi(q^a, q^b; q^c; q, z)$ where $|q| < 1$ is called the basic hypergeometric series, where basic refers to the base q. This series is the extension or analogue of the ordinary hypergeometric series which was anticipated at the beginning of the chapter. The series can be further generalised by replacing q^a, q^b and q^c by complex parameters a, b and c.

Lemma 6.1.7 (cf. [20], p.3)**.** *The series $\phi(a, b; c; q, z)$ converges absolutely for $|z| < 1$ and $|q| < 1$.*

Proof. The ratio test gives

$$\lim_{n\to\infty} \left| \frac{(a;q)_{n+1} (b;q)_{n+1} z^{n+1}}{(c;q)_{n+1} (q;q)_{n+1}} \frac{(c;q)_n (q;q)_n}{(a;q)_n (b;q)_n z^n} \right|$$

$$= \lim_{n\to\infty} \left| \frac{(1 - aq^n)(1 - bq^n)}{(1 - cq^n)(1 - q^{n+1})} z \right| = |z| < 1$$

for $|q| < 1$. □

Definition 6.1.8 (cf. [20], p.4)**.** *The generalised basic hypergeometric series is given by*

$$\begin{aligned} &{}_r\phi_s(a_1, a_2, \ldots, a_r; b_1, b_1, \ldots, b_s; q, z) \\ &= \sum_{n=0}^{\infty} \frac{(a_1;q)_n (a_2;q)_n \ldots (a_r;q)_n}{(q;q)_n (b_1;q)_n \ldots (b_s;q)_n} \left[(-1)^n q^{\binom{n}{2}} \right]^{1+s-r} z^n \end{aligned}$$

where the symbol $\binom{n}{2}$ is as above and $q \neq 0$ when $s + 1 < r$.

The q-binomial theorem is a q-analogue for the ordinary binomial theorem and can be used to derive a q-analogue of the ordinary exponential function.

Theorem 6.1.9 (cf. [7], p. 350)**.**

$$ {}_1\phi_0(a;-;q,z) = \sum_{n=0}^{\infty} \frac{(a;q)_n}{(q;q)_n} z^n = \frac{(az;q)_\infty}{(z;q)_\infty} \tag{6.1.2}$$

for $|z| < 1$, $|q| < 1$.

Proof. Let $h_a(z)$ be the series

$$\sum_{n=0}^{\infty} \frac{(a;q)_n}{(q;q)_n} z^n.$$

Then since $(a,q)_0 = (aq;q)_0 = 1$,

$$\begin{aligned}
h_a(z) - h_{aq}(z) &= \sum_{n=1}^{\infty} \frac{(a;q)_n - (aq;q)_n}{(q;q)_n} z^n \\
&= \sum_{n=1}^{\infty} \frac{(aq;q)_{n-1}}{(q;q)_n} [1 - a - (1 - aq^n)] z^n \\
&= -a \sum_{n=1}^{\infty} \frac{(1-q^n)(aq;q)_{n-1}}{(q;q)_n} z^n \\
&= -a \sum_{n=0}^{\infty} \frac{(aq;q)_n}{(q;q)_n} z^{n+1} \\
&= -az h_{aq}(z).
\end{aligned}$$

Similarly

$$\begin{aligned}
h_a(z) - h_a(qz) &= \sum_{n=1}^{\infty} \frac{(a;q)_n}{(q;q)_n} (z^n - q^n z^n) \\
&= \sum_{n=1}^{\infty} \frac{(a;q)_n}{(q;q)_{n-1}} z^n \\
&= \sum_{n=0}^{\infty} \frac{(a;q)_{n+1}}{(q;q)_n} z^{n+1} \\
&= (1-a) z h_{aq}(z).
\end{aligned}$$

So $h_{aq}(z) = \dfrac{h_a(z) - h_a(qz)}{(1-a)z}$ and

$$h_a(z) - \frac{h_a(z) - h_a(qz)}{(1-a)z} = -az \frac{h_a(z) - h_a(qz)}{(1-a)z}$$

$$h_a(z)(z - az + az - 1) = h_a(qz)(az - 1)$$
$$h_a(z) = \frac{(1 - az)}{(1 - z)} h_a(qz).$$

Iteration gives

$$h_a(z) = \frac{(az; q)_n}{(z; q)_n} h_a(q^n z).$$

and as $n \to \infty$

$$h_a(z) = \frac{(az; q)_\infty}{(z; q)_\infty} h_a(0) = \frac{(az; q)_\infty}{(z; q)_\infty}.$$

□

If $a = q^c$, then as $q \to 1^-$ in (6.1.2) the ordinary binomial series is obtained, i.e.

$${}_1F_0(c; -; z) = \sum_{n=0}^{\infty} \frac{(c)_n}{n!} z^n.$$

Ramanujan's sum formula will feature in the following considerations.

Theorem 6.1.10 (cf. [20], p.126)**.**

$$\sum_{n=-\infty}^{\infty} \frac{(a; q)_n}{(b; q)_n} z^n = \frac{\left(q, \dfrac{b}{a}, az, \dfrac{q}{az}; q\right)_\infty}{\left(b, \dfrac{q}{a}, z, \dfrac{b}{az}; q\right)_\infty},$$

for $\left|\dfrac{b}{a}\right| < |z| < 1.$

Analogues of the well-known special functions play an important role in the theory of basic hypergeometric series. In particular the analogue of the gamma function is used here.

Definition 6.1.11 (cf. [20], p.17)**.** *The q-analogue of the gamma function is defined by*

$$\Gamma_q(x) = \frac{(q; q)_\infty}{(q^x; q)_\infty} (1 - q)^{1-x},$$

$0 < q < 1.$

It can be shown (cf. [20], p. 17) that

$$\lim_{q\to 1^-}\Gamma_q(x)=\Gamma(x).$$

Lemma 6.1.12 (cf. [20], p. 17)**.**

$$\Gamma_q(x+1)=\frac{1-q^x}{1-q}\Gamma_q(x)$$

Proof.

$$\begin{aligned}\Gamma_q(x+1)&=\frac{(q;q)_\infty}{(q^{x+1};q)_\infty}(1-q)^{-x}\\&=\frac{1-q^x}{1-q}\frac{(q;q)_\infty}{(1-q^x)\prod_{k=0}^{\infty}(1-q^{x+1+k})}(1-q)^{1-x}\\&=\frac{1-q^x}{1-q}\Gamma_q(x).\end{aligned}$$

□

Recalling that $\lim\limits_{q\to 1^-}\dfrac{1-q^x}{1-q}=x$ shows that this relation is the q-analogue of the familiar relation

$$\Gamma(x+1)=x\Gamma(x),$$

satisfied by the ordinary gamma function. Iteration gives

$$\Gamma_q(x+n)=\frac{(q^x;q)_n}{(1-q)^n}\Gamma_q(x). \tag{6.1.3}$$

Lemma 6.1.13.

$$\frac{\Gamma_q(x+n)}{\Gamma_q(x)}=\frac{(q^x;q)_n}{(1-q)^n}.$$

Proof.

$$\begin{aligned}\frac{\Gamma_q(x+n)}{\Gamma_q(x)}&=\frac{(q^x;q)_n}{(1-q)^n}\frac{\Gamma_q(x)}{\Gamma_q(x)}\\&=\frac{(q^x;q)_n}{(1-q)^n}.\end{aligned}$$

□

6.2 q-Laguerre polynomials

Definition 6.2.1 (cf. [20], p.19)**.** *Let $0<q<1$, then $\int_0^\infty f(x)d_qx$ is used to denote the series*

$$(1-q)\sum_{n=-\infty}^{\infty}f(q^n)q^n \tag{6.2.1}$$

The set of points $\{q^n\}_{n=-\infty}^{\infty}$ forms a partition of the interval $[0,\infty)$ and q^n is unbounded for large enough negative n. Furthermore, as $q \to 1^-$, the partitions get finer so that in the limit (cf. [20], p.19)

$$\lim_{q\to 1^-}(1-q)\sum_{n=-\infty}^{\infty} f(q^n)q^n = \int_0^{\infty} f(x)dx,$$

for any continuous function $f(x)$. For this reason the operator given by (6.2.1) can be seen as a q-analogue of the Riemann integral on the interval $[0,\infty)$.

Suppose a distribution function $\alpha(x)$ is given. Then if $c > 0$ is a real constant, $c\alpha(x)$ is still bounded, non-decreasing and has finite moments, i.e. $c\alpha(x)$ is a distribution function and by Lemma 2.1.10

$$\int_{-\infty}^{\infty} f(x)d[c\alpha(x)] = c\int_{-\infty}^{\infty} f(x)d\alpha(x).$$

Let q be fixed such that $0 < q < 1$, and let $\alpha(x)$ be a function constant except for jumps of size $a_n q^n$ at the points q^n, where each $a_n > 0$ is chosen so that the series

$$\sum_{n=-\infty}^{\infty} |P_n(q^n)a_n q_n|$$

converges for an arbitrary polynomial $P_n(x)$, then $\alpha(x)$ is a distribution function and so is $(1-q)\alpha(x)$. The constants $\{a_n\}_{n=0}^{\infty}$ can be thought of together as generated by a weight function $w(x)$ defined so that $w(q^n) = a_n$. The distribution $(1-q)\alpha(x)$ generates a canonical inner product on polynomials $\langle\cdot,\cdot\rangle$ which is given by

$$\langle P_n(x), P_m(x)\rangle = \int_0^{\infty} P_n(x)P_m(x)w(x)d_qx.$$

Using this inner product a set of orthogonal polynomials corresponding to the weight function $w(x)$ can be generated. In contrast with the classical polynomials this weight function only needs to be defined at the points q^n.

The q-binomial theorem can be used to derive a q-analogue of the exponential function. The ordinary binomial series gives

$$e^z =_1 F_0(0;-;z) = \sum_{n=0}^{\infty}\frac{z^n}{n!}.$$

Definition 6.2.2 (cf. [20], p. 9)**.** *Let* $|z| < 1$, $|q| < 1$ *and define* $e_q(z)$ *by*

$$e_q(z) = {}_1\phi_0(0; -; q, z) = \sum_{n=0}^{\infty} \frac{z^n}{(q;q)_n} = \frac{1}{(z;q)_\infty}.$$

By prior considerations $\lim_{q \to 1^-} e_q(z(1-q)) = e^z$.

Recall that the Laguerre polynomials satisfied the orthogonality condition

$$\int_0^\infty L_n^\alpha(x) L_m^\alpha(x) x^\alpha e^{-x} dx = h_n \delta_{mn}.$$

The q-Laguerre polynomials $L_n^\alpha(x;q)$ can be defined by the orthogonality relation (cf. [38], p.24)

$$\begin{aligned}\frac{1}{A} \sum_{k=-\infty}^{\infty} L_n^\alpha(q^k;q) L_m^\alpha(q^k;q) \frac{q^{k\alpha+k}}{(-(1-q)q^k;q)_\infty} \\ &= \frac{(q^{\alpha+1};q)_n}{q^n(q;q)_n}, \quad m = n \\ &= 0, \quad m \neq n,\end{aligned}$$

where

$$A = \sum_{n=-\infty}^{\infty} \frac{q^{k\alpha+k}}{(-(1-q)q^k;q)_\infty}.$$

Using the q-integral operator this definition can be written more compactly.

Definition 6.2.3 (cf. [38],p.24)**.** *The* q*-Laguerre polynomials are defined by the orthogonality condition*

$$\begin{aligned}\frac{1}{A(1-q)} \int_0^\infty L_n^\alpha(x;q) L_m^\alpha(x;q) x^\alpha e_q(-(1-q)x) d_q x \\ &= \frac{(q^{\alpha+1};q)_n}{q^n(q;q)_n}, \quad m = n \\ &= 0, \quad m \neq n,\end{aligned} \tag{6.2.2}$$

where A *is as above.*

A can be chosen arbitrarily with the polynomials invariant up to normalisation, but for this specific choice of A the 0^{th} moment

$$\mu_0 = \frac{1}{A(1-q)} \int_0^\infty x^\alpha e_q(-(1-q)x) d_q x = 1.$$

The classical Laguerre polynomials can be represented by the hypergeometric series

$$\frac{(\alpha+1)_n}{n!}\ {}_1F_1(-n;\alpha+1;x).$$

Theorem 6.2.4 (cf. [38], p.21)**.** *The q-Laguerre polynomials $L_n^\alpha(x;q)$ can be represented by*

$$\begin{aligned} L_n^\alpha(x;q) &= \frac{(q^{\alpha+1};q)_n}{(q;q)_n}\sum_{k=0}^{n}\frac{(q^{-n};q)_k q^{\binom{k}{2}}(1-q)^k(q^{n+\alpha+1}x)^k}{(q^{\alpha+1};q)_k(q;q)_k} \\ &= \frac{(q^{\alpha+1};q)_n}{(q;q)_n}\ {}_1\phi_1(q^{-n};q^{\alpha+1};q,-(1-q)q^{n+\alpha+1}x). \end{aligned}$$

Theorem 6.2.5 (cf. [38], p.26)**.** *The q-Laguerre polynomials satisfy a three-term recurrence relation of the form*

$$\begin{aligned} &xL_n^\alpha(x;q) = -\frac{(1-q^{n+1})}{(1-q)q^{2n+\alpha+1}}L_{n+1}^\alpha(x;q) \\ &+\left\{\frac{(1-q^n)}{(1-q)q^{2n+\alpha}}+\frac{(1-q^{n+\alpha+1})}{(1-q)q^{2n+\alpha+1}}\right\}L_n^\alpha(x;q)-\frac{(1-q^{n+\alpha})}{(1-q)q^{2n+\alpha}}L_{n-1}^\alpha(x;q). \\ &L_0^\alpha(x;q)=1,\quad L_1^\alpha(x;q)=-q^{\alpha+1}x+\frac{(1-q^{\alpha+1})}{1-q}. \end{aligned}$$

6.3 The q-Laguerre moment problem

Theorem 6.1.10 can be used to calculate the moments $\{\mu_n\}_{n=0}^\infty$ associated with the q-Laguerre polynomials.

Lemma 6.3.1 (cf. [38], p.24)**.**

$$\sum_{k=-\infty}^{\infty}\frac{q^{\beta k}}{(aq^k;q)_\infty}=\frac{\left(aq^\beta,\dfrac{q^{1-\beta}}{a},q;q\right)_\infty}{\left(q^\beta,a,\dfrac{q}{a};q\right)_\infty}. \tag{6.3.1}$$

Proof. Theorem 6.1.10 gives

$$\sum_{n=-\infty}^{\infty}\frac{(a;q)_n}{(b;q)_n}x^n=\frac{\left(q,\dfrac{b}{a},ax,\dfrac{q}{ax};q\right)_\infty}{\left(b,\dfrac{q}{a},x,\dfrac{b}{ax};q\right)_\infty},$$

for $\left|\frac{b}{a}\right| < |x| < 1$. Using Lemma 6.1.2, this can be rewritten

$$\sum_{k=-\infty}^{\infty} \frac{(bq^k;q)_\infty (a;q)_\infty}{(aq^k;q)_\infty (b;q)_\infty} x^k = \frac{\left(q, \frac{b}{a}, ax, \frac{q}{ax}; q\right)_\infty}{\left(b, \frac{q}{a}, x, \frac{b}{ax}; q\right)_\infty}$$

$$\sum_{k=-\infty}^{\infty} \frac{(bq^k;q)_\infty}{(aq^k;q)_\infty} x^k = \frac{\left(q, \frac{b}{a}, ax, \frac{q}{ax}; q\right)_\infty}{\left(a, \frac{q}{a}, x, \frac{b}{ax}; q\right)_\infty}$$

Setting $b = 0$, $x = q^\beta$ and noting that $(0;q)_\infty = 1$ gives the result. $\square$

With this formula a different form for the constant A can be given.

$$A = \frac{\left(-(1-q)q^{\alpha+1}, \frac{q^{-\alpha}}{-(1-q)}, q; q\right)_\infty}{\left(q^{\alpha+1}, -(1-q), \frac{q}{-(1-q)}; q\right)_\infty}.$$

Theorem 6.3.2 (cf. [38], p. 25)**.** *The n^{th} moment μ_n associated with the q-Laguerre polynomials when it is stipluated that $\mu_0 = 1$ is given by*

$$\mu_n = \frac{(q^{\alpha+1};q)_n q^{-\alpha n - \binom{n+1}{2}}}{(1-q)^n}. \tag{6.3.2}$$

Proof. The n^{th} moment is given by

$$\begin{aligned}
&\frac{1}{A(1-q)} \int_0^\infty x^n x^\alpha e_q(-(1-q)x) d_q x \\
&\quad = \sum_{k=-\infty}^{\infty} \frac{(q^k)^n (q^k)^\alpha q^k}{A(-1(1-q)q^k;q)_\infty} \\
&\quad = \frac{1}{A} \frac{\left(-(1-q)q^{n+\alpha+1}, \frac{q^{-n-\alpha}}{-(1-q)}, q; q\right)_\infty}{\left(q^{n+\alpha+1}, -(1-q), \frac{q}{-(1-q)}; q\right)_\infty} \\
&\quad = \frac{\left(q^{\alpha+1}, -(1-q), -\frac{q}{(1-q)}, -(1-q)q^{\alpha+n+1}, \frac{-q^{-\alpha-n}}{(1-q)}, q; q\right)_\infty}{\left(-(1-q)q^{\alpha+1}, \frac{-q^{-\alpha}}{(1-q)}, q, q^{\alpha+n+1}, -(1-q), -\frac{q}{1-q}; q\right)_\infty},
\end{aligned}$$

where (6.3.1) was used in the last step. Using Lemma 6.1.2 and cancelling like terms in numerator and denominator this reduces to

$$\frac{\left(q^{\alpha+1}, -\dfrac{q^{-n-\alpha}}{(1-q)}; q\right)_n}{(-(1-q)q^{\alpha+1}; q)_n}.$$

In Lemma 6.1.4 set $a = -q^{\alpha+1}(1-q)$ to get

$$\frac{(q^{\alpha+1}; q)_n q^{-\alpha n - \binom{n+1}{2}}}{(1-q)^n},$$

which is what was required. □

Lemma 6.3.3 (cf. [7], p.353)**.**

$$\int_0^\infty x^\alpha e_q(-(1-q)x)dx = \frac{\Gamma(-\alpha)\Gamma(\alpha+1)}{\Gamma_q(-\alpha)}, \quad 0 < q < 1, \Re(\alpha) > -1.$$

Let $x \in [0, \infty)$ and $0 < q < 1$, then $(1-q)x > 0$ and

$$\begin{aligned} e_q(-(1-q)x) &= \frac{1}{(-(1-q)x; q)_\infty} \\ &= \frac{1}{\prod_{k=0}^{\infty}(1 - (-(1-q)x)q^k)} \\ &= \frac{1}{\prod_{k=0}^{\infty}(1 + (1-q)x)q^k} > 0 \end{aligned}$$

Let $\alpha > -1$ and $x \in [0, \infty)$ then $x^\alpha > 0$. Consequently the function $\alpha(x)$ defined by the integral

$$\alpha(x) = \frac{\Gamma_q(-\alpha)}{\Gamma(-\alpha)\Gamma(\alpha+1)} \int_0^x x^\alpha e_q(-(1-q)x)dx,$$

where $\alpha > -1$, $0 < q < 1$, exists because the integrand is continuous and positive and the integral is bounded above by

$$\frac{\Gamma_q(-\alpha)}{\Gamma(-\alpha)\Gamma(\alpha+1)} \int_0^\infty x^\alpha e_q(-(1-q)x)dx = 1.$$

Because the integrand is positive $\alpha(x)$ is non-decreasing.

Theorem 6.3.4 (cf. [38], p. 24)**.** *The n^{th} moment of $\alpha(x)$ is given by*

$$\mu_n = \frac{(q^{\alpha+1}; q)_n q^{-\alpha n - \binom{n+1}{2}}}{(1-q)^n}.$$

Proof. First note that

$$n + \binom{n}{2} = n + \frac{n(n-1)}{2} = \frac{n^2 + 2n - n}{2} = \binom{n+1}{2}.$$

By Lemma 6.3.3,

$$\begin{aligned}\mu_n &= \frac{\Gamma_q(-\alpha)}{\Gamma(-\alpha)\Gamma(\alpha+1)} \int_0^\infty x^{\alpha+n} e_q(-(1-q)x)dx \\ &= \frac{\Gamma_q(-\alpha)\Gamma(-\alpha-n)\Gamma(\alpha+n+1)}{\Gamma_q(-\alpha-n)\Gamma(-\alpha)\Gamma(\alpha+1)}.\end{aligned}$$

Lemma 6.1.13 gives

$$\frac{\Gamma_q(-\alpha)}{\Gamma_q(-\alpha-n)} = \frac{(q^{-\alpha-n}; q)_n}{(1-q)^n}.$$

Using $a = q^{\alpha+1}$ in Lemma 6.1.4 then gives

$$\frac{\Gamma_q(-\alpha)}{\Gamma_q(-\alpha-n)} = \frac{(-1)^n (q^{\alpha+1}; q)_n q^{-\alpha n - \binom{n+1}{2}}}{(1-q)^n}.$$

The reflection formula for the gamma function gives

$$\begin{aligned}\frac{\Gamma(-\alpha-n)\Gamma(\alpha+n+1)}{\Gamma(-\alpha)\Gamma(\alpha+1)} &= \frac{\pi}{\sin(-\pi\alpha)} \frac{\sin(-\pi\alpha - \pi n)}{\pi} \\ &= \frac{(-1)^n \sin(-\pi\alpha)}{\sin(-\pi\alpha)}.\end{aligned}$$

Combining the above gives

$$\mu_n = \frac{(q^{\alpha+1}; q)_n q^{-\alpha n - \binom{n+1}{2}}}{(1-q)^n}.$$ □

This establishes that $\alpha(x)$ is a distribution function. Futhermore the moments of $\alpha(x)$ are identical with the moments associated with the q-Laguerre polynomials given in (6.3.2). As a result the q-Laguerre polynomials are orthogonal with respect to at least two distributions which can't be equivalent because one is continuous and the other discrete (cf. [38], pp.24, 25). The moment problem associated with the q-Laguerre polynomials is indeterminate (cf. [7], p.354).

Denote by $\psi(x)$ the distribution

$$\psi(x) = C \int_0^x x^\alpha e_q(-(1-q)x)dx$$

$$C = \frac{\Gamma_q(-\alpha)}{\Gamma(-\alpha)\Gamma(\alpha+1)}.$$

$\psi(x)$ is a solution of a Stieltjes moment problem and an associated symmetric distribution $\alpha(x)$ can be constructed via $\alpha(x) = \operatorname{sgn}(x)\psi(x^2)$. Departing slightly from the notation in the discussion of symmetric moment problems the polynomial that was denoted by $P_n(x)$ will be denoted by $p_n(x)$. Then $p_n(x) = C_1 L_n^\alpha(x;q)$, where C_1 is chosen so that $p_n(x)$ is monic. The polynomials $K_n(x)$ are the monic polynomials orthogonal with respect to $xd\psi(x)$. These polynomials are just $C_2 L_n^{\alpha+1}(x;q)$, where C_2 is chosen for the monic normalisation. The monic symmetric polynomials $S_n(x)$ orthogonal with respect to $\alpha(x)$ are then obtained from the relations

$$S_{2n}(x) = p_n(x^2)$$
$$S_{2n+1}(x) = xK_n(x^2).$$

The recurrence for the q-Laguerre polynomials is given in Theorem 6.2.5. This is not the monic or the orthonormal form of the three-term recurrence. Recall that the leading coefficient of $L_n^\alpha(x;q)$ is

$$k_n = \frac{(q^{-n};q)_n q^{\binom{n+1}{2}+n\alpha+n^2}(1-q)^n}{(q,q;q)_n},$$

and that

$$(q^{-n};q)_n = (q;q)_n(-1)^n q^{-\binom{n+1}{2}},$$

so that

$$k_n = \frac{q^{n(n+\alpha)}(q-1)^n}{(q;q)_n}.$$

This gives

$$p_n(x) = \frac{(q;q)_n}{q^{n(n+\alpha)}(q-1)^n} L_n^{\alpha}(x;q)$$
$$K_n(x) = \frac{(q;q)_n}{q^{n(n+\alpha+1)}(q-1)^n} L_n^{\alpha+1}(x;q).$$

The recurrence relation for the polynomials $\{p_n(x)\}_{n=0}^{\infty}$ is then

$$x\frac{q^{n(n+\alpha)}(q-1)^n}{(q;q)_n} p_n(x) = -\frac{(1-q^{n+1})}{(1-q)q^{2n+\alpha+1}} \frac{q^{n^2+2n+n\alpha+\alpha+1}(q-1)^{n+1}}{(q;q)_{n+1}} p_{n+1}(x)$$
$$+ \frac{(1+q-q^{n+1}-q^{n+\alpha+1})}{(1-q)q^{2n+\alpha+1}} \frac{q^{n(n+\alpha)}(q-1)^n}{(q;q)_n} p_n(x)$$
$$- \frac{(1-q^{n+\alpha})}{(1-q)q^{2n+\alpha}} \frac{q^{n^2-2n+n\alpha-\alpha+1}(q-1)^{n-1}}{(q;q)_{n-1}} p_{n-1}(x).$$

Dividing the above by

$$\frac{q^{n(n+\alpha)}(q-1)^n}{(q;q)_n},$$

gives (cf. [30], p.159)

$$xp_n(x) = p_{n+1}(x) + \frac{(1+q-q^{n+1}-q^{n+\alpha+1})}{(1-q)q^{2n+\alpha+1}} p_n(x)$$
$$+ \frac{(1-q^n)(1-q^{n+\alpha})}{(1-q)^2 q^{4n+2\alpha-1}} p_{n-1}(x), \tag{6.3.3}$$
$$p_0(x) = 1, \quad p_1(x) = x - \frac{(1-q^{\alpha+1})}{(1-q)q^{\alpha+1}}. \tag{6.3.4}$$

This is the monic form of the recurrence relation i.e.

$$p_{n+1}(x) = (x-e_n)p_n(x) - d_{n-1}^2 p_{n-1}(x).$$

and immediately gives the sequences associated with the orthonormal polynomials

$$d_n = \sqrt{\frac{(1-q^n)(1-q^{n+\alpha})}{(1-q)^2 q^{4n+2\alpha-1}}},$$
$$e_n = \frac{(1+q-q^{n+1}-q^{n+\alpha+1})}{(1-q)q^{2n+\alpha+1}}.$$

The polynomials $\{p_n(x)\}_{n=0}^{\infty}$ are the denominators of a Jacobi continued fraction. The associated numerators $\{q_n(x)\}_{n=0}^{\infty}$ are the polynomials which satisfy (6.3.4), but with initial conditions

$$q_0(x) = 0, \quad q_1(x) = 1.$$

From the sets $\{q_n(x)\}_{n=0}^{\infty}$ and $\{p_n(x)\}_{n=0}^{\infty}$ the orthonormal polynomials $\{P_n(x)\}_{n=0}^{\infty}$ and $\{Q_n(x)\}_{n=0}^{\infty}$ can be obtained using (2.6.8)

$$P_n(x) = \frac{1}{d_0 d_1 \dots d_{n-1}} p_n(x)$$

$$Q_n(x) = \frac{1}{d_0 d_1 \dots d_{n-1}} q_n(x).$$

Recall the polynomials used in deriving the Nevanlinna parametrisation

$$A_n(z) = d_{n-1}\{Q_{n-1}(0)Q_n(z) - Q_n(0)Q_{n-1}(z)\},$$

$$B_n(z) = d_{n-1}\{Q_{n-1}(0)P_n(z) - Q_n(0)P_{n-1}(z)\},$$

$$C_n(z) = d_{n-1}\{P_{n-1}(0)Q_n(z) - P_n(0)Q_{n-1}(z)\},$$

$$D_n(z) = d_{n-1}\{P_{n-1}(0)P_n(z) - P_n(0)P_{n-1}(z)\}.$$

Theorem 6.3.5 ([38], p.36)**.** *Let*

$$L_\infty^\alpha(x;q) = \lim_{n\to\infty} L_n^\alpha(x;q),$$

then $L_\infty^\alpha(x;q)$ is an entire function and is given by

$$L_\infty^\alpha(x;q) = \frac{(q^{\alpha+1};q)_\infty}{(q;q)_\infty} \sum_{k=0}^{\infty} \frac{q^{k^2+\alpha k}(1-q)^k(-x)^k}{(q^{\alpha+1};q)_k (q;q)_k}.$$

The functions $A(z)$, $B(z)$, $C(z)$ and $D(z)$ have been computed and are given by (cf.[30], pp.163-166)

$$A(z) = -\frac{1-q}{(q^\alpha,-z;q)_\infty}\{\sum_{n=0}^{\infty} \frac{z^n q^{\frac{n(n+1)}{2}}}{(q;q)_n(1-q^{n-\alpha})}$$
$$\big\{(q^a;q)_\infty L_\infty^\alpha(z;q) + ((q;q)_\infty - (q^{\alpha+1};q)_\infty) L_\infty^{\alpha-1}(z;q)\big\}$$
$$- \frac{(q^\alpha;q)_\infty}{(q^{-\alpha};q)_\infty} L_\infty^{-\alpha}(z;q) - ((q;q)_\infty - (q^{\alpha+1};q)_\infty)\frac{(q;q)_\infty}{(q^{-\alpha};q)_\infty} L_\infty^{1-\alpha}(z;q)\}$$

$$
\begin{aligned}
B(z) &= \left[\frac{1}{1-q^{\alpha}} - \frac{(q;q)_\infty}{(q^\alpha;q)_\infty}\right] L_\infty^{\alpha-1}(z;q) - L_\infty^{\alpha}(z;q) \\
C(z) &= \frac{z}{(-z;q)_\infty} L_\infty^{\alpha+1}(z;q) \sum_{n=0}^{\infty} \frac{z^n q^{\frac{n(n+1)}{2}}}{(q;q)_n(1-q^{n-\alpha})} \\
&\quad + \frac{(q;q)_\infty}{(-z,q^{-\alpha};q)_\infty} L_\infty^{\alpha-1}(z;q) \\
D(z) &= z L_\infty^{\alpha+1}(z;q).
\end{aligned}
$$

Bibliography

[1] N. I. Akhiezer, *The Classical Moment Problem*, Oliver and Boyd, 1965.

[2] N. I. Akhiezer, I. M. Glazman *Theory of Linear Operators in Hilbert Space*, volume 1, Frederick Ungar Publishing Company, New York, 1961, (Dover reprint 1993).

[3] N. I. Akhiezer, I. M. Glazman *Theory of Linear Operators in Hilbert Space*, volume 2, Frederick Ungar Publishing Company, New York, 1963, (Dover reprint 1993).

[4] N. I. Akhiezer, M. G. Krein, *Some Questions in the Theory of Moments*, American Mathematical Society, Providence, Rhode Island, 1962.

[5] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Cambridge University Press, Cambridge, 1999.

[6] H. Anton, C. Rorres, *Elementary Linear Algebra*, John Wiley and Sons, 9th edition, 2005.

[7] R. Askey, Ramanujan's extension of the gamma and beta functions, *Amer. Math. Monthly*, 87(5) (1980), 346-359.

[8] C. Berg, J. Christensen, Density questions in the classical theory of moments, *Ann. Inst. Fourier (Grenoble)*, 31(3), (1981), 99-114.

[9] M. Barnsley, J. Geronimo, A. Harrington, Condensed Julia sets with an application to a fractal lattice model hamiltonian, *Trans. Amer. Math. Soc.*, 288(2), (1985), 537-561.

[10] K. Case, J. Geronimo, Scattering theory and orthogonal polynomials on the real line, *Trans. Amer. Math. Soc.*, 258(2), (1980), 467-494.

[11] P. Chebyshev, *Ouvres*, Chelsea Publishing Company, New York (two volumes), 1961.

[12] T. Chihara, Chain sequences and orthogonal polynomials, *Trans. Amer. Math. Soc.*, 104(1), (1962), 1-16.

[13] T. Chihara, Hamburger moment problems and orthogonal polynomials, *Trans. Amer. Math. Soc.*, 315(1), (1989), 189-203.

[14] T. Chihara, Indeterminate symmetric moment problems, *J. Math. Anal. Appl.*, 85, (1982) 331-346.

[15] T. Chihara, *Introduction to Orthogonal Polynomials*, Gordon and Breach, 1978.

[16] T. Chihara, On co-recursive orthogonal polynomials, *Proc. Amer. Math. Soc.*, 8(5), (1957), 899-905.

[17] T. Chihara, Non-linear recurrence relations for classical orthogonal polynomials, *Amer. Math. Monthly*, 65(3), (1958), 195-197.

[18] T. Chihara, The parameters of a chain sequence, *Proc. Amer. Math. Soc.*, 108(3), (1990), 775-780.

[19] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, American Mathematical Society, Providence, Rhode Island, 1999 (Reprint 2000).

[20] G. Gasper, M. Rahman, *Basic Hypergeometric Series*, Cambridge, 1997.

[21] J. Geronimo, A relation between the coefficients in the recurrence formula and the spectral function for orthogonal polynomials, *Trans. Amer. Math. Soc.*, 260(1), (1980) 65-82.

[22] J. Geronimus, *Orthogonal Polynomials: Estimates, Asymptotic Formulae and Series of Polynomials Orthogonal on the Unit Circle and on an Interval*, Consultants Bureau, New York, 1961.

[23] J. Geronimus, On the trigonometric moment problem, *Ann. of Math. (2)*, 47(4), (1946) 742-761.

[24] F. Gesztesy, B. Simon, m-Functions and inverse speactral analysis for finite and semi-infinite Jacobi matrices, *J. Anal. Math.*, 73 (1997) 267-297.

[25] R. Gilman, A class of functions continuous but not absolutely continuous, *Ann. of Math. (2)*, 33(3), (1932), 433-442.

[26] H. Hamburger, Ueber eine erweiterung des Stieltjes'schen momentenproblems, *Math. Ann.*, 81, (1920), 235-319, 82, (1921), 120-164, 168-187.

[27] J. Hanna, J. Rowland, *Fourier Series, Transforms, and Boundary Value Problems*, Dover, 2008.

[28] F. Hausdorff, Momentprobleme für ein endliches intervall, *Math. Z.*, 16, (1923), 220-248.

[29] H. Horst, Riemann-Stieltjes and Lebesgue-Stieltjes integrability, *Amer. Math. Monthly*, 91(9), (1984), 551-559.

[30] M. Ismail, G. Rahman, The q-Laguerre polynomials and related moment problems, *J. Math. Anal. Appl.*, 218, (1998) 155-174.

[31] R. Kannan, C. Krueger, *Advanced Analysis on the Real Line*, Springer, New York, 1996.

[32] S. Karlin, J. McGregor, The classification of birth and death processes, *Trans. Amer. Math. Soc.*, 86(2) (1957) 366-400.

[33] S. Karlin, J. McGregor, The differential equations of birth and death processes and the Stieltjes moment problem, *Trans. Amer. Math. Soc.*, 85(2), (1957) 489-546.

[34] S. Khrushchev, Continued fractions and orthogonal polynomials on the unit circle, *J. Comput. Appl. Math.*, 178, (2005) 267-303.

[35] S. Khrushchev, *Orthogonal Polynomials and Continued Fractions*, Cambridge University Press, Cambridge, 2008.

[36] S. Khrushchev, Schur's algorithm, orthogonal polynomials and convergence of Wall's continued fractions in $L^2(\mathbb{T})$, *J. Approx. Theory*, 108, (2001), 161-248.

[37] D. Lubinsky, Singularly continuous measures in Nevai's class M, *Proc. Amer. Math. Soc.*, 111(2), (1991), 413-420.

[38] D. Moak, The q-analogue of the Lageurre polynomials, *J. Math. Anal. Appl.* 81, (1981) 20-47.

[39] F. Olver, *Asymptotics and Special Functions*, Academic Press, New York, 1976.

[40] J. Oxtoby, *Measure and Category*, Springer-Verlag, 1971.

[41] G. Pólya, G. Szegö, *Problems and Theorems in Analysis*, volume 2, Berlin, 1976.

[42] E. D. Rainville, *Special Functions*, MacMillan, 1960.

[43] E. D. Rainville, *Intermediate Differential Equations*, MacMillan, 1964.

[44] F. Riesz, B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar Publishing Company, New York, 1955 (Dover reprint 1990).

[45] J. Shohat, J. Tamarkin *The Problem of Moments*, American Mathematical Society, Providence, Rhode Island, 1943.

[46] G. Shilov, B. Gurevich, *Integral Measure and Derivative*, Prentice-Hall, 1966, Englewood Cliffs, New Jersey.

[47] B. Simon, The classical moment problem as a self-adjoint finite difference operator, *Adv. Math.*, 137, (1998), 82-203.

[48] T. Stieltjes, Reserches sur les fractions continues, *Anns. Fac. Sci. Univ. Toulouse*, 8, (1894), J1-J122, 9, (1895), A5-A47.

[49] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, 2003 (reprint).

[50] N. Temme, *Special Functions*, John Wiley and Sons, 1996.

[51] D. Varberg, On absolutely continuous functions, *Amer. Math. Monthly*, 72(8) (1965), 831-841.

[52] H. Wall, *Continued Fractions*, D. Van Nostrand Company, 1948.

[53] H. Wall, Continued fractions and totally monotone sequences, *Trans. Amer. Math. Soc.*, 48(2), (1940), 165-184.

[54] H. Wall, Contributions to the analytic theory of J-fractions, *Trans. Amer. Math. Soc.*, 55(3), (1944), 373-392.

[55] E. Whitakker, G. Watson, *Modern Analysis*, Merchant Books, 2008 (reprint of 1915 2nd edition).