Encyclopedia of Mathematics and its Applications 98

CLASSICAL AND QUANTUM ORTHOGONAL POLYNOMIALS IN ONE VARIABLE

Mourad E. H. Ismail

CAMBRIDGE

Volume 98
Classical and Quantum Orthogonal Polynomials in One Variable

This is first modern treatment of orthogonal polynomials from the viewpoint of special functions. The coverage is encyclopedic, including classical topics such as Jacobi, Hermite, Laguerre, Hahn, Charlier and Meixner polynomials as well as those, e.g. Askey–Wilson and Al-Salam–Chihara polynomial systems, discovered over the last 50 years: multiple orthogonal polynomials are dicussed for the first time in book form. Many modern applications of the subject are dealt with, including birth and death processes, integrable systems, combinatorics, and physical models. A chapter on open research problems and conjectures is designed to stimulate further research on the subject.

Exercises of varying degrees of difficulty are included to help the graduate student and the newcomer. A comprehensive bibliography rounds off the work, which will be valued as an authoritative reference and for graduate teaching, in which role it has already been successfully class-tested.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

All the titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing visit
http://publishing.cambridge.org/stm/mathematics/eom/

88. Teo Mora *Solving Polynomial Equation Systems I*
89. Klaus Bichteler *Stochastic Integration with Jumps*
90. M. Lothaire *Algebraic Combinatorics on Words*
91. A.A. Ivanov & S.V. Shpectorov *Geometry of Sporadic Groups, 2*
92. Peter McMullen & Egon Schulte *Abstract Regular Polytopes*
93. G. Gierz et al. *Continuous Lattices and Domains*
94. Steven R. Finch *Mathematical Constants*
95. Youssef Jabri *The Mountain Pass Theorem*
96. George Gasper & Mizan Rahman *Basic Hypergeometric Series 2nd ed.*
97. Maria Cristina Pedicchio & Walter Tholen *Categorical Foundations*
98. Mourad Ismail *Classical and Quantum Orthogonal Polynomials in One Variable*
99. Teo Mora *Solving Polynomial Equation Systems II*
100. Enzo Olivieri & Maria Eulalia Vares *Large Deviations and Metastability*
101. A. Kushner, V. Lychagin & V. Roubtsov *Contact Geometry and Nonlinear Differential Equations*
102. R.J. Wilson & L. Beineke *Topics in Algebraic Graph Theory*

Classical and Quantum Orthogonal Polynomials in One Variable

Mourad E.H. Ismail

University of Central Florida

With two chapters by

Walter Van Assche

Catholic University of Leuven

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press,
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

www.cambridge.org
Information on this title: www.cambridge.org/9780521782012

First published 2005

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10/13pt *System* LaTeX 2_ε [AUTHOR]

A catalogue record for this book is available from the British Library

ISBN-13 978–0–521–78201–2 hardback
ISBN-10 0–521–78201–5 hardback

Contents

Foreword

There are a number of ways of studying orthogonal polynomials. Gabor Szegő's book "Orthogonal Polynomials" (Szegő, 1975) had two main topics. Most of this book dealt with polynomials which are orthogonal on the real line, with a chapter on polynomials orthogonal on the unit circle and a short chapter on polynomials orthogonal on more general curves. About two-thirds of Szegő's book deals with the classical orthogonal polynomials of Jacobi, Laguerre and Hermite, which are orthogonal with respect to the beta, gamma and normal distributions, respectively. The rest deals with more general sets of orthogonal polynomials, some general theory, and some asymptotics.

Barry Simon has recently written a very long book on polynomials orthogonal on the unit circle, (Simon, 2004). His book has very little on explicit examples, so its connection with Szegő's book is mainly in the general theory, which has been developed much more deeply than it had been in 1938 when Szegő's book appeared.

The present book, by Mourad Ismail, complements Szegő's book in a different way. It primarily deals with specific sets of orthogonal polynomials. These include the classical polynomials mentioned above and many others. The classical polynomials of Jacobi, Laguerre and Hermite satisfy second-order linear homogeneous differential equations of the form

$$a(x)y''(x) + b(x)y'(x) + \lambda_n y(x) = 0$$

where $a(x)$ and $b(x)$ are polynomials of degrees 2 and 1, respectively, which are independent of n, and λ_n is independent of x. They have many other properties in common. One is that the derivative of $p_n(x)$ is a constant times $q_{n-1}(x)$ where $\{p_n(x)\}$ is in one of these classes of polynomials and $\{q_n(x)\}$ is also. These are the only sets of orthogonal polynomials with the property that their derivatives are also orthogonal.

Many of the classes of polynomials studied in this book have a similar nature, but with the derivative replaced by another operator. The first operator which was used is

$$\Delta f(x) = f(x+1) - f(x),$$

a standard form of a difference operator. Later, a q-difference operator was used

$$D_q f(x) = [f(qx) - f(x)]/[qx - x].$$

Still later, two divided difference operators were introduced. The orthogonal polynomials which arise when the q-divided difference operator is used contain a set of polynomials introduced by L. J. Rogers in a remarkable series of papers which appeared in the 1890s. One of these sets of polynomials was used to derive what we now call the Rogers–Ramanujan identities. However, the orthogonality of Rogers's polynomials had to wait decades before it was found. Other early work which leads to polynomials in the class of these generalized classical orthogonal polynomials was done by Chebyshev, Markov and Stieltjes.

To give an idea about the similarities and differences of the classical polynomials and some of the extensions, consider a set of polynomials called ultraspherical or Gegenbauer polynomials, and the extension Rogers found. Any set of polynomials which is orthogonal with respect to a positive measure on the real line satisfies a three term recurrence relation which can be written in a number of equivalent ways. The ultraspherical polynomials $C_n^{\nu}(x)$ are orthogonal on $(-1,1)$ with respect to $\left(1-x^2\right)^{\nu-1/2}$. Their three-term recurrence relation is

$$2\,(n+\nu)\,xC_n^{\nu}(x) = (n+1)\,C_{n+1}^{\nu}(x) + (n+2\nu-1)\,C_{n-1}^{\nu}(x)$$

The three-term recurrence relation for the continuous q-ultraspherical polynomials of Rogers satisfy a similar recurrence relation with every $(n+a)$ replaced by $1-q^{n+a}$. That is a natural substitution to make, and when the recurrence relation is divided by $1-q$, letting q approach 1 gives the ultraspherical polynomials in the limit.

Both of these sets of polynomials have nice generating functions. For the ultraspherical polynials one nice generating function is

$$\left(1-2xr+r^2\right)^{-\nu} = \sum_{n=0}^{\infty} C_n^{\nu}(x)\,r^n$$

The extension of this does not seem quite as nice, but when the substitution $x = \cos\theta$ is used on both, they become similar enough for one to guess what the left-hand side should be. Before the substitution it is

$$\prod_{n=0}^{\infty} \frac{\left(1-2xq^{\nu+n}r+q^{2\nu+2n}r^2\right)}{\left(1-2xq^n r+q^{2n}r^2\right)} = \sum_{n=0}^{\infty} C_n(x;q^{\nu}\,|\,q)\,r^n.$$

The weight function is a completely different story. To see this, it is sufficient to state it:

$$w\left(x,q^{\nu}\right) = \left(1-x^2\right)^{-1/2} \prod_{n=0}^{\infty} \frac{\left(1-\left(2x^2-1\right)q^n+q^{2n}\right)}{\left(1-\left(2x^2-1\right)q^{n+\nu}+q^{2n+2\nu}\right)}.$$

These polynomials of Rogers were rediscovered about 1940 by two mathematicians, (Feldheim, 1941b) and (Lanzewizky, 1941). Enough had been learned about orthogonal polynomials by then for them to know they had sets of orthogonal polynomials, but neither could find the orthogonality relation. One of these two mathematicians, E. Feldheim, lamented that he was unable to find the orthogonality relation. Stieltjes and Markov had found theorems which would have allowed Felheim to work out the orthogonality relation, but there was a war going on when Feldheim did his work and he was unaware of this old work of Stieltjes and Markov. The limiting case when

$\nu \to \infty$ gives what are called the continuous q-Hermite polynomials. It was these polynomials which Rogers used to derive the Rogers-Ramanujan identities.

Surprisingly, these polynomials have recently come up in a very attractive problem in probability theory which has no q in the statement of the problem. See Bryc (Bryc, 2001) for this work.

Stieltjes solved a minimum problem which can be considered as coming from one dimensional electrostatics, and in the process found the discriminant for Jacobi polynomials. The second-order differential equation they satisfy played an essential role. When I started to study special functions and orthogonal polynomials, it seemed that the only orthogonal polynomials which satisfied differential equations nice enough to be useful were Jacobi, Laguerre and Hermite. For a few classes of orthogonal polynomials nice enough differential equations existed, but they were not well known. Now, thanks mainly to a conjecture of G. Freud which he proved in two very special cases, and work by quite a few people including Nevai and some of his students, we know that nice enough differential equations exist for polynomials orthogonal with respect to $\exp(-v(x))$ when $v(x)$ is smooth enough. The work of Stieltjes can be partly extended to this much wider class of orthogonal polynomials. Some of this is done in Chapter 3.

Chapter 4 deals with the classical polynomials. For Hermite polynomials there is an explicit expression for the analogue of the Poisson kernel for Fourier series which was found by Mehler in the 19th century. An important multivariable extension of this formula found independently by Kibble and Slepian is in Chapter 4. Chapter 5 contains some information about the Pollaczek polynomials on the unit interval. Their recurrence relation is a slight variant of the one for ultraspherical polynomials listed above. The weight function is drastically different, having infinitely many point masses outside the interval where the absolutely continuous part is supported or vanishing very rapidly at one or both of the end points of the interval supporting the absolutely continuous part of the orthogonality measure.

Chapter 6 deals with extensions of the classical orthogonal polynomials whose weight function is discrete. Here the classical discriminant seemingly cannot be found in a useful form, but a variant of it has been computed for the Hahn polynomials. This extends the result of Stieltjes on the discriminant for Jacobi polynomails. Hahn polynomials extend Jacobi polynomials and are orthogonal with respect to the hypergeometric distribution. Transformations of them occur in the quantum theory of angular momentum and they and their duals occur in some settings of coding theory.

The polynomials considered in the first 10 chapters which have explicit formulas are given as generalized hypergeometric series. These are series whose term ratio is a rational function of n. In Chapters 11 to 19 a different setting occurs, that of basic hypergeometric series. These are series whose term ratio is a rational function of q^n.

In the 19th century Markov and Stieltjes found examples of orthogonal polynomials which can be written as basic hypergeometric series and found an explicit orthogonality relation. As mentioned earlier, Rogers also found some polynomials which are orthogonal and can be given as basic hypergeometric series, but he was unaware they were orthogonal. A few other examples were found before Wolfgang Hahn wrote a major paper, (Hahn, 1949b) in which he found basic hypergeometric extensions of the classical polynomials and the discrete ones up to the Hahn polynomial level. There is one level higher than this where orthogonal polynomials exist which have properties very similar to many of those known for the classical orthogonal polynomials. In particular, they satisfy a second-order divided q-difference equation and this divided q-difference operator applied to them gives another set of orthogonal polynomials. When this was first published, the polynomials were treated directly without much motivation. Here simpler cases are done first and then a boot-strap argument allows one to obtain more general polynomials, eventually working up to the most general classical type sets of orthogonal polynomials.

The most general of these polynomials has four free parameters in addition to the q of basic hypergeometric series. When three of the parameters are held fixed and the fourth is allowed to vary, the coefficients which occur when one is expanded in terms of the other are given as products. The resulting identity contains a very important transformation formula between a balanced ${}_4\phi_3$ and a very-well-poised ${}_8\phi_7$ which Watson found in the 1920s as the master identity which contains the Rogers-Ramanujan identities as special cases and many other important formulas. There are many ways to look at this identity of Watson, and some of these ways lead to interesting extensions. When three of the four parameters are shifted and this connection problem is solved, the coefficients are single sums rather than the double sums which one expects. At present we do not know what this implies, but surprising results are usually important, even if it takes a few decades to learn what they imply.

The fact that there are no more classical type polynomials beyond those mentioned in the last paragraph follows from a theorem of Leonard (Leonard, 1982). This theorem has been put into a very attractive setting by Terwilliger, some of whose work has been summarized in Chapter 20. However, that is not the end since there are biorthogonal rational functions which have recently been discovered. Some of this work is contained in Chapter 18. There is even one higher level than basic hypergeometric functions, elliptic hypergeometric functions. Gasper and Rahman have included a chapter on them in (Gasper & Rahman, 2004).

Chapters 22 and 23 were written by Walter Van Assche. The first is on the Riemann-Hilbert method of studying orthogonal polynomials. This is a very powerful method for deriving asymptotics of wide classes of orthogonal polynomials. The second chapter is on multiple orthogonal polynomials. These are polynomials in one variable which are orthogonal with respect to r different measures. The basic ideas go back to the 19th century, but except for isolated work which seems to start with Angelesco in 1919, it has only been in the last 20 or so years that significant work has been done on them.

There are other important results in this book. One which surprised me very much is the q-version of Airy functions, at least as the two appear in asymptotics. See, for example, Theorem 21.7.3.

When I started to work on orthogonal polynomials and special functions, I was told by a number of people that the subject was out-of-date, and some even said dead. They were wrong. It is alive and well. The one variable theory is far from finished, and the multivariable theory has grown past its infancy but not enough for us to be able to predict what it will look like in 2100.

Madison, WI
April 2005

Richard A. Askey

Preface

I first came across the subject of orthogonal polynomials when I was a student at Cairo University in 1964. It was part of a senior-level course on special functions taught by the late Professor Foad M. Ragab. The instructor used his own notes, which were very similar in spirit to the way Rainville treated the subject. I enjoyed Ragab's lectures and, when I started graduate school in 1968 at the Univerity of Alberta, I was fortunate to work with Waleed Al-Salam on special functions and q-series. Jerry Fields taught me asymptotics and was very generous with his time and ideas. In the late 1960s, courses in special functions were a rarity at North American universities and have been replaced by Bourbaki-type mathematics courses. In the early 1970s, Richard Askey emerged as the leader in the area of special functions and orthogonal polynomials, and the reader of this book will see the enormous impact he made on the subject of orthogonal polynomials. At the same time, George Andrews was promoting q-series and their applications to number theory and combinatorics. So when Andrews and Askey joined forces in the mid-1970s, their combined expertise advanced the subject in leaps and bounds. I was very fortunate to have been part of this group and to participate in these developments. My generation of special functions / orthogonal polynomials people owes Andrews and Askey a great deal for their ideas which fueled the subject for a while, for the leadership role they played, and for taking great care of young people.

This book project started in the early 1990s as lecture notes on q-orthogonal polynomials with the goal of presenting the theory of the Askey–Wilson polynomials in a way suitable for use in the classroom. I taught several courses on orthogonal polynomials at the University of South Florida from these notes, which evolved with time. I later realized that it would be better to write a comprehensive book covering all known systems of orthogonal polynomials in one variable. I have attempted to include as many applications as possible. For example, I included treatments of the Toda lattice and birth and death processes. Applications of connection relations for q-polynomials to the evaluation of integrals and the Rogers–Ramanujan identities are also included. To the best of my knowledge, my treatment of associated orthogonal polynomials is a first in book form. I tried to include all systems of orthogonal polynomials but, in order to get the book out in a timely fashion, I had to make some compromises. I realized that the chapters on Riemann–Hilbert problems and multiple orthogonal polynomials should be written by an expert on the subject, and

Walter Van Assche kindly agreed to write this material. He wrote Chapters 22 and 23, except for §22.8. Due to the previously mentioned time constraints, I was unable to treat some important topics. For example, I covered neither the theories of matrix orthogonal polynomials developed by Antonio Durán, Yuan Xu and their collaborators, nor the recent interesting explicit systems of Grünbaum and Tirao and of Durán and Grünbaum. I hope to do so if the book has a second edition. Regrettably, neither the Sobolov orthogonal polynomials nor the elliptic biorthogonal rational functions are treated.

Szegő's book on orthogonal polynomials inspired generations of mathematicians. The character of this volume is very different from Szegő's book. We are mostly concerned with the special functions aspects of orthogonal polynomials, together with some general properties of orthogonal polynomial systems. We tried to minimize the possible overlap with Szegő's book. For example, we did not treat the refined bounds on zeros of Jacobi, Hermite and Laguerre polynomials derived in (Szegő, 1975) using Sturmian arguments. Although I tried to cover a broad area of the subject matter, the choice of the material is influenced by the author's taste and personal bias.

Dennis Stanton has used parts of this book in a graduate-level course at the University of Minnesota and kindly supplied some of the exercises. His careful reading of the book manuscript and numerous corrections and suggestions are greatly appreciated. Thanks also to Richard Askey and Erik Koelink for reading the manuscript and providing a lengthy list of corrections and additional information. I am grateful to Paul Terwilliger for his extensive comments on §20.3.

I hope this book will be useful to students and researchers alike. It has a collection of open research problems in Chapter 24 whose goal is to challenge the reader's curiosity. These problems have varying degrees of difficulty, and I hope they will stimulate further research in this area.

Many people contributed to this book directly or indirectly. I thank the graduate students and former graduate students at the University of South Florida who took orthogonal polynomials and special functions classes from me and corrected misprints. In particular, I thank Plamen Simeonov, Jacob Christiansen, and Jemal Gishe. Mahmoud Annaby and Zeinab Mansour from Cairo University also sent me helpful comments. I learned an enormous amount of mathematics from talking to and working with Richard Askey, to whom I am eternally grateful. I am also indebted to George Andrews for personally helping me on many occasions and for his work which inspired parts of my research and many parts of this book. The book by Gasper and Rahman (Gasper & Rahman, 1990) has been an inspiration for me over many years and I am happy to see the second edition now in print (Gasper & Rahman, 2004). It is the book I always carry with me when I travel, and I "never leave home without it." I learned a great deal of mathematics and picked up many ideas from collaboration with other mathematicians. In particular I thank my friends Christian Berg, Yang Chen, Ted Chihara, Jean Letessier, David Masson, Martin Muldoon, Jim Pitman, Mizan Rahman, Dennis Stanton, Galliano Valent, and Ruiming Zhang for the joy of having them share their knowledge with me and for the pleasure of working with them. P. G. (Tim) Rooney helped me early in my career, and was very generous with his time. Thanks, Tim, for all the scientific help and post-doctorate support.

This book was mostly written at the University of South Florida (USF). All the typesettng was done at USF, although during the last two years I was employed by the University of Central Florida. I thank Marcus McWaters, the chair of the Mathematics Department at USF, for his encouragement and continued support which enabled me to complete this book. It is my pleasure to acknowledge the enormous contribution of Denise L. Marks of the University of South Florida. She was always there when I needed help with this book or with any of my edited volumes. On many occasions, she stayed after office hours in order for me to meet deadlines. Working with Denise has always been a pleasure, and I will greatly miss her in my new job at the University of Central Florida.

In closing I thank the staff at Cambridge University Press, especially David Tranah, for their support and cooperation during the preparation of this volume and I look forward to working with them on future projects.

Orlando, FL　　Mourad E.H. Ismail
April 2005

1
Preliminaries

In this chapter we collect results from linear algebra and real and complex analysis which we shall use in this book. We will also introduce the definitions and terminology used. Some special functions are also introduced in the present chapter, but the q-series and related material are not defined until Chapter 11. See Chapters 11 and 12 for q-series.

1.1 Hermitian Matrices and Quadratic Forms

Recall that a matrix $A = (a_{j,k})$, $1 \leq j, k \leq n$ is called Hermitian if

$$\overline{a_{j,k}} = a_{k,j}, \quad 1 \leq j, k \leq n. \tag{1.1.1}$$

We shall use the following inner product on the n-dimensional complex space $\mathbb{C}^n$,

$$(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} x_j \, \overline{y_j}, \quad \mathbf{x} = (x_1, \dots, x_n)^T, \quad \mathbf{y} = (y_1, \dots, y_n)^T, \tag{1.1.2}$$

where A^T is the transpose of A. Clearly

$$(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}, \qquad (a\mathbf{x}, \mathbf{y}) = a(\mathbf{x}, \mathbf{y}), \quad a \in \mathbb{C}.$$

Two vectors $\mathbf{x}$ and $\mathbf{y}$ are called orthogonal if $(\mathbf{x}, \mathbf{y}) = 0$. The adjoint A^* of A is the matrix satisfying

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y}). \tag{1.1.3}$$

It is easy to see that if $A = (a_{j,k})$ then $A^* = (\overline{a_{k,j}})$. Thus, A is Hermitian if and only if $A^* = A$. The eigenvalues of Hermitian matrices are real. This is so since $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq 0$ then

$$\lambda(\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, \mathbf{x}) = (\mathbf{x}, A^*\mathbf{x}) = (\mathbf{x}, \lambda\mathbf{x}) = \overline{\lambda}(\mathbf{x}, \mathbf{x}).$$

Furthermore, the eigenvectors corresponding to distinct eigenvalues are orthogonal. This is the case because if $A\mathbf{x} = \lambda_1\mathbf{x}$ and $A\mathbf{y} = \lambda_2\mathbf{y}$ then

$$\lambda_1(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y}) = \lambda_2(\mathbf{x}, \mathbf{y}),$$

hence $(\mathbf{x}, \mathbf{y}) = 0$.

Any Hermitian matrix generates a quadratic form

$$\sum_{j,k=1}^{n} a_{j,k}\,\overline{x_j}\,x_k, \tag{1.1.4}$$

and conversely any quadratic form with $\overline{a_{j,k}} = a_{k,j}$ determines a Hermitian matrix A through

$$\sum_{j,k=1}^{n} a_{j,k}\,\overline{x_j}\,x_k = \mathbf{x}^* A\mathbf{x} = (A\mathbf{x},\mathbf{x}). \tag{1.1.5}$$

In an infinite dimensional Hilbert space $\mathcal{H}$, the adjoint is defined by (1.1.3) provided it holds for all $x, y \in \mathcal{H}$. A linear operator A defined in $\mathcal{H}$ is called self-adjoint if $A = A^*$. On the other hand, A is called symmetric if

$$(Ax, y) = (x, Ay)$$

whenever both sides are defined.

Theorem 1.1.1 *Assume that the entries of a matrix A satisfy $|a_{j,k}| \le M$ for all j, k and that each row of A has at most ℓ nonzero entries. Then all the eigenvalues of A satisfy*

$$|\lambda| \le \ell M.$$

Proof Take $\mathbf{x}$ to be an eigenvector of A with an eigenvalue λ, and assume that $\|\mathbf{x}\| = 1$. Observe that the Cauchy–Schwartz inequality implies

$$\begin{aligned} |\lambda|^2 = |(A\mathbf{x},\mathbf{x})|^2 &= \left|\sum_{j,k=1}^{n} a_{j,k}\,\overline{x_j}\,x_k\right|^2 \le \|x\|^2 \sum_{j=1}^{n}\left|\sum_{k=1}^{n} a_{j,k}x_k\right|^2 \\ &\le \ell^2 M^2. \end{aligned}$$

Hence the theorem is proved. □

A quadratic form (1.1.4) is *positive definite* if $(A\mathbf{x},\mathbf{x}) > 0$ for any nonzero $\mathbf{x}$. Recall that a matrix U is unitary if $U^* = U^{-1}$. The spectral theorem for Hermitian matrices is:

Theorem 1.1.2 *For every Hermitian matrix A there is a unitary matrix U whose columns are the eigenvectors of A such that*

$$A = U^*\Lambda U, \tag{1.1.6}$$

and Λ is the diagonal matrix formed by the corresponding eigenvalues of A.

For a proof see (Horn & Johnson, 1992). An immediate consequence of Theorem 1.1.2 is the following corollary.

Corollary 1.1.3 *The quadratic form* (1.1.4) *is reducible to a sum of squares,*

$$\sum_{j,k=1}^{n} a_{j,k}\,\overline{x_j}\,x_k = \sum_{k=1}^{n} \lambda_k\,|y_k|^2\,, \tag{1.1.7}$$

where $\mathbf{y} = U\mathbf{x}$*, and* $\lambda_1, \dots, \lambda_n$ *are the eigenvalues of* A*.*

The following important characterization of positive definite forms follows from Corollary 1.1.3.

Theorem 1.1.4 *The quadratic form* (1.1.4)–(1.1.5) *is positive definite if and only if the eigenvalues of* A *are positive.*

We next state the Sylvester criterion for positive definiteness (Shilov, 1977), (Horn & Johnson, 1992).

Theorem 1.1.5 *The quadratic form* (1.1.5) *is positive definite if and only if the principal minors of* A*, namely*

$$a_{1,1}, \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}, \dots, \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}, \tag{1.1.8}$$

are positive.

Recall that a matrix $A = (a_{j,k})$ is called **strictly diagonally dominant** if

$$2\,|a_{j,j}| > \sum_{k=1}^{n} |a_{j,k}|\,. \tag{1.1.9}$$

The following criterion for positive definiteness is in (Horn & Johnson, 1992, Theorem 6.1.10).

Theorem 1.1.6 *Let* A *be* $n \times n$ *matrix which is Hermitian, strictly diagonally dominant, and its diagonal entries are positive. Then* A *is positive definite.*

1.2 Some Real and Complex Analysis

We need some standard results from real and complex analysis which we shall state without proofs and provide references to where proofs can be found. We shall normalize functions of bounded variations to be continuous on the right.

Theorem 1.2.1 (Helly's selection principle) *Let* $\{\psi_n(x)\}$ *be a sequence of uniformly bounded nondecreasing functions. Then there is a subsequence* $\{\psi_{k_n}(x)\}$ *which converges to a nondecreasing bounded function,* ψ*. Moreover if for every* n *the moments* $\int_{\mathbb{R}} x^m\,d\psi_n(x)$ *exist for all* m*,* $m = 0, 1, \dots$*, then the moments of* ψ *exist and* $\int_{\mathbb{R}} x^m\,d\psi_{n_k}(x)$ *converges to* $\int_{\mathbb{R}} x^m\,d\psi(x)$*. Furthermore if* $\{\psi_n(x)\}$ *does not converge, then there are at least two such convergent subsequences.*

For a proof we refer the reader to Section 3 of the introduction to Shohat and Tamarkin (Shohat & Tamarkin, 1950).

Theorem 1.2.2 (Vitali) *Let $\{f_n(z)\}$ be a sequence of functions analytic in a domain $\mathcal{D}$ and assume that $f_n(z) \to f(z)$ pointwise in $\mathcal{D}$. Then $f_n(z) \to f(z)$ uniformly in any subdomain bounded by a contour C, provided that C is contained in $\mathcal{D}$.*

A proof is in Titchmarsh (Titchmarsh, 1964, page 168).

We now briefly discuss the Lagrange inversion and state two useful identities that will be used in later chapters.

Theorem 1.2.3 (Lagrange) *Let $f(z)$ and $\phi(z)$ be functions of z analytic on and inside a contour C containing the point a in its interior. Let t be such that $|t\phi(z)| < |z-a|$ on the contour C. Then the equation*

$$\zeta = a + t\phi(\zeta), \tag{1.2.1}$$

regarded as an equation in ζ, has one root interior to C; and further any function of ζ analytic on the closure of the interior of C can be expanded as a power series in t by the formula

$$f(\zeta) = f(a) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left[\frac{d^{n-1} f(x)[\phi(x)]^n}{dx^{n-1}}\right]_{x=a}. \tag{1.2.2}$$

See Whittaker and Watson (Whittaker & Watson, 1927, §7.32), or Polya and Szegő (Pólya & Szegő, 1972, p. 145). An equivalent form is

$$\frac{f(\zeta)}{1 - t\phi'(\zeta)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[\frac{d^n f(x)[\phi(x)]^n}{dx^n}\right]_{x=a} \tag{1.2.3}$$

Two important special cases are $\phi(z) = e^z$, or $\phi(z) = (1+z)^\beta$. These cases lead to:

$$e^{\alpha z} = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha+n)^{n-1}}{n!} w^n, \quad w = ze^{-z}, \tag{1.2.4}$$

$$(1+z)^\alpha = 1 + \alpha \sum_{n=1}^{\infty} \binom{\alpha + \beta n - 1}{n-1} \frac{w^n}{n}, \quad w = z(1+z)^{-\beta}. \tag{1.2.5}$$

We say that (Olver, 1974)

$$f(z) = O(g(z)), \quad \text{as } z \to a,$$

if $f(z)/g(z)$ is bounded in a neighborhood of $z = a$. On the other hand we write

$$f(z) = o(g(z)), \quad \text{as } z \to a$$

if $f(z)/g(z) \to 0$ as $z \to a$.

A very useful method to determine the large n behavior of orthogonal polynomials $\{p_n(x)\}$ is Darboux's asymptotic method.

Theorem 1.2.4 *Let $f(z)$ and $g(z)$ be analytic in $\{z : |z| < r\}$ and assume that*

$$f(z) = \sum_{n=0}^{\infty} f_n z^n, \quad g(z) = \sum_{n=0}^{\infty} g_n z^n, \quad |z| < r. \tag{1.2.6}$$

If $f - g$ is continuous on the closed disc $\{z : |z| \le r\}$ then

$$f_n = g_n + o\left(r^{-n}\right). \tag{1.2.7}$$

This form of Darboux's method is in (Olver, 1974, Ch. 8) and, in view of Cauchy's formulas, is just a restatement of the Riemann–Lebesgue lemma. For a given function f, g is called a comparison function. Another proof of Darboux's lemma is in (Knuth & Wilf, 1989).

In order to apply Darboux's method to a sequence $\{f_n\}$ we need first to find a generating function for the f_n's, that is, find a function whose Taylor series expansion around $z = 0$ has coefficients $c_n f_n$, for some simple sequence $\{c_n\}$. In this work we pay particular attention to generating functions of orthogonal polynomials and Darboux's method will be used to derive asymptotic expansions for some of the orthogonal polynomials treated in this work. The recent work (Wong & Zhao, 2005) shows how Darboux's method can be used to derive uniform asymptotic expansions. This is a major simplification of the version in (Fields, 1967). Wang and Wong developed a discrete version of the Liouville–Green approximation (WKB) in (Wang & Wong, 2005a). This gives uniform asymptotic expansions of a basis of solutions of three-term recurrence relations. This technique is relevant, because all orthogonal polynomials satisfy three-term recurrence relations.

The Perron–Stieltjes inversion formula, see (Stone, 1932, Lemma 5.2), is

$$F(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{z - t}, \quad z \notin \mathbb{R} \tag{1.2.8}$$

if and only if

$$\mu(t) - \mu(s) = \lim_{\epsilon \to 0^+} \int_s^t \frac{F(x - i\epsilon) - F(x + i\epsilon)}{2\pi i}\, dx. \tag{1.2.9}$$

The above inversion formula enables us to recover μ from knowing its Stieltjes transform $F(z)$.

Remark 1.2.1 *It is clear that if μ has an isolated atom u at $x = a$ then $z = a$ will be a pole of F with residue equal to u. Conversely, the poles of F determine the location of the isolated atoms of μ and the residues determine the corresponding masses. Formula* (1.2.9) *captures this behavior and reproduces the residue at an isolated singularity.*

Remark 1.2.2 *Formula* (1.2.9) *shows that the absolutely continuous component of μ is given by*

$$\mu'(x) = \left[F\left(x - i0^+\right) - F\left(x + i0^+\right)\right] / (2\pi i). \tag{1.2.10}$$

An analytic function defined on a closed disc is bounded and its absolute value attains its maximum on the boundary.

Definition 1.2.1 *Let f be an entire function. The maximum modulus is*

$$M(r;f) := \sup\{|f(z)| : |z| \le r\}, \quad r > 0. \tag{1.2.11}$$

The order of f, $\rho(f)$ is defined by

$$\rho(f) := \limsup_{r\to\infty} \frac{\ln\ln M(r,f)}{\ln r}. \tag{1.2.12}$$

Theorem 1.2.5 ((Boas, Jr., 1954)) *If $\rho(f)$ is finite and is not equal to a positive integer, then f has infinitely many zeros.*

If f has finite order, its type σ is

$$\sigma = \inf\{K : M(r) < \exp(Kr^\rho)\}. \tag{1.2.13}$$

For an entire function of finite order and type we define the Phragmén–Lindelöf indicator $h(\theta)$ as

$$h(\theta) = \overline{\lim_{r\to\infty}} \frac{\ln\left|f\left(re^{i\theta}\right)\right|}{r^\rho}. \tag{1.2.14}$$

Consider the infinite product

$$P = \prod_{n=1}^{\infty}(1 + a_n). \tag{1.2.15}$$

We say the P converges to ℓ, $\ell \neq 0$, if

$$\lim_{m\to\infty} \prod_{n=1}^{m}(1 + a_n) = \ell.$$

If $\ell = 0$ we say P diverges to zero. One can prove, see (Rainville, 1960, Chapter 1), that $a_n \to 0$ is necessary for P to converge. Similarly, one can define absolute convergence of infinite products. When $a_n = a_n(z)$ are functions of z, say, we say that P converges uniformly in a domain $\mathcal{D}$ if the partial products

$$\prod_{n=1}^{m}(1 + a_n(z))$$

converge uniformly in $\mathcal{D}$ to a function with no zeros in $\mathcal{D}$.

Definition 1.2.2 *Given a set of distinct points $\{x_j : 1 \le j \le n\}$, the Lagrange fundamental polynomial $\ell_k(x)$ is*

$$\ell_k(x) = \prod_{\substack{j=1\\ j\neq k}}^{n} \frac{(x - x_j)}{(x_k - x_j)} = \frac{S_n(x)}{S_n'(x_k)(x - x_k)}, \quad 1 \le k \le n, \tag{1.2.16}$$

where $S_n(x) = \prod_1^n (x - x_j)$. *The Lagrange interpolation polynomial of a function* $f(x)$ *at the nodes* $x_1, \ldots, x_n$ *is the unique polynomial* $L(x)$ *of degree* $n-1$ *such that* $f(x_j) = L(x_j)$.

It is easy to see that $L(x)$ in Definition 1.2.2 is

$$L(x) = \sum_{k=1}^{n} \ell_k(x) f(x_k) = \sum_{k=1}^{n} f(x_k) \frac{S_n(x)}{S_n'(x_k)(x - x_k)}. \tag{1.2.17}$$

Theorem 1.2.6 (Poisson Summation Formula) *Let* $f \in L_1(\mathbb{R})$ *and* F *be its Fourier transform,*

$$F(t) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-ixt}\, dx, \quad t \in \mathbb{R}.$$

Then

$$\sum_{k=-\infty}^{\infty} f(2k\pi) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-inx}\, dx.$$

For a proof, see (Zygmund, 1968, §II.13).

Theorem 1.2.7 *Given two differential equations in the form*

$$\frac{d^2u}{dz^2} + f(z)u(z) = 0, \quad \frac{d^2v}{dz^2} + g(z)v(z) = 0,$$

then $y = uv$ *satisfies*

$$\frac{d}{dz}\left\{\frac{y''' + 2(f+g)y' + (f'+g')\,y}{f-g}\right\} + (f-g)y = 0, \text{ if } f \neq g \tag{1.2.18}$$

$$y''' + 4fy' + 2f'y = 0, \quad \text{if } f = g. \tag{1.2.19}$$

A proof of Theorem 1.2.6 is in Watson (Watson, 1944, §5.4), where he attributes the theorem to P. Appell.

Lemma 1.2.8 *Let* $y = y(x)$ *satisfy the differential equation*

$$\phi(x)y''(x) + y(x) = 0, \quad a < x < b \tag{1.2.20}$$

where $\phi(x) > 0$, *and* $\phi'(x)$ *is positive* (*negative*) *and continuous on* (a, b). *Then the successive relative maxima of* $|y|$, *increase* (*decrease*) *with* x *in* (a, b) *if* ϕ *increases* (*decreases*) *on* (a, b).

Proof Let

$$f(x) := \{y(x)\}^2 + \phi(x)\{y'(x)\}^2, \tag{1.2.21}$$

so that $f(x) = \{y(x)\}^2$ if $y'(x) = 0$. Clearly

$$\begin{aligned} f'(x) &= y'(x)\{2y(x) + \phi'(x)y'(x) + 2\phi(x)y''(x)\} \\ &= \phi'(x)\{y'(x)\}^2. \end{aligned}$$

Thus sign $f'(x) = \operatorname{sign} \phi'$ in between the consecutive successive maxima of $|y|$ and the result follows. $\square$

1.3 Some Special Functions

Standard references in this area are (Andrews et al., 1999), (Bailey, 1935), (Rainville, 1960), (Erdélyi et al., 1953b), (Slater, 1964), (Whittaker & Watson, 1927).

The gamma and beta functions are probably the most important functions in mathematics beyond the exponential and logarithmic functions. Recall that

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt, \quad \text{Re } z > 0, \tag{1.3.1}$$

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt, \quad \text{Re } x > 0, \quad \text{Re } y > 0. \tag{1.3.2}$$

They are related through

$$B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y). \tag{1.3.3}$$

The functional equation

$$\Gamma(z+1) = z\Gamma(z) \tag{1.3.4}$$

extends the gamma function to a meromorphic function with poles at $z = 0, -1, \dots,$ and also extends $B(x,y)$ to a meromorphic function of x and y. The Mittag–Leffler expansion for Γ'/Γ is (Whittaker & Watson, 1927, §12.3)

$$\frac{\Gamma'(z)}{\Gamma(z)} = -\gamma - \frac{1}{z} - \sum_{n=1}^{\infty} \left[\frac{1}{z+n} - \frac{1}{n}\right], \tag{1.3.5}$$

where γ is the Euler constant, (Rainville, 1960, §7).

The shifted factorial is

$$(a)_0 := 1, \quad (a)_n = a(a+1)\cdots(a+n-1), \quad n > 0, \tag{1.3.6}$$

hence (1.3.4) gives

$$(a)_n = \Gamma(a+n)/\Gamma(a). \tag{1.3.7}$$

The shifted factorial is also called Pochhammer symbol. Note that (1.3.7) is meaningful for any complex n, when $a+n$ is not a pole of the gamma function. The gamma function and the shifted factorial satisfy the duplication formulas

$$\Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z+1/2)/\sqrt{\pi}, \quad (2a)_{2n} = 2^{2n}(a)_n(a+1/2)_n. \tag{1.3.8}$$

We also have the reflection formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}. \tag{1.3.9}$$

We define the multishifted factorial as

$$(a_1, \cdots, a_m)_n = \prod_{j=1}^{m} (a_j)_n\, .$$

Some useful identities are

$$(a)_m(a+m)_n = (a)_{m+n}, \quad (a)_{N-k} = \frac{(a)_N(-1)^k}{(-a-N+1)_k}. \tag{1.3.10}$$

A hypergeometric series is

$$\begin{aligned} {}_rF_s\left(\begin{matrix} a_1,\dots,a_r \\ b_1,\dots,b_s \end{matrix}\,\middle|\, z\right) &= {}_rF_s\left(a_1,\dots,a_r;b_1,\dots,b_s;z\right) \\ &= \sum_{n=0}^{\infty} \frac{(a_1,\dots,a_r)_n}{(b_1,\dots,b_s)_n}\frac{z^n}{n!}. \end{aligned} \tag{1.3.11}$$

If one of the numerator parameters is a negative integer, say $-k$, then the series (1.3.11) becomes a finite sum, $0 \le n \le k$ and the ${}_rF_s$ series is called *terminating*. As a function of z nonterminating series is entire if $r \le s$, is analytic in the unit disc if $r = s+1$. The hypergeometric function ${}_2F_1(a,b;c;z)$ satisfies the hypergeometric differential equation

$$z(1-z)\frac{d^2y}{dz^2} + [c-(a+b+1)z]\frac{dy}{dz} - aby = 0. \tag{1.3.12}$$

The confluent hypergeometric function (Erdélyi et al., 1953b, §6.1)

$$\Phi(a,c;z) := {}_1F_1(a;c;z) \tag{1.3.13}$$

satisfies the differential equation

$$z\frac{d^2y}{dz^2} + (c-z)\frac{dy}{dz} - ay = 0, \tag{1.3.14}$$

and $\lim\limits_{b\to\infty} {}_2F_1(a,b;cz/b) = {}_1F_1(a;c;z)$. The Tricomi Ψ function is a second linear independent solution of (1.3.14) and is defined by (Erdélyi et al., 1953b, §6.5)

$$\Psi(a,c;x) := \frac{\Gamma(1-c)}{\Gamma(a-c+1)}\Phi(a,c;x) + \frac{\Gamma(c-1)}{\Gamma(a)}x^{1-c}\Phi(a-c+1,2-c;x). \tag{1.3.15}$$

The function of Ψ has the integral presentation (Erdélyi et al., 1953a, §6.5)

$$\Psi(a,c;x) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-xt}t^{a-1}(1+t)^{c-a-1}\,dt, \tag{1.3.16}$$

for $\operatorname{Re} a > 0$, $\operatorname{Re} x > 0$.

The Bessel function J_ν and the modified Bessel function I_ν, (Watson, 1944) are

$$\begin{aligned} J_\nu(z) &= \sum_{n=0}^{\infty}\frac{(-1)^n\,(z/2)^{\nu+2n}}{\Gamma(n+\nu+1)\,n!}, \\ I_\nu(z) &= \sum_{n=0}^{\infty}\frac{(z/2)^{\nu+2n}}{\Gamma(n+\nu+1)\,n!}. \end{aligned} \tag{1.3.17}$$

Clearly $I_\nu(z) = e^{-i\pi\nu/2}J_\nu\left(ze^{i\pi/2}\right)$. Furthermore

$$J_{1/2}(z) = \sqrt{\frac{2}{\pi z}}\sin z, \quad J_{-1/2}(z) = \sqrt{\frac{2}{\pi z}}\cos z. \tag{1.3.18}$$

The Bessel functions satisfy the recurrence relation

$$\frac{2\nu}{z} J_\nu(z) = J_{\nu+1}(z) + J_{\nu-1}(z). \tag{1.3.19}$$

The Bessel functions J_ν and $J_{-\nu}$ satisfy

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + \left(x^2 - \nu^2\right) y = 0. \tag{1.3.20}$$

When ν is not an integer J_ν and $J_{-\nu}$ are linear independent solutions of (1.3.20) whose Wronskian is (Watson, 1944, §3.12)

$$W\{J_\nu(x), J_{-\nu}(x)\} = -\frac{2\,\sin(\nu\pi)}{\pi\,x}, \quad W\{f, g\} := f g' - g f'. \tag{1.3.21}$$

The function I_ν satisfies the differential equation

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - \left(x^2 + \nu^2\right) y = 0, \tag{1.3.22}$$

whose second solution is

$$\begin{aligned} K_\nu(x) &= \frac{\pi}{2}\,\frac{I_{-\nu}(x) - I_\nu(x)}{\sin(\pi\nu)}, \\ K_n(x) &= \lim_{\nu\to n} K_\nu(x), \quad n = 0, \pm 1, \dots . \end{aligned} \tag{1.3.23}$$

We also have

$$\begin{aligned} I_{\nu-1}(x) - I_{\nu+1}(x) &= \frac{2\nu}{x}\, I_\nu(x), \\ K_{\nu+1}(x) - K_{\nu-1}(x) &= \frac{2\nu}{x}\, K_\nu(x). \end{aligned} \tag{1.3.24}$$

Theorem 1.3.1 *When $\nu > -1$, the function $z^{-\nu} J_\nu(z)$ has only real and simple zeros. Furthermore, the positive (negative) zeros of $J_\nu(z)$ and $J_{\nu+1}(z)$ interlace for $\nu > -1$.*

We shall denote the positive zeros of $J_\nu(z)$ by $\{j_{\nu,k}\}$, that is

$$0 < j_{\nu,1} < j_{\nu,2} < \cdots < j_{\nu,n} < \cdots . \tag{1.3.25}$$

The Bessel functions satisfy the differential recurrence relations, (Watson, 1944)

$$z J_\nu'(z) = \nu J_\nu(z) - z J_{\nu+1}(z), \tag{1.3.26}$$

$$z Y_\nu'(z) = \nu Y_\nu(z) - z Y_{\nu+1}(z), \tag{1.3.27}$$

$$z I_\nu(z) = z I_{\nu+1}(z) + \nu I_\nu(z), \tag{1.3.28}$$

$$z K_\nu'(z) = \nu K_\nu(z) - z K_{\nu+1}(z), \tag{1.3.29}$$

where $Y_\nu(z)$ is

$$\begin{aligned} Y_\nu(z) &= \frac{J_\nu(z)\cos\nu\pi - J_{-\nu}(z)}{\sin\nu\pi}, \quad \nu \neq 0, \pm 1, \dots, \\ Y_n(z) &= \lim_{\nu\to n} Y_\nu(z), \quad n = 0, \pm 1, \dots . \end{aligned} \tag{1.3.30}$$

The functions $J_\nu(z)$ and $Y_\nu(z)$ are linearly independent solutions of (1.3.20).

The Bessel functions are special cases of ${}_1F_1$ in the sense

$$e^{-iz}\,{}_1F_1(\nu+1/2;2\nu+1;2iz)=\Gamma(\nu+1)(z/2)^{-\nu}J_\nu(z), \tag{1.3.31}$$

(Erdélyi et al., 1953b, §6.9.1).

Two interesting functions related to special cases of Bessel functions are the functions

$$\begin{aligned} k(x)&=\frac{\pi}{3}\left(\frac{x}{3}\right)^{\frac{1}{2}}J_{-1/3}\left(2(x/3)^{3/2}\right)=\frac{\pi}{3}\sum_{n=0}^{\infty}\frac{(-x/3)^{3n}}{n!\,\Gamma(n+2/3)},\\ \ell(x)&=\frac{\pi}{3}\left(\frac{x}{3}\right)^{\frac{1}{2}}J_{1/3}\left(2(x/3)^{3/2}\right)=\frac{\pi}{9}\,x\sum_{n=0}^{\infty}\frac{(-x/3)^{3n}}{n!\,\Gamma(n+4/3)}. \end{aligned} \tag{1.3.32}$$

Indeed $\{k(x),\ell(x)\}$ is a basis of solutions of the Airy equation

$$\frac{d^2y}{dx^2}+\frac{1}{3}\,xy=0. \tag{1.3.33}$$

Moreover

$$k(x)=-\ell(x)(1+o(1))=\frac{3^{-\frac{3}{4}}}{2}\,\sqrt{\pi}\,|x|^{-\frac{1}{4}}\exp\left\{2(|x|/3)^{3/2}\right\}(1+o(1)),$$

as $x\to-\infty$. Thus the only solution of (1.3.33) which is bounded as $x\to-\infty$ is $k(x)+\ell(x)$. Set

$$A(x):=k(x)+\ell(x), \tag{1.3.34}$$

which has the asymptotic behavior

$$A(x)=\frac{\sqrt{\pi}}{2\,3^{1/4}}\,|x|^{-\frac{1}{4}}\exp\left\{-2(|x|/3)^{3/2}\right\}(1+o(1))$$

as $x\to-\infty$. The function $A(x)$ is called the Airy function and plays an important role in the theory of orthogonal polynomials with exponential weights, random matrix theory, as well as other parts of mathematical physics. The function $A(x)$ is positive on $(-\infty,0)$ and has only positive simple zeros.

We shall use the notation

$$0<i_1<i_2<\cdots, \tag{1.3.35}$$

for the positive zeros of the Airy function.

The Appell functions generalize the hypergeometric function to two variables.

They are defined by (Appell & Kampé de Fériet, 1926), (Erdélyi et al., 1953b)

$$F_1\left(a;b,b';c;x,y\right)=\sum_{m,n=0}^{\infty}\frac{(a)_{n+m}(b)_m\left(b'\right)_n}{(c)_{m+n}\,m!\,n!}\,x^m y^n, \tag{1.3.36}$$

$$F_2\left(a;b,b';c,c';x,y\right)=\sum_{m,n=0}^{\infty}\frac{(a)_{n+m}(b)_m\left(b'\right)_n}{(c)_m(c')_n\,m!\,n!}\,x^m y^n, \tag{1.3.37}$$

$$F_3\left(a,a';b,b';c;x,y\right)=\sum_{m,n=0}^{\infty}\frac{(a)_m\left(a'\right)_n(b)_m\left(b'\right)_n}{(c)_{m+n}\,m!\,n!}\,x^m y^n, \tag{1.3.38}$$

$$F_4\left(a,b;c,c';x,y\right)=\sum_{m,n=0}^{\infty}\frac{(a)_{n+m}(b)_{m+n}}{(c)_m(c')_n\,m!\,n!}\,x^m y^n. \tag{1.3.39}$$

The complete elliptic integrals of the first and second kinds are (Erdélyi et al., 1953a)

$$\mathbf{K}=\mathbf{K}(k)=\int_0^1\frac{du}{\sqrt{(1-u^2)\,(1-k^2u^2)}}, \tag{1.3.40}$$

$$\mathbf{E}=\mathbf{E}(k)=\int_0^1\sqrt{\frac{1-k^2u^2}{1-u^2}}\,du \tag{1.3.41}$$

respectively. Indeed

$$\mathbf{K}(k)=\frac{\pi}{2}\,{}_2F_1\left(1/2,1/2;1;k^2\right), \tag{1.3.42}$$

$$\mathbf{E}(k)=\frac{\pi}{2}\,{}_2F_1\left(-1/2,1/2;1;k^2\right). \tag{1.3.43}$$

We refer to k as the modulus, while the complementary modulus k' is

$$k'=\left(1-k^2\right)^{1/2}. \tag{1.3.44}$$

1.4 Summation Theorems and Transformations

In the shifted factorial notation the binomial theorem is

$$\sum_{n=0}^{\infty}\frac{(a)_n}{n!}\,z^n=(1-z)^{-a}, \tag{1.4.1}$$

when $|z|<1$ if a is not a negative integer.

The Gauss sum is

$${}_2F_1\left(\left.\begin{matrix}a,b\\c\end{matrix}\,\right|1\right)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)},\quad \mathrm{Re}\{c-a-b\}>0. \tag{1.4.2}$$

The terminating version of (1.4.2) is the Chu–Vandermonde sum

$${}_2F_1\left(\left.\begin{matrix}-n,b\\c\end{matrix}\,\right|1\right)=\frac{(c-b)_n}{(c)_n}. \tag{1.4.3}$$

A hypergeometric series (1.3.11) is called *balanced* if $r = s + 1$ and

$$1 + \sum_{k=1}^{s+1} a_k = \sum_{k=1}^{s} b_k \tag{1.4.4}$$

The Pfaff–Saalschütz theorem is

$${}_3F_2\left(\left.\begin{matrix}-n, a, b\\ c, d\end{matrix}\,\right|\, 1\right) = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}, \tag{1.4.5}$$

if the balanced condition, $c + d = 1 - n + a + b$, is satisfied. In Chapter 13 we will give proofs of generalizations of (1.4.2) and (1.4.5) to q-series.

The Stirling formula for the gamma function is

$$\operatorname{Log}\Gamma(z) = \left(z - \frac{1}{2}\right)\operatorname{Log} z - z + \frac{1}{2}\ln(2\pi) + O\left(z^{-1}\right), \tag{1.4.6}$$

$|\arg z| \le \pi - \epsilon$, $\epsilon > 0$. An important consequence of Stirling's formula is

$$\lim_{z\to\infty} z^{b-a}\,\frac{\Gamma(z+a)}{\Gamma(z+b)} = 1. \tag{1.4.7}$$

The hypergeometric function has the Euler integral representation

$$\begin{aligned}&{}_2F_1\left(\left.\begin{matrix}a, b\\ c\end{matrix}\,\right|\, z\right)\\ &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}\,dt,\end{aligned} \tag{1.4.8}$$

for $\operatorname{Re} b > 0$, $\operatorname{Re}(c-b) > 0$. The Pfaff–Kummer transformation is

$${}_2F_1\left(\left.\begin{matrix}a, b\\ c\end{matrix}\,\right|\, z\right) = (1-z)^{-a}{}_2F_1\left(\left.\begin{matrix}a, c-b\\ c\end{matrix}\,\right|\, \frac{z}{z-1}\right), \tag{1.4.9}$$

and is valid for $|z| < 1$, $|z| < |z-1|$. An iterate of (1.4.9) is

$${}_2F_1\left(\left.\begin{matrix}a, b\\ c\end{matrix}\,\right|\, z\right) = (1-z)^{c-a-b}{}_2F_1\left(\left.\begin{matrix}c-a, c-b\\ c\end{matrix}\,\right|\, z\right), \tag{1.4.10}$$

for $|z| < 1$. Since ${}_1F_1(a;\zeta;z) = \lim\limits_{b\to\infty} {}_2F_1(a, b; c; z/b)$, (1.4.10) yields

$${}_1F_1(a; c; z) = e^z{}_1F_1(c-a; c; -z). \tag{1.4.11}$$

Many of the summation theorems and transformation formulas stated in this section have q-analogues which will be stated and proved in §12.2.

In particular we shall give a new proof of the terminating case of Watson's theorem (Slater, 1964, (III.23))

$$\begin{aligned}&{}_3F_2\left(\left.\begin{matrix}a, b, c\\ (a+b+1)/2, 2c\end{matrix}\,\right|\, 1\right)\\ &= \frac{\Gamma(1/2)\Gamma(c+1/2)\Gamma((a+b+1)/2)\Gamma(c+(1-a-b)/2)}{\Gamma((a+1)/2)\Gamma((b+1)/2)\Gamma(c+(1-a)/2)\Gamma(c+(1-b)/2)}\end{aligned} \tag{1.4.12}$$

in §9.2. We also give a new proof of the q-analogue of (1.4.12) in the terminating and nonterminating cases. As $q \to 1$ we get (1.4.12) in its full generality. This is just one of many instances where orthogonal polynomials shed new light on the theory of evaluation of sums and integrals.

We shall make use of the quadratic transformation (Erdélyi et al., 1953a, (2.11.37))

$$\begin{aligned} {}_2F_1\left(\left.\begin{matrix} a, b \\ 2b \end{matrix}\right| z\right) &= \left(\frac{1+(1-z)^{1/2}}{2}\right)^{-2a} \\ &\times {}_2F_1\left(\left.\begin{matrix} a, a-b+\frac{1}{2} \\ b+\frac{1}{2} \end{matrix}\right| \left[\frac{1-(1-z)^{1/2}}{1+(1-z)^{1/2}}\right]^2\right). \end{aligned} \tag{1.4.13}$$

In particular,

$${}_2F_1\left(\left.\begin{matrix} a, a+1/2 \\ 2a+1 \end{matrix}\right| z\right) = \left(\frac{1+(1-z)^{1/2}}{2}\right)^{-2a}. \tag{1.4.14}$$

Exercises

1.1 Prove that if $c, b_1, b_2, \dots, b_n$ are distinct complex numbers, then (Gosper et al., 1993)

$$\begin{aligned} \prod_{k=1}^{n} \frac{x+a_k-b_k}{x-b_k} &= \prod_{k=1}^{n} \frac{c+a_k-b_k}{c-b_k} \\ &+ \prod_{k=1}^{n} \frac{a_k(x-c)}{(b_k-c)\,(x-b_k)} \prod_{\substack{j=1 \\ j\neq k}} \frac{b_k+a_j-b_j}{b_k-b_j}. \end{aligned}$$

This formula is called the nonlocal derangement identity.

1.2 Use Exercise 1.1 to prove that (Gosper et al., 1993)

$$\begin{aligned} &\sum_{j=1}^{\infty} \left(1-(-1)^j \cos\left(\sqrt{j(j+a)}\,\pi\right)\right) \\ &\times \frac{\Gamma\left(\left(j+\sqrt{j(j+a)}\right)/2\right)\Gamma\left(\left(j-\sqrt{j(j+a)}\right)/2\right)}{j!\,j} = -\frac{\pi^4 a}{12}, \end{aligned}$$

where $\operatorname{Re} a < 4$.

1.3 Derive Sonine's first integral

$$J_{\alpha+\beta+1}(z) = \frac{2^{-\beta} z^{\beta+1}}{\Gamma(\beta+1)} \int_0^1 x^{2\beta+1}\left(1-x^2\right)^{\alpha/2} J_\alpha\left(z\sqrt{1-x^2}\right) dx,$$

where $\operatorname{Re}\alpha > -1$ and $\operatorname{Re}\beta > -1$.

1.4 Prove the identities (1.2.4) and (1.2.5).

1.5 Prove that (Gould, 1962)

$$f(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{a+bk}{n} g(n)$$

if and only if

$$\binom{a+bn}{n} g(n) = \sum_{k=0}^{n} \frac{a+bk-k}{a+bn-k} \binom{a+bn-k}{n-k} f(k).$$

Hint: Express the exponential generating function $\sum_{n=0}^{\infty} f(n)(-t)^n/n!$ in terms of $\sum_{n=0}^{\infty} g(n)(-t)^n/n!$ by using (1.2.5).

1.6 Prove the generating function

$$\sum_{n=0}^{\infty} \frac{(\lambda)_n}{n!} \phi_n(x) t^n$$
$$= (1-t)^{-\lambda} {}_{p+2}F_s \left(\left. \begin{matrix} \lambda/2, (\lambda+1)/2, a_1, \ldots, a_p \\ b_1, \ldots, b_s \end{matrix} \right| \frac{-4tx}{(1-t)^2} \right),$$

when $s \ge p+1$, $\left| tx/(1-t)^2 \right| < 1/4$, $|t| < 1$, where

$$\phi_n(x) = {}_{p+2}F_s \left(\left. \begin{matrix} -n, n+\lambda, a_1, \ldots, a_p \\ b_1, \ldots, b_s \end{matrix} \right| x \right).$$

Note: This formula and Darboux's method can be used to determine the large n asymptotics of $\phi_n(x)$, see §7.4 in (Luke, 1969a), where $\{\phi_n(x)\}$ are called extended Jacobi polynomials.

2

Orthogonal Polynomials

This chapter develops properties of general orthogonal polynomials. These are polynomials orthogonal with respect to positive measures. An application to solving the Toda lattice equations is given in §2.8.

We start with a given positive Borel measure μ on $\mathbb{R}$ with infinite support and whose moments $\int_{\mathbb{R}} x^n \, d\mu(x)$ exist for $n = 0, 1, \ldots$. Recall that the distribution function $F_\mu(x)$ of a finite Borel measure μ is $F_\mu(x) = \mu((-\infty, x])$. A distribution function is nondecreasing, right continuous, nonnegative, bounded, and $\lim_{x \to -\infty} F_\mu(x) = 0$. Conversely any function satisfying these properties is a distribution function for a measure μ and $\int_{\mathbb{R}} f(x) \, dF_\mu(x) = \int_{\mathbb{R}} f(x) \, d\mu(x)$, see (McDonald & Weiss, 1999, §4.7). Because of this fact we will use μ to denote measures or distribution functions and we hope this will not cause any confusion to our readers.

By a **polynomial sequence** we mean a sequence of polynomials, say $\{\varphi_n(x)\}$, such that $\varphi_n(x)$ has precise degree n. A polynomial sequence $\{\varphi_n(x)\}$ is called **monic** if $\varphi_n(x) - x^n$ has degree at most $n - 1$.

2.1 Construction of Orthogonal Polynomials

Given μ as above we observe that the moments

$$\mu_j := \int_{\mathbb{R}} x^j \, d\mu(x), \quad j = 0, 1, \ldots, \tag{2.1.1}$$

generate a quadratic form

$$\sum_{j,k=0}^{n} \mu_{k+j} \, \overline{x_j} \, x_k. \tag{2.1.2}$$

We shall always normalize the measures to have $\mu_0 = 1$, that is

$$\int_{\mathbb{R}} d\mu = 1. \tag{2.1.3}$$

The form (2.1.2) is positive definite since μ has infinite support and the expression in (2.1.2) is

$$\int_{\mathbb{R}} \left| \sum_{j=0}^{n} x_j t^j \right|^2 d\mu(t).$$

Let D_n denote the determinant

$$D_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}. \tag{2.1.4}$$

The Sylvester criterion, Theorem 1.1.5, implies the positivity of the determinants $D_0, D_1, \ldots$. The determinant D_n is a Hankel determinant.

Theorem 2.1.1 *Given a positive Borel measure μ on $\mathbb{R}$ with infinite support and finite moments, there exists a unique sequence of monic polynomials $\{P_n(x)\}_0^\infty$,*

$$P_n(x) = x^n + \textit{lower order terms}, \quad n = 0, 1, \ldots,$$

and a sequence of positive numbers $\{\zeta_n\}_0^\infty$, with $\zeta_0 = 1$ such that

$$\int_{\mathbb{R}} P_m(x) P_n(x)\, d\mu(x) = \zeta_n\, \delta_{m,n}. \tag{2.1.5}$$

Proof We prove (2.1.5) for $m, n = 0, \ldots, N$, and $N = 0, \ldots,$ by induction on N. Define $P_0(x)$ to be 1. Assume $P_0(x), \ldots, P_N(x)$ have been defined and (2.1.5) holds. Set $P_{N+1}(x) = x^{N+1} + \sum_{j=0}^{N} c_j x^j$. For $m < N+1$, we have

$$\int_{\mathbb{R}} x^m P_{N+1}(x)\, d\mu(x) = \mu_{N+m+1} + \sum_{j=0}^{N} c_j \mu_{j+m}.$$

We construct the coefficients c_j by solving the system of equations

$$\sum_{j=0}^{N} c_j \mu_{j+m} = -\mu_{N+m+1}, \quad m = 0, 1, \ldots, N,$$

whose determinant is D_N and $D_N > 0$. Thus the polynomial $P_{N+1}(x)$ has been found and ζ_{N+1} is $\int_{\mathbb{R}} P_{N+1}^2(x)\, d\mu(x)$. $\square$

As a consequence of Theorem 2.1.1 we see that

$$P_n(x) = \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix}, \; \zeta_n = \frac{D_n}{D_{n-1}}, \tag{2.1.6}$$

since the right-hand side of (2.1.6) satisfies the requirements of Theorem 2.1.1. The reason is that for $m < n$, $\int_{\mathbb{R}} x^m P_n(x)\,d\mu(x)$ is a determinant whose $n+1$ and $m+1$ rows are equal. To evaluate ζ_n, note that

$$\zeta_n = \int_{\mathbb{R}} P_n^2(x)\,d\mu(x) = \int_{\mathbb{R}} x^n P_n(x)\,dx$$
$$= \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & & \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} = \frac{D_n}{D_{n-1}}.$$

Remark 2.1.1 *The odd moments will vanish if μ is symmetric about the origin, hence* (2.1.6) *shows that $P_n(x)$ contains only the terms x^{n-2k}, $0 \le k \le n/2$.*

We now prove the following important result of Heine.

Theorem 2.1.2 *The monic orthogonal polynomials $\{P_n(x)\}$ have the Heine integral representation*

$$P_n(x) = \frac{1}{n!\,D_{n-1}} \int_{\mathbb{R}^n} \prod_{i=1}^{n} (x - x_i) \prod_{1 \le j < k \le n} (x_k - x_j)^2\,d\mu(x_1)\cdots d\mu(x_n). \tag{2.1.7}$$

Proof In the determinant in (2.1.6) write μ_j as $\int_{\mathbb{R}} x_k^j\,d\mu(x_k)$ in row k, $1 \le k \le n$. Thus

$$P_n(x) = \frac{1}{D_{n-1}} \int_{\mathbb{R}^n} \begin{vmatrix} 1 & x_1 & \cdots & x_1^n \\ x_2 & x_2^2 & \cdots & x_2^{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ x_n^{n-1} & x_n^n & \cdots & x_n^{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} d\mu(x_1)\cdots d\mu(x_n). \tag{2.1.8}$$

If we choose the integration variables to be $x_{k_1}, x_{k_2}, \ldots, x_{k_n}$ where $(k_1, k_2, \ldots, k_n) = \sigma(1, 2, \ldots, n)$, and σ is a permutation of $(1, 2, \ldots, n)$, then (2.1.8) becomes

$$P_n(x) = \frac{\operatorname{sign}\sigma}{D_{n-1}}$$
$$\times \int_{\mathbb{R}^n} \begin{vmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^n \\ 1 & x & \cdots & x^n \end{vmatrix} x_1^{k_1}\cdots x_n^{k_n}\,d\mu(x_1)\cdots d\mu(x_n),$$

where $\operatorname{sign}\sigma$ is (-1) raised to the number of inversions in $\sigma(1, \ldots, n) = (k_1, \ldots, k_n)$, see (Shilov, 1977, Chapter 1). Now sum over all permutations σ of $(1, 2, \ldots, n)$, which are $n!$ in number, then divide by $n!$ we get the factor

$$\sum_{\sigma}(\operatorname{sign}\sigma)x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

inside the integrand. The above sum is a Vandermonde determinant hence is equal to $\prod_{1\le j<k\le n}(x_k - x_j)$. Finally the determinant in the integrand is a Vandermonde determinant with variables $x_1, x_2, \ldots, x_n, x$, and it evaluates to

$$\prod_{i=1}^{n}(x - x_i)\prod_{1\le j<k\le n}(x_k - x_j),$$

hence the result. □

By equating coefficients of x^n on both sides of (2.1.7) we prove the following corollary to Heine's formula.

Corollary 2.1.3 *We have*

$$\int_{\mathbb{R}^{n+1}}\prod_{1\le j<k\le n+1}(x_k - x_j)^2\,d\mu(x_1)\cdots d\mu(x_{n+1}) = (n+1)!\,D_n. \tag{2.1.9}$$

Both Heine's formula and its corollary have been used extensively in the theory of random matrices, (Mehta, 1991), (Deift, 1999). Heine's formula when $\mu'(x) = x^\alpha(1-x)^\beta$, $x\in[0,1]$ has been generalized by Selberg to what is known as the Selberg integral

$$\begin{aligned}\int_{[0,1]^n}\prod_{k=1}^{n}\left\{x_k^{\alpha-1}(1-x_k)^{\beta-1}\right\}\prod_{1\le i<j\le n}|x_i - x_j|^{2\gamma}\,dx_1\cdots dx_n\\ = \prod_{j=1}^{n}\frac{\Gamma(\alpha+(j-1)\gamma)\,\Gamma(\beta+(j-1)\gamma)\,\Gamma(1+j\gamma)}{\Gamma(\alpha+\beta+(n+j-2)\gamma)\,\Gamma(1+\gamma)},\end{aligned}$$

(Andrews et al., 1999, §8.1).

Another determinant representation for $\{P_n(x)\}$ is

$$P_n(x) = \frac{1}{D_{n-1}}\begin{vmatrix}\mu_0x-\mu_1 & \mu_1x-\mu_2 & \cdots & \mu_{n-1}x-\mu_n\\ \mu_1x-\mu_2 & \mu_2x-\mu_3 & \cdots & \mu_nx-\mu_{n+1}\\ \vdots & \vdots & \vdots & \vdots\\ \mu_{n-1}x-\mu_n & \mu_nx-\mu_{n+1} & \cdots & \mu_{2n-2}x-\mu_{2n-1}\end{vmatrix}. \tag{2.1.10}$$

To prove (2.1.10), multiply column j in (2.1.6) by $-x$ and add it to the $j+1$ column for $j = n, n-1, \ldots, 1$.

It is important to note that the construction of the monic orthogonal polynomials in (2.1.6) and Theorem 2.1.1 depends only on the moments and not on the original measure.

The orthonormal polynomials will be denoted by $\{p_n(x)\}$

$$p_n(x) = P_n(x)/\sqrt{\zeta_n},$$

so that

$$p_n(x) = \frac{1}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix}. \tag{2.1.11}$$

We shall use the notation

$$p_n(x) = \gamma_n x^n + \text{lower order terms} \tag{2.1.12}$$

with

$$\gamma_n = \sqrt{D_{n-1}/D_n}. \tag{2.1.13}$$

Consequently,

$$D_n = 1/\left[\gamma_1^2 \cdots \gamma_n^2\right]. \tag{2.1.14}$$

Sometimes it is more convenient to use polynomial bases other than the monomials. Let $\{\phi_n(x)\}$ be a polynomial sequence with real coefficients and with $\phi_0(x) = 1$. For a given probability measure with finite moments set

$$\phi_{jk} = \int_{\mathbb{R}} \phi_j(x)\phi_k(x)\, d\mu(x). \tag{2.1.15}$$

Theorem 2.1.4 *The matrices $\{\phi_{jk} : 0 \le j, k \le n\}$, $n = 0, 1, \ldots,$ are positive definite. With*

$$\tilde{D}_n = \begin{vmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n} \\ \vdots & & & \\ \phi_{n,0} & \phi_{n,1} & \cdots & \phi_{n,n} \end{vmatrix}, \quad n \ge 0, \tag{2.1.16}$$

the polynomials orthonormal with respect to μ are

$$p_n(x) = \frac{1}{\sqrt{\tilde{D}_n \tilde{D}_{n-1}}} \begin{vmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n} \\ \ddots & & & \\ \phi_{n-1,0} & \phi_{n-1,1} & \cdots & \phi_{n-1,n} \\ \phi_0(x) & \phi_1(x) & \cdots & \phi_n(x) \end{vmatrix}. \tag{2.1.17}$$

Proof Clearly $\phi_{jk} = \phi_{kj}$ and the quadratic form $\sum \phi_{jk} x_j x_k$ is $\int_{\mathbb{R}} \left|\sum x_j \phi_j(x)\right|^2 d\mu(x)$ which implies the positive definiteness of the matrices $\{\phi_{jk} : 0 \le j, k \le n\}$, $n = 0, 1, \ldots$. Hence $\tilde{D}_n > 0$, $n \ge 0$. The proof of the orthonormality of $\{P_n\}$ is similar to the proof for $\phi_n(x) = x^n$. $\square$

It is our opinion that the determinant representations of orthogonal polynomials are underused. For a clever use of these representations see (Wilson, 1991).

A matrix of the form

$$\begin{pmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n} \\ \vdots & \vdots & \cdots & \vdots \\ \phi_{n,0} & \phi_{n,1} & \cdots & \phi_{n,n} \end{pmatrix}$$

is called a **Gram matrix**. A Gram matrix is always positive definite.

Remark 2.1.2 *Observe that the construction of the monic orthogonal polynomials only used the fact that $D_n \neq 0$, so the positivity of D_n was used to construct the orthonormal polynomials. Therefore, one can allow the measure μ to be a signed measure but assume that $D_n \neq 0$. When $\{P_n(x)\}$ are monic polynomials orthogonal with respect to a signed measure μ with $D_n \neq 0$, we shall call $\{P_n(x)\}$ signed orthogonal polynomials.*

Theorem 2.1.5 *Let $f \in L_2(\mu, \mathbb{R})$ and N be a positive integer. If $\{p_j(x) : 0 \leq j \leq N\}$ are orthonormal with respect to μ, then*

$$\inf \left\{ \int_{\mathbb{R}} \left| f(x) - \sum_{j=0}^{N} f_j \, p_j(x) \right|^2 d\mu(x) : f_j \in \mathbb{R}, \; 0 \leq j \leq N \right\}$$

is attained if and only if $f_j = \int_{\mathbb{R}} f(x)\, p_j(x)\, d\mu(x)$. The infimum is

$$\int_{\mathbb{R}} |f(x)|^2 d\mu(x) - \sum_{j=0}^{N} |f_j|^2 .$$

Proof Clearly

$$\int_{\mathbb{R}} \left| f(x) - \sum_{j=0}^{N} f_j \, p_j(x) \right|^2 d\mu(x)$$

$$= \int_{\mathbb{R}} |f(x)|^2 d\mu - 2\,\mathrm{Re} \left\{ \sum_{j=0}^{N} f_j \int_{\mathbb{R}} p_j(x)\, \overline{f(x)}\, d\mu(x) \right\} + \sum_{j=0}^{N} |f_j|^2 .$$

Let $a_j = \int_{\mathbb{R}} p_j(x)\, f(x)\, d\mu(x)$. The above expression is

$$\int_{\mathbb{R}} |f(x)|^2 d\mu(x) - \sum_{j=0}^{N} |a_j|^2 + \sum_{j=0}^{N} |f_j - a_j|^2 ,$$

which is minimized if and only if $f_j = a_j$ for $j = 0, 1, \ldots, N$. ☐

Theorem 2.1.6 *Let N be a positive integer and let*

$$q_N(x) = x^N + \sum_{j=0}^{N-1} c_j \, x^j .$$

Then

$$\inf\left\{\int_{\mathbb{R}} |q_N(x)|^2\, d\mu(x) : c_0, \ldots, c_{N-1} \in \mathbb{R}\right\}$$

is attained if and only if $q_N(x) = p_N(x)/\gamma_N$.

Proof Rewrite $q_N(x)$ as $\sum_{j=0}^{N} d_j\, p_j(x)$ with $d_N = 1/\gamma_N$. The rest follows from the orthogonality. □

Since an $N \times N$ Hankel matrix formed by moments of a positive measure is positive definite, it is natural to study the limiting behavior of its smallest eigenvalue, $\lambda_1(N)$. Surprisingly

$$\lim_{N\to\infty} \lambda_1(N) > 0$$

if and only if the moments μ_n do not determine a unique measure μ, see (Berg et al., 2002). Berg, Chen and Ismail gave a positive lower bound for the above limit.

2.2 Recurrence Relations

Since $\{P_n\}$ forms a basis for the vector space of polynomials over $\mathbb{R}$ then

$$xP_n(x) = P_{n+1}(x) + \sum_{j=0}^{n} d_j P_j(x),$$

with $d_j\zeta_j = \int_{\mathbb{R}} xP_n(x)P_j(x)\, d\mu(x)$, $0 \le j \le n$. On the other hand $xP_n(x)P_j(x)$ is $P_n(x)$ multiplied by a polynomial of degree $j+1$, hence the integrals $\int_{\mathbb{R}} xP_n(x)P_j(x)\, d\mu(x)$ vanish for $j + 1 < n$, and we find that $d_0 = \cdots = d_{n-2} = 0$.

Theorem 2.2.1 *A monic sequence of orthogonal polynomials satisfies a three-term recurrence relation*

$$xP_n(x) = P_{n+1}(x) + \alpha_n P_n(x) + \beta_n P_{n-1}(x), \quad n > 0, \tag{2.2.1}$$

with

$$P_0(x) = 1, \quad P_1(x) = x - \alpha_0, \tag{2.2.2}$$

where α_n *is real, for* $n \ge 0$ *and* $\beta_n > 0$ *for* $n > 0$.

Proof We need only to prove that $\beta_n > 0$. Clearly (2.2.1) yields

$$\begin{aligned}\beta_n\zeta_{n-1} &= \int_{\mathbb{R}} xP_n(x)P_{n-1}(x)\, d\mu(x)\\ &= \int_{\mathbb{R}} P_n(x)\,[P_n(x) + \text{lower order terms}]\, d\mu(x) = \zeta_n,\end{aligned}$$

hence $\beta_n > 0$. □

Note that we have actually proved that

$$\zeta_n = \beta_1 \cdots \beta_n. \tag{2.2.3}$$

Theorem 2.2.2 *The Christoffel–Darboux identities hold for $N > 0$*

$$\sum_{k=0}^{N-1} \frac{P_k(x)P_k(y)}{\zeta_k} = \frac{P_N(x)P_{N-1}(y) - P_N(y)P_{N-1}(x)}{\zeta_{N-1}\,(x-y)}, \tag{2.2.4}$$

$$\sum_{k=0}^{N-1} \frac{P_k^2(x)}{\zeta_k} = \frac{P_N'(x)P_{N-1}(x) - P_N(x)P_{N-1}'(x)}{\zeta_{N-1}}. \tag{2.2.5}$$

Proof Multiply (2.2.1) by $P_n(y)$ and subtract the result from the same expression with x and y interchanged. With $\Delta_k(x,y) = P_k(x)P_{k-1}(y) - P_k(y)P_{k-1}(x)$, we find

$$(x-y)P_n(x)P_n(y) = \Delta_{n+1}(x,y) - \beta_n\Delta_n(x,y),$$

which can be written in the form

$$(x-y)\frac{P_n(x)P_n(y)}{\zeta_n} = \frac{\Delta_{n+1}(x,y)}{\zeta_n} - \frac{\Delta_n(x,y)}{\zeta_{n-1}},$$

in view of (2.2.3). Formula (2.2.4) now follows from the above identity by telescopy. Formula (2.2.5) is the limiting case $y \to x$ of (2.2.4). □

Remark 2.2.1 *It is important to note that Theorem* 2.2.2 *followed from* (2.2.1) *and* (2.2.2)*, hence the Christoffel–Darboux identity* (2.2.4) *will hold for any solution of* (2.2.1)*, with possibly an additional term $c/(x-y)$ depending on the initial conditions. Similarly an identity like* (2.2.5) *will also hold.*

Theorem 2.2.3 *Assume that α_{n-1} is real and $\beta_n > 0$ for all $n = 1, 2, \ldots$. Then the zeros of the polynomials generated by* (2.2.1)–(2.2.2) *are real and simple. Furthermore the zeros of P_n and P_{n-1} interlace.*

Proof Let $x = u$ be a complex zero of P_N. Since the polynomials $\{P_n(x)\}$ have real coefficients, then $x = \overline{u}$ is also a complex zero of $P_N(x)$. With $x = u$ and $y = \overline{u}$ we see that the right-hand side of (2.2.4) vanishes while its left-hand side is larger than 1. Therefore all the zeros of P_N are real for all N. On the other hand, a multiple zero of P_N will make the right-hand side of (2.2.5) vanish while its left-hand side is positive.

Let

$$x_{N,1} > x_{N,2} > \cdots > x_{N,N} \tag{2.2.6}$$

be the zeros of P_N. Since $P_N(x) > 0$ for $x > x_{N,1}$ we see that $(-1)^{j-1}P_N'(x_{N,j}) > 0$. From this and (2.2.5) it immediately follows that $(-1)^{j-1}P_{N-1}(x_{N,j}) > 0$, hence P_{N-1} has a change of sign in each of the $N-1$ intervals $(x_{N,j+1}, x_{N,j})$, $j = 1, \ldots, N-1$, so it must have at least one zero in each such interval. This accounts for all the zeros of P_{N-1} and the theorem follows by Theorem 2.2.3. □

Another way to see the reality of zeros of polynomials generated by (2.2.1)–(2.2.2) is to relate P_n to characteristic polynomials of real symmetric matrices. Let $A_n = (a_{j,k} : 0 \le j, k < n)$ be the tridiagonal matrix

$$\begin{aligned} a_{j,j} &= \alpha_j, \quad a_{j,j+1} = 1, \\ a_{j+1,j} &= \beta_{j+1}, \quad a_{j,k} = 0, \quad \text{for } |j-k| > 1. \end{aligned} \tag{2.2.7}$$

Let $S_n(\lambda)$ be the characteristic polynomial of A_n, that is the determinant of $\lambda I - A_n$. By expanding the determinant expression for S_n about the last row it follows that $S_n(x)$ satisfies the recurrence relation (2.2.1). On the other hand $S_1(x) = x - \alpha_0$, $S_2(x) = (x-\alpha_0)(x-\alpha_1) - \beta_1$, so $S_1 = P_1$ and $S_2 = P_2$. Therefore S_n and P_n agree for all n. This establishes the following theorem.

Theorem 2.2.4 *The monic polynomials have the determinant representation*

$$P_n(x) = \begin{vmatrix} x-\alpha_0 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -\beta_1 & x-\alpha_1 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & & -\beta_{n-2} & x-\alpha_{n-2} & -1 \\ 0 & 0 & 0 & \cdots & 0 & -\beta_{n-1} & x-\alpha_{n-1} \end{vmatrix}. \tag{2.2.8}$$

It is straightforward to find an invertible diagonal matrix D so that $A_n = D^{-1}B_nD$ where $B_n = (b_{j,k})$ is a real symmetric tridiagonal matrix with $b_{j,j} = \alpha_j$, $b_{j,j+1} = \sqrt{\beta_{j+1}}$, $j = 0, \dots, n-1$. Thus the zeros of P_n are real.

Theorem 2.2.5 *Let $\{P_n\}$ be a sequence of orthogonal polynomials satisfying* (2.1.5) *and let $[a,b]$ be the smallest closed interval containing the support of μ. Then all zeros of P_n lie in $[a,b]$.*

Proof Let $c_1, \dots c_j$ be the zeros of P_n lying inside $[a,b]$. If $j < n$ then the orthogonality implies $\int\limits_{\mathbb{R}} P_n(x) \prod\limits_{k=1}^{j} (x-c_j)\, d\mu = 0$, which contradicts the fact that the integrand does not change sign on $[a,b]$ by Theorem 2.2.3. □

Several authors studied power sums of zeros of orthogonal polynomials and special functions. Let

$$s_k = \sum_{j=1}^{n} (x_{n,j})^k, \tag{2.2.9}$$

where $x_{n,1} > x_{n,2} > \cdots > x_{n,n}$ are the zeros of $P_n(x)$. Clearly

$$\frac{P_n'(x)}{P_n(x)} = \sum_{j=1}^{n} \frac{1}{x - x_{n,j}} = \sum_{k=0}^{\infty} \frac{1}{x^{k+1}} \sum_{j=1}^{n} (x_{n,j})^k.$$

Thus

$$P_n'(z)/P_n(z) = \sum_{k=0}^{\infty} s_k z^{-k-1}, \tag{2.2.10}$$

for $|z| > \max\{|x_{n,j}| : 1 \le j \le n\}$. Power sums for various special orthogonal polynomials can be evaluated using (2.2.10). If no $x_{n,j} = 0$, we can define s_k for $k < 0$ and apply

$$\frac{P_n'(z)}{P_n(z)} = -\sum_{j=1}^{n} \frac{1}{x_{n,j}} \sum_{k=0}^{\infty} \left(\frac{x}{x_{n,j}}\right)^k,$$

to conclude that

$$P_n'(z)/P_n(z) = -\sum_{k=0}^{\infty} z^k s_{-k-1}, \tag{2.2.11}$$

for $|z| < \min\{|x_{n,j}| : 1 \le j \le n\}$. Formula (2.2.11) also holds when P_n is replaced by a function with the factor product representation

$$f(z) = \prod_{k=1}^{\infty} (1 - z/x_k).$$

An example is $f(z) = \Gamma(\nu+1)(2/z)^\nu J_\nu(z)$. Examples of (2.2.10) and (2.2.11) and their applications are in (Ahmed et al., 1979), (Ahmed et al., 1982) and (Ahmed & Muldoon, 1983). The power sums of zeros of Bessel polynomials have a remarkable property as we shall see in Theorems 4.10.4 and 4.10.5.

The **Poisson kernel** $P_r(x, y)$ of a system of orthogonal polynomials is

$$P_r(x, y) = \sum_{n=0}^{\infty} P_n(x) P_n(y) \frac{r^n}{\zeta_n}. \tag{2.2.12}$$

One would expect $\lim\limits_{r \to 1^-} P_r(x, y)$ to be a Dirac measure $\delta(x - y)$. Indeed under certain conditions

$$\lim_{r \to 1^-} \int_{\mathbb{R}} P_r(x, y) f(y)\, d\mu(y) = f(x), \tag{2.2.13}$$

for $f \in L^2(\mu)$. A crucial step in establishing (2.2.13) for a specific system of orthogonal polynomials is the nonnegativity of the Poisson kernel on the support of μ.

Definition 2.2.1 *The kernel polynomials $\{K_n(x, y)\}$ of a distribution function F_μ are*

$$K_n(x, y) = \sum_{k=0}^{n} \overline{p_k(x)}\, p_k(y) = \sum_{k=0}^{n} \overline{P_k(x)}\, P_k(y)/\zeta_k, \tag{2.2.14}$$

$n = 0, 1, \ldots.$

Theorem 2.2.6 *Let $\pi(x)$ be a polynomial of degree at most n and normalized by*

$$\int_{\mathbb{R}} |\pi(x)|^2 d\mu(x) = 1. \tag{2.2.15}$$

Then the maximum of $|\pi(x_0)|^2$ taken over all such $\pi(x)$ is attained when

$$\pi(x) = \zeta K_n(x_0, x) / \sqrt{K_n(x_0, x_0)}, \quad |\zeta| = 1.$$

The maximum is $K_n(x_0, x_0)$.

Proof Assume that $\pi(x)$ satisfies (2.2.15) and let $\pi(x) = \sum_{k=0}^{n} c_k p_k(x)$. Then

$$|\pi(x_0)|^2 \le \left(\sum_{k=0}^{r} |c_k|^2\right)\left(\sum_{k=0}^{n} |p_k(x_0)|^2\right) = K_n(x_0, x_0)$$

and the theorem follows. □

Remark 2.2.2 *Sometimes it is convenient to use neither the monic nor the orthonormal polynomials. If $\{\phi_n(x)\}$ satisfy*

$$\int_{\mathbb{R}} \phi_m(x)\phi_n(x)\, d\mu(x) = \zeta_n \delta_{m,n} \tag{2.2.16}$$

then

$$x\phi_n(x) = A_n\phi_{n+1}(x) + B_n\phi_n(x) + C_n\phi_{n-1}(x) \tag{2.2.17}$$

and

$$\zeta_n = \frac{C_1 \cdots C_n}{A_0 \cdots A_{n-1}}\, \zeta_0. \tag{2.2.18}$$

The interlacing property of the zeros of P_n and P_{n+1} extend to eigenvalues of general Hermitian matrices.

Theorem 2.2.7 (Cauchy Interlace Theorem) *Let A be a Hermitian matrix of order n, and let B be a principal submatrix of A of order $n-1$. If $\lambda_n \le \lambda_{n-1} \le \cdots \le \lambda_2 \le \lambda_1$ lists the eigenvalues of A and $\mu_n \le \mu_{n-1} \le \cdots \le \mu_3 \le \mu_2$ the eigenvalues of B, then $\lambda_n \le \mu_n \le \lambda_{n-1} \le \mu_{n-1} \le \cdots \le \lambda_2 \le \mu_2 \le \lambda_1$.*

Recently, Hwang gave an elementary proof of Theorem 2.2.7 in (Hwang, 2004). The standard proof of Theorem 2.2.7 uses Sylvester's law of inertia (Parlett, 1998). For a proof using the Courant–Fischer minmax principle, see (Horn & Johnson, 1992).

2.3 Numerator Polynomials

Consider (2.2.1) as a difference equation in the variable n,

$$x\, y_n = y_{n+1} + \alpha_n y_n + \beta_n y_{n-1}. \tag{2.3.1}$$

Thus (2.3.1) has two linearly independent solutions, one of them being $P_n(x)$. We introduce a second solution, P_n^*, defined initially by

$$P_0^*(x) = 0, \quad P_1^*(x) = 1. \tag{2.3.2}$$

It is clear that $P_n^*(x)$ is monic and has degree $n-1$ for all $n > 0$. Next consider the Casorati determinant (Milne-Thomson, 1933)

$$\Delta_n(x) := P_n(x)P_{n-1}^*(x) - P_n^*(x)P_{n-1}(x).$$

From (2.2.1) we see that $\Delta_{n+1}(x) = \beta_n \Delta_n(x)$, and in view of the initial conditions (2.2.2) and (2.3.2) we establish

$$P_n(x)P_{n-1}^*(x) - P_n^*(x)P_{n-1}(x) = -\beta_1 \cdots \beta_{n-1}. \tag{2.3.3}$$

From the theory of difference equations (Jordan, 1965), (Milne-Thomson, 1933) we know that two solutions of a second order linear difference equation are linearly independent if and only if their Casorati determinant is not zero. Thus P_n and P_n^* are linearly independent solutions of (2.3.1).

Theorem 2.3.1 *For $n > 0$, the zeros of $P_n^*(x)$ are all real and simple and interlace with the zeros of $P_n(x)$.*

Proof Clearly (2.3.3) shows that $P_n^*(x_{n,j})\, P_{n-1}(x_{n,j}) > 0$. Then P_n^* and P_{n-1} have the same sign at the zeros of P_n. Now Theorem 2.2.4 shows that P_n^* has a zero in all intervals $(x_{n,j}, x_{n,j-1})$ for all j, $2 \le j < n$. □

Definition 2.3.1 *Let $\{P_n(x)\}$ be a family of monic orthogonal polynomials generated by* (2.2.1) *and* (2.2.2). *The associated polynomials $\{P_n(x; c)\}$ of order c of $P_n(x)$ are polynomials satisfying* (2.2.1) *with n replaced by $n + c$ and given initially by*

$$P_0(x; c) := 1, \quad P_1(x; c) = x - \alpha_c. \tag{2.3.4}$$

The above procedure is well-defined when $c = 1, 2, \ldots$. In general, the above definition makes sense if the recursion coefficients are given by a pattern amenable to replacing n by $n + c$. It is clear that $P_{n-1}(x; 1) = P_n^*(x)$. Clearly $\{P_n(x; c)\}$ are orthogonal polynomials if $c = 1, 2, \ldots$.

Theorem 2.3.2 *The polynomials $\{P_n^*(z)\}$ have the integral representation*

$$P_n^*(z) = \int_{\mathbb{R}} \frac{P_n(z) - P_n(y)}{z - y}\, d\mu(y), \quad n \ge 0. \tag{2.3.5}$$

Proof Let $r_n(z)$ denote the right-hand side of (2.3.5). Then $r_0(z) = 0$ and $r_1(x) = 1$. For nonreal z and $n > 0$,

$$\begin{aligned} &r_{n+1}(z) - (z - \alpha_n)\, r_n(z) + \beta_n r_{n-1}(z) \\ &= \int_{\mathbb{R}} \frac{P_{n+1}(y) - (z - \alpha_n)\, P_n(y) + \beta_n P_{n-1}(y)}{y - z}\, d\mu(y) \\ &= \int_{\mathbb{R}} \frac{y - z}{y - z}\, P_n(y)\, d\mu(y), \end{aligned}$$

which vanishes for $n > 0$. Thus $r_n(z) = P_n^*(z)$ and the restriction on z can now be removed. □

2.4 Quadrature Formulas

Let $\{P_n(x)\}$ be a sequence of monic orthogonal polynomials satisfying (2.1.5) with zeros as in (2.2.6).

Theorem 2.4.1 *Given N there exists a sequence of positive numbers $\{\lambda_k : 1 \le k \le N\}$ such that*

$$\int_{\mathbb{R}} p(x)\, d\mu(x) = \sum_{k=1}^{N} \lambda_k p\left(x_{N,k}\right), \tag{2.4.1}$$

for all polynomials p of degree at most $2N-1$. The λ's depend on N, and $\mu_0, \dots, \mu_N$ but not on p. Moreover, the λ's have the representations

$$\lambda_k = \int_{\mathbb{R}} \frac{P_N(x)\, d\mu(x)}{P_N'\left(x_{N,k}\right)\left(x - x_{N,k}\right)} \tag{2.4.2}$$

$$= \int_{\mathbb{R}} \left[\frac{P_N(x)}{P_N'\left(x_{N,k}\right)\left(x - x_{N,k}\right)}\right]^2 d\mu(x). \tag{2.4.3}$$

Furthermore if (2.4.1) *holds for all p of degree at most $2N-1$ then the λ's are unique and are given by* (2.4.2).

Proof Let L be the Lagrange interpolation polynomial of p at the nodes $x_{N,k}$, see (1.2.16)–(1.2.17). Since $L(x) = p(x)$ at $x = x_{N,j}$, for all j, $1 \le j \le N$ then $p(x) - L(x) = P_N(x) r(x)$, with r a polynomial of degree $\le N-1$. Therefore

$$\begin{aligned}
\int_{\mathbb{R}} p(x)\, d\mu(x) &= \int_{\mathbb{R}} L(x)\, d\mu(x) + \int_{\mathbb{R}} P_N(x) r(x)\, d\mu(x) \\
&= \sum_{k=1}^{N} p(x_{N,k}) \int_{\mathbb{R}} \frac{P_N(x)\, d\mu(x)}{P_N'\left(x_{N,k}\right)\left(x - x_{N,k}\right)}.
\end{aligned}$$

This establishes (2.4.2). Applying (2.4.1) to $p(x) = P_N(x)^2 / \left(x - x_{N,k}\right)^2$ we establish (2.4.3) and the uniqueness of the λ's. The positivity of the λ's follows from (2.4.3). □

The numbers $\lambda_1, \dots, \lambda_N$ are called Christoffel numbers.

Theorem 2.4.2 *The Christoffel numbers have the properties*

$$\sum_{k=1}^{N} \lambda_k = \mu(\mathbb{R}), \tag{2.4.4}$$

$$\lambda_k = -\zeta_N / \left[P_{N+1}\left(x_{N,k}\right) P_N'\left(x_{N,k}\right)\right], \tag{2.4.5}$$

$$\frac{1}{\lambda_k} = \sum_{j=0}^{N} P_j^2\left(x_{N,k}\right) / \zeta_j =: K_N\left(x_{N,k}, x_{N,k}\right). \tag{2.4.6}$$

Proof Apply (2.4.1) with $p(x) \equiv 1$ to get (2.4.4). Next replace N by $N+1$ in (2.2.4) then set $y = x_{N,k}$ and integrate with respect to μ. The result is

$$1 = -\frac{P_{N+1}\left(x_{N,k}\right)}{\zeta_N} \int_{\mathbb{R}} \frac{P_N(x)}{x - x_{N,k}}\, d\mu(x),$$

and the rule (2.4.2) implies (2.4.5). Formula (2.4.6) follows from (2.4.5) and (2.2.5). $\square$

We now come to the Chebyshev–Markov–Stieltjes separation theorem and inequalities. Let $[a, b]$ be the convex hull of the support of μ. For $N \geq 2$, we let

$$u_k = x_{N,N-k}, \tag{2.4.7}$$

so that $u_1 < u_2 < \cdots < u_N$. Let F_μ be the distribution function of μ. In view of the positivity of the λ's and (2.4.4), there exist numbers $y_1 < y_2 < \cdots < y_{N-1}$, $a < y_1$, $y_{N-1} < b$, such that

$$\lambda_k = F_\mu\left(y_k\right) - F_\mu\left(y_{k-1}\right), \quad 1 \leq k \leq N, \quad y_0 := a, \quad y_N := b. \tag{2.4.8}$$

Theorem 2.4.3 (Separation Theorem) *The points $\{y_k\}$ interlace with the zeros $\{u_k\}$; that is*

$$u_k < y_k < u_{k+1}, \quad 1 \leq k \leq N-1.$$

Equivalently

$$F_\mu\left(u_k\right) < F_\mu\left(y_k\right) = \sum_{j=1}^{k} \lambda_j < F_\mu\left(u_{k+1}^-\right), \quad 1 \leq k < N.$$

An immediate consequence of the separation theorem is the following corollary.

Corollary 2.4.4 *Let I be an open interval formed by two consecutive zeros of $P_N(x)$. Then $\mu(I) > 0$.*

Proof of Theorem 2.4.3 Define a right continuous step function V by $V(a) = 0$, $V(x) = \sum_{j=1}^{N} \lambda_j$ for $x \geq x_{N,N}$ and V has a jump λ_j at $x = u_j$. Formula (2.4.1) is

$$\int_a^b p(x)\, d(\mu - V) = 0.$$

Hence,

$$0 = \int_a^b p'(x) \left[F_\mu(x) - V(x)\right] dx. \tag{2.4.9}$$

Set $\beta(x) = F_\mu(x) - V(x)$. Let $I_1 = (a, u_1)$, $I_{j+1} = (u_j, u_{j+1})$, $1 \leq j < N$, $I_{N+1} = (u_N, b)$. Clearly, $\beta(x)$ is nondecreasing on I_j, V_j. Moreover, $\beta(x) \geq 0$, $\beta(x) \not\equiv 0$, on I_1, but $\beta(x) \leq 0$, $\beta(x) \not\equiv 0$ on I_{N+1}. On I_j, $1 < j \leq N$, β either has a constant sign or changes sign from negative to positive at some point within the interval. Such points where the change of sign may occur must be among the y points defined in (2.4.8). Thus, $[a, b]$ can be subdivided into at most $2N$ subintervals where $\beta(x)$ has a constant sign on each subinterval. If the number of intervals of β constant signs is $< 2N$, then we can choose p' in (2.4.9) to have degree at most $2n-2$ such that $p'(x)\beta(x) \geq 0$ on $[a, b]$, which gives a contradiction. Thus we must have at least $2N$ intervals where $\beta(x)$ keeps a constant sign. By the pigeonhole principle, we must have $y_j \in I_{j+1}$, $1 \leq j < N$ and the theorem follows. □

Szegő gives two additional proofs of the separation theorem; see (Szegő, 1975, §3.41).

2.5 The Spectral Theorem

The main result in this section is Theorem 2.5.2 which we call the spectral theorem for orthogonal polynomials. In some of the literature it is called Favard's theorem, (Chihara, 1978), (Szegő, 1975), because Favard (Favard, 1935) proved it in 1935. Shohat claimed to have had an unpublished proof several years before Favard published his paper (Shohat, 1936). More interestingly, the theorem is stated and proved in Wintner's book on spectral theory of Jacobi matrices (Wintner, 1929). The theorem also appeared in Marshal Stone's book (Stone, 1932, Thm. 10.27) without attributing it to any particular author. The question of uniqueness of μ is even discussed in Theorem 10.30 in (Stone, 1932).

We start with a sequence of polynomials $\{P_n(x)\}$ satisfying (2.2.1)–(2.2.2), with $\alpha_{n-1} \in \mathbb{R}$ and $\beta_n > 0$, for $n > 0$. Fix N and arrange the zeros of P_N as in (2.2.6). Define a sequence

$$\rho\left(x_{N,j}\right) = \zeta_{N-1} / \left[P_N'\left(x_{N,j}\right) P_{N-1}\left(x_{N,j}\right)\right], \quad 1 \leq j \leq N. \tag{2.5.1}$$

We have established the positivity of $\rho\left(x_{N,j}\right)$ in Theorem 2.2.4. With x a zero of $P_N(x)$, rewrite (2.2.4) and (2.2.5) as

$$\rho(x_{N,r}) \sum_{k=0}^{N-1} P_k\left(x_{N,r}\right) P_k\left(x_{N,s}\right) / \zeta_k = \delta_{r,s}. \tag{2.5.2}$$

Indeed (2.5.2) says that the real matrix U,

$$U = \left(u_{r,k}\right), \ 1 \leq r, k \leq N, \quad u_{r,k} := \sqrt{\rho\left(x_{N,r}\right)}\, \frac{P_{k-1}\left(x_{N,r}\right)}{\sqrt{\zeta_k - 1}},$$

satisfies $UU^T = I$, whence $U^TU = I$, that is

$$\sum_{r=1}^{N} \rho\left(x_{N,r}\right) P_k\left(x_{N,r}\right) P_j\left(x_{N,r}\right) = \zeta_k \delta_{j,k}, \quad j, k = 0, \ldots, N-1. \tag{2.5.3}$$

We now introduce a sequence of right continuous step functions $\{\psi_N\}$ by

$$\psi_N(-\infty) = 0, \quad \psi_N\left(x_{N,j} + 0\right) - \psi_N\left(x_{N,j} - 0\right) = \rho\left(x_{N,j}\right). \tag{2.5.4}$$

Theorem 2.5.1 *The moments* $\int_{\mathbb{R}} x^j \, d\psi_N(x)$, $1 \leq j \leq 2N-2$ *do not depend on* ζ_k *for* $k > \lfloor (j+1)/2 \rfloor$ *where* $\lfloor a \rfloor$ *denotes the integer part of* a.

Proof For fixed j choose $N > 1 + j/2$ and write x^j as $x^s x^\ell$, with $0 \leq \ell, s \leq N-1$. Then express x^s and x^ℓ as linear combinations of $P_0(x), \ldots, P_{N-1}(x)$. Thus the evaluation of $\int_{\mathbb{R}} x^j \, d\psi_N(x)$ involves only $\zeta_0, \ldots, \zeta_{\lfloor (j+1)/2 \rfloor}$. □

Theorem 2.5.2 *Given a sequence of polynomials* $\{P_n(x)\}$ *generated by* (2.2.1)–(2.2.2) *with* $\alpha_{n-1} \in \mathbb{R}$ *and* $\beta_n > 0$ *for all* $n > 0$, *then there exists a distribution function* μ *such that*

$$\int_{\mathbb{R}} P_m(x) P_n(x) \, d\mu(x) = \zeta_n \delta_{m,n}, \tag{2.5.5}$$

and ζ_n *is given by* (2.2.3).

Proof Since

$$1 = \zeta_0 = \int_{\mathbb{R}} d\psi_N(x) = \psi_N(\infty) - \psi_N(-\infty)$$

then the ψ_N's are uniformly bounded. From Helly's selection principle it follows that there is a subsequence ϕ_{N_k} which converges to a distribution function μ. The rest follows from Theorems 2.5.1 and 1.2.1. It is clear that the limiting function μ of any subsequence will have infinitely many points of increase. □

Shohat (Shohat, 1936) proved that if $\alpha_n \in \mathbb{R}$ and $\beta_{n+1} \neq 0$ for all $n \geq 0$, then there is a real signed measure μ with total mass 1 such that (2.5.5) holds, with $\zeta_0 = 1$, $\zeta_n = \beta_1 \cdots \beta_n$, see also (Shohat, 1938).

The distribution function μ in Theorem 2.5.2 may not be unique, as can be seen from the following example due to Stieltjes.

Example 2.5.3 *Consider the weight function*

$$w(x;\alpha) = [1 + \alpha \sin(2\pi c \ln x)] \exp\left(-c \ln^2 x\right), \quad x \in (0, \infty),$$

for $\alpha \in (-1, 1)$ *and* $c > 0$. *The moments are*

$$\begin{aligned}
\mu_n &:= \int_0^\infty x^n w(x;\alpha) \, dx = \int_{-\infty}^\infty e^{u(n+1)} [1 + \alpha \sin(2\pi c u)] \, \exp\left(-cu^2\right) du \\
&= \exp\left(\frac{(n+1)^2}{4c}\right) \int_{-\infty}^\infty \exp\left(-c\left(u - \frac{n+1}{2c}\right)^2\right) [1 + \alpha \sin(2\pi c u)] \, du \\
&= \exp\left(\frac{(n+1)^2}{4c}\right) \int_{-\infty}^\infty \exp\left(-cv^2\right) [1 - (-1)^n \alpha \sin(2\pi c v)] \, dv.
\end{aligned}$$

In the last step we have set $u = v + (n+1)/(2c)$. *Clearly*

$$\mu_n = \exp\left(\frac{(n+1)^2}{4c}\right) \int_{-\infty}^\infty \exp\left(-cv^2\right) dv = \sqrt{\frac{\pi}{c}} \exp\left(\frac{(n+1)^2}{4c}\right)$$

which is independent of α. Therefore the weight functions $w(x;\alpha)$, for all $\alpha \in (-1,1)$, have the same moments.

We will see that μ is unique when the α_n's and β_n's are bounded. Let

$$\xi := \lim_{n\to\infty} x_{n,n}, \quad \eta := \lim_{n\to\infty} x_{n,1}. \tag{2.5.6}$$

Both limits exist since $\{x_{n,1}\}$ increases with n while $\{x_{n,n}\}$ decreases with n, by Theorem 2.2.4. Theorem 2.2.6 and the construction of the measures μ in Theorem 2.5.2 motivate the following definition.

Definition 2.5.1 *The **true interval of orthogonality** of a sequence of polynomials $\{P_n\}$ generated by* (2.2.1)–(2.2.2) *is the interval* $[\xi,\eta]$.

It is clear from Theorem 2.4.3 that $[\xi,\eta]$ is a subset of the convex hull of $\operatorname{supp}(\mu)$.

Theorem 2.5.4 *The support of every μ of Theorem* 2.5.2 *is bounded if $\{\alpha_n\}$ and $\{\beta_n\}$ are bounded sequences.*

Proof If $|\alpha_n| < M$, $\beta_n \le M$, then apply Theorem 2.2.5 to identify the points $x_{N,j}$ as eigenvalues of a tridiagonal matrix then apply Theorem 1.1.1 to see that $|x_{n,j}| < \sqrt{3}\,M$ for all N and j. Thus the support of each ψ_N is contained in $\left(-\sqrt{3}\,M, \sqrt{3}\,M\right)$ and the result follows. □

Theorem 2.5.5 *If $\{\alpha_n\}$ and $\{\beta_n\}$ are bounded sequences then the measure of orthogonality μ is unique.*

Proof By Theorem 2.5.4 we may assume that the support of one measure μ is compact. Let ν be any other measure. For any $a > 0$, we have

$$\begin{aligned}\int\limits_{|x|\ge a} d\nu(x) &\le a^{-2n} \int\limits_{|x|\ge a} x^{2n} d\nu(x) \le a^{-2n} \int\limits_{\mathbb{R}} x^{2n} d\nu(x) \\ &= a^{-2n} \int\limits_{\mathbb{R}} x^{2n} d\mu(x).\end{aligned}$$

Assume $|x_{N,j}| \le A$, $j = 1,2,\dots,N$ and for all $N \ge 1$. Apply (2.4.1) with $N = n+1$. Then

$$\begin{aligned}\int\limits_{|x|\ge a} d\nu(x) &\le a^{-2n} \sum_{k=1}^{n+1} \lambda_k \left(x_{n+1,k}\right)^{2n} \\ &\le (A/a)^{2n} \sum_{k=1}^{n+1} \lambda_k = (A/a)^n.\end{aligned}$$

If $a > A$, then $\int\limits_{|x|\ge a} d\nu(x) = 0$, hence $\operatorname{supp}\nu \subset [-A, A]$. We now prove that $\mu = \nu$.

Clearly for $|x| \geq 2A$, $\sum_{k=0}^{n} t^k x^{-k-1}$ converges to $1/(x-t)$ for all $t \in [-A, A]$. Therefore

$$\begin{aligned}\int_{\mathbb{R}} \frac{d\mu(t)}{x-t} &= \int_{\mathbb{R}} \lim_{n\to\infty} \sum_{k=0}^{n} \frac{t^k}{x^{k+1}}\, d\mu(t) \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} \sum_{k=0}^{n} \frac{t^k}{x^{k+1}}\, d\mu(t) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\mu_k}{x^{k+1}},\end{aligned}$$

where in the last step we used the Lebesgue dominated convergence theorem, since $|t/x| \leq 1/2$. The last limit depends only on the moments, hence is the same for all μ's that have the same μ_k's. Thus $F(x) := \int_{\mathbb{R}} d\mu(t)/(x-t)$ is uniquely determined for x outside the circle $|x| = 2A$. Since F is analytic in $x \in \mathbb{C} \setminus [-A, A]$, by the identity theorem for analytic functions F is unique and the theorem follows from the Perron–Stieltjes inversion formula. □

Observe that the proof of Theorem 2.5.5 shows that $\operatorname{supp}\mu$ is the true interval of orthogonality when $\{\alpha_n\}$ and $\{\beta_n\}$ are bounded sequences.

A very useful result in the theory of moment problems is the following theorem, whose proofs can be found in (Shohat & Tamarkin, 1950; Akhiezer, 1965).

Theorem 2.5.6 *Assume that μ is unique and $\int_{\mathbb{R}} d\mu = 1$. Then μ has an atom at $x = u \in \mathbb{R}$ if and only if the series*

$$S := \sum_{n=0}^{\infty} P_n^2(u)/\xi_n \tag{2.5.7}$$

converges. Furthermore, if μ has an atom at $x = u$, then

$$\mu(\{u\}) = 1/S. \tag{2.5.8}$$

Some authors prefer to work with a positive linear functional $\mathcal{L}$ defined by $\mathcal{L}(x^n) = \mu_n$. The question of constructing the orthogonality measure then becomes a question of finding a representation of $\mathcal{L}$ as an integral with respect to a positive measure. This approach is used in Chihara (Chihara, 1978). Some authors even prefer to work with a linear functional which is not necessarily positive, but the determinants D_n of (2.1.4) are assumed to be nonzero. Such functionals are called regular. An extensive theory of polynomials orthogonal with respect to regular linear functionals has been developed by Pascal Maroni and his students and collaborators.

Boas (Boas, Jr., 1939) proved that any sequence of real numbers $\{c_n\}$ has a moment representation $\int_E x^n\, d\mu(x)$ for a nontrivial finite signed measure μ. In particular, the sequence $\{\mu_n\}$, $\mu_n = 0$ for all $n = 0, 1, \ldots$, is a moment sequence. Here, E

can be $\mathbb{R}$ or $[a, \infty)$, $a \in \mathbb{R}$. For example

$$0 = \int_0^\infty x^n \sin(2\pi c \ln x) \exp\left(-c \ln x^2\right) dx, \quad c > 0 \tag{2.5.9}$$

$$0 = \int_0^\infty x^n \sin x^{1/4} \exp\left(-x^{1/4}\right) dx. \tag{2.5.10}$$

Formula (2.5.9) follows from Example 2.5.3. To prove (2.5.10), observe that its right-hand side, with $x = u^4$, is

$$\begin{aligned} -2i \int_0^\infty & u^{4n+3} \left[\exp(-u(1+i)) - \exp(-u(1-i))\right] du \\ &= \frac{2i(4n+3)!}{(1+i)^{4n+3}} - \frac{2i(4n+3)!}{(1-i)^{4n+3}} = 0. \end{aligned}$$

Thus the signed measure in Boas' theorem is never unique. A nontrivial measure whose moments are all zero is called a polynomial killer.

Although we cannot get into the details of the connection between constructing orthogonality measures and the spectral problem on a Hilbert space, we briefly describe the connection. Consider the operator T which is multiplication by x. The three-term recurrence relation gives a realization of T as a tridiagonal matrix operator

$$T = \begin{pmatrix} \alpha_0 & a_1 & 0 & \cdots \\ a_1 & \alpha_1 & a_2 & \cdots \\ & \ddots & \ddots & \ddots \end{pmatrix} \tag{2.5.11}$$

defined on a dense subset of ℓ^2. It is clear that T is symmetric. When T is self-adjoint there exists a unique measure supported on $\sigma(T)$, the spectrum of T, such that

$$T = \int_{\sigma(T)} \lambda dE_\lambda.$$

In other words

$$(y, p(T)x) = \int_{\sigma(T)} p(\lambda)\,(y, dE_\lambda x)\,, \tag{2.5.12}$$

for polynomial p, and for all $x, y \in \ell^2$. By choosing the basis $e_0, e_1, \ldots$ for ℓ^2, $e_n = (u_1, u_2, \ldots)$, $u_k = \delta_{kn}$, we see that $(e_0, dE_\lambda e_0)$ is a positive measure. This is the measure of orthogonality of $\{P_n(x)\}$. One can evaluate $(e_n, dE_\lambda e_m)$, for all $m, n \geq 0$, from the knowledge of $(e_0 dE_\lambda e_0)$. Hence dE_λ can be computed from the knowledge of the measure with respect to which $\{P_n(x)\}$ are orthogonal. The details of this theory are in (Akhiezer, 1965) and (Stone, 1932).

One can think of the operator T in (2.5.11) as a discrete Schrödinger operator (Cycon et al., 1987). This analogy is immediate when $a_n = 1$. One can think of the diagonal entries $\{\alpha_n\}$ as a potential. There is extensive theory known for doubly infinite Jacobi matrices with 1's on the super and lower diagonals and a random potential α_n; see (Cycon et al., 1987, Chapter 9). The theory of general doubly infinite tridiagonal matrices is treated in (Berezans'kiĭ, 1968). Many problems in this area remain open.

2.6 Continued Fractions

A continued J-fraction is

$$\frac{A_0}{A_0 z + B_0 -}\;\frac{C_1}{A_1 z + B_1 -}\cdots. \tag{2.6.1}$$

The nth convergent of the above continued fraction is the rational function

$$\frac{A_0}{A_0 z + B_0 -}\;\frac{C_1}{A_1 z + B_1 -}\cdots\frac{C_{n-1}}{A_{n-1} z + B_{n-1}}. \tag{2.6.2}$$

Write the nth convergent as $N_n(z)/D_n(z)$, $n > 1$, and the first convergent is $\dfrac{A_0}{A_0 z + B_0}$.

Definition 2.6.1 *The J-fraction* (2.6.1) *is of positive type if* $A_n A_{n-1} C_n > 0$, $n = 1, 2, \ldots$.

Theorem 2.6.1 *Assume that* $A_n C_{n+1} \neq 0$, $n = 0, 1, \ldots$. *Then the polynomials* $N_n(z)$ *and* $D_n(z)$ *are solutions of the recurrence relation*

$$y_{n+1}(z) = [A_n z + B_n]\, y_n(z) - C_n\, y_{n-1}(z), \quad n > 0, \tag{2.6.3}$$

with the initial values

$$D_0(z) := 1,\; D_1(z) := A_0 z + B_0, \quad N_0(z) := 0,\; N_1(z) := A_0. \tag{2.6.4}$$

Proof It is easy to check that $N_1(z)/D_1(z)$ and $N_2(z)/D_2(z)$ agree with what (2.6.3) and (2.6.4) give for $D_2(z)$ and $N_2(z)$. Now assume that $N_n(z)$ and $D_n(z)$ satisfy (2.6.3) for $n = 1, \ldots, N-1$. Since the $N+1$ convergent is

$$\begin{aligned} &\frac{N_{N+1}(z)}{D_{N+1}(z)} \\ &= \frac{A_0}{A_0 z + B_0 -}\;\frac{C_1}{A_1 z + B_1 -}\cdots\frac{C_{N-1}}{A_{N-1} z + B_{N-1} - C_N/\left(A_N z + B_N\right)} \end{aligned} \tag{2.6.5}$$

then $N_{N+1}(z)$ and $D_{N+1}(z)$ follow from $N_N(z)$ and $D_N(z)$ by replacing C_{N-1} and $A_{N-1}z+B_{N-1}$ by $(A_N z + B_N)\, C_{N-1}$ and $(A_{N-1} z + B_{N-1})\,(A_N z + B_N) - C_N$, respectively. In other words

$$\begin{aligned} D_{N+1} = {} & [(A_{N-1} z + B_{N-1})\,(A_N z + B_N) - C_N]\, D_{N-1} \\ & - C_{N-1}\,(A_N z + B_n)\, D_{N-2}, \end{aligned}$$

which yields (2.6.3) for $n = N$ and $y_N = D_N$. Similarly we establish the recursions for the N_n's. $\square$

When $A_n = 1$ then D_n and N_n become P_n and P_n^* of §2.3, respectively.

Theorem 2.6.2 (Markov) *Assume that the true interval of orthogonality* $[\xi, \eta]$ *is bounded. Then*

$$\lim_{n\to\infty} \frac{P_n^*(z)}{P_n(z)} = \int_\xi^\eta \frac{d\mu(t)}{z-t}, \quad z \notin [\xi,\eta], \tag{2.6.6}$$

and the limit is uniform on compact subsets of $\mathbb{C} \setminus [\xi,\eta]$.

Proof The Chebyshev–Markov–Stieltjes inequalities and Theorem 2.5.5 imply that μ is unique and supported on E, say, $E \subset [\xi,\eta]$. Since $\int_E x^m d\psi_n(x) \to \int_E x^m d\mu$ for all m then $\int_E f d\psi_n \to \int f d\mu$ for every continuous function f. The function $1/(z-t)$ is continuous for $t \in E$, and $\operatorname{Im} z \neq 0$, whence

$$\frac{P_n^*(z)}{P_n(z)} = \int_\xi^\eta \frac{d\psi_n(t)}{z-t} \to \int_\xi^\eta \frac{d\mu(t)}{z-t}, \quad z \notin [\xi,\eta].$$

The uniform convergence follows from Vitali's theorem. □

Markov's theorem is very useful in determining orthogonality measures for orthogonal polynomials from the knowledge of the recurrence relation they satisfy.

Definition 2.6.2 *A solution* $\{u_n(z)\}$ *of* (2.6.3) *is called a minimal solution at* ∞ *if* $\lim_{n\to\infty} u_n(z)/v_n(z) = 0$ *for any other linear independent solution* $v_n(z)$. *A minimal solution at* $-\infty$ *is similarly defined.*

The minimal solution is the discrete analogue of the principal solution of differential equation.

It is clear that if the minimal solution exists then it is unique, up to a multiplicative function of z. The following theorem of Pincherle characterizes convergence of continued fractions in terms of the existence of the minimal solution.

Theorem 2.6.3 (Pincherle) *The continued fraction* (2.6.2) *converges at* $z = z_0$ *if and only if the recurrence relation* (2.6.3) *has a minimal solution* $\left\{y_n^{(\min)}(z)\right\}$. *Furthermore if a minimal solution* $\left\{y_n^{(\min)}(z)\right\}$ *exists then the continued fraction converges to* $y_0^{(\min)}(z)/y_{-1}^{(\min)}(z)$.

For a proof, see (Jones & Thron, 1980, pp. 164–166) or (Lorentzen & Waadeland, 1992, pp. 202–203).

Pincherle's theorem is very useful in finding the functions to which the convergents of a continued fraction converge. On the other hand, finding minimal solutions is not always easy but has been done in many interesting specific cases by David Masson and his collaborators. The following theorem whose proof appears in (Lorentzen & Waadeland, 1992, §4.2.2) is useful in verifying whether a solution is minimal.

Theorem 2.6.4 *Let $\{u_n(z)\}$ be a solution to* (2.2.1) *and assume that $u_n(\zeta) \neq 0$ for all n and a fixed ζ. Then $\{u_n(\zeta)\}$ is a minimal solution to* (2.6.3) *at ∞ if and only if*

$$\sum_{n=1}^{\infty} \frac{\prod_{m=1}^{n} \beta_m}{(u_n(\zeta)u_{n+1}(\zeta))} = \infty.$$

2.7 Modifications of Measures: Christoffel and Uvarov

Given a distribution function F_μ and the corresponding orthogonal polynomials $P_n(x)$, an interesting question is what can we say about the polynomials orthogonal with respect to $\Phi(x)\,d\mu(x)$ where Φ is a function positive on the support of μ, and their recursion coefficients. In this section we give a formula for the polynomials whose measure of orthogonality is $\Phi(x)\,d\mu(x)$, when Φ is a polynomial or a rational function. The modification of a measure by multiplication by a polynomial or a rational function can also be explained through the Darboux transformation of integrable systems. For details and some of the references to the literature the reader may consult (Bueno & Marcellán, 2004). In the next section we study the recursion coefficients when Φ is an exponential function.

Theorem 2.7.1 (Christoffel) *Let $\{P_n(x)\}$ be monic orthogonal polynomials with respect to μ and let*

$$\Phi(x) = \prod_{k=1}^{m} (x - x_k) \tag{2.7.1}$$

be nonnegative on the support of $d\mu$. If the x_k's are simple zeros then the polynomials $S_n(x)$ defined by

$$C_{n,m}\Phi(x)S_n(x) = \begin{vmatrix} P_n(x_1) & P_{n+1}(x_1) & \cdots & P_{n+m}(x_1) \\ P_n(x_2) & P_{n+1}(x_2) & \cdots & P_{n+m}(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ P_n(x_m) & P_{n+1}(x_m) & \cdots & P_{n+m}(x_m) \\ P_n(x) & P_{n+1}(x) & \cdots & P_{n+m}(x) \end{vmatrix} \tag{2.7.2}$$

with

$$C_{n,m} = \begin{vmatrix} P_n(x_1) & P_{n+1}(x_1) & \cdots & P_{n+m-1}(x_1) \\ P_n(x_2) & P_{n+1}(x_2) & \cdots & P_{n+m-1}(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ P_n(x_m) & P_{n+1}(x_m) & \cdots & P_{n+m-1}(x_m) \end{vmatrix}, \tag{2.7.3}$$

are orthogonal with respect to $\Phi(x)\,d\mu(x)$, and S_n has degree n. If the zero x_k has multiplicity $r > 1$, then we replace the corresponding rows of (2.7.2) *by derivatives of order $0, 1, \dots, r-1$ at x_k.*

Proof If $C_{n,m} = 0$, there are constants $c_0, \dots, c_{m-1}$, not all zero, such that the polynomial $\pi(x) := \sum_{k=0}^{m-1} c_k P_{n+k}(x)$, vanishes at $x = x_1, \dots, x_m$. Therefore $\pi(x) =$

$\Phi(x)G(x)$, and G has degree at most $n-1$. But this makes $\int_{\mathbb{R}} \Phi(x)G^2(x)\, d\mu(x) = 0$, in view of the orthogonality of the P_n's. Thus $G(x) \equiv 0$, a contradiction. Now assume that all the x_k's are simple. It is clear that the right-hand side of (2.7.2) vanishes at $x = x_1, \ldots, x_m$. Define $S_n(x)$ by (2.7.2), hence S_n has degree $\leq n$. Obviously the right-hand side of (2.7.2) is orthogonal to any polynomial of degree $< n$ with respect to $d\mu$. If the degree of $S_n(x)$ is $< n$, then $\int_{\mathbb{R}} \Phi(x)S_n^2(x)\, d\mu(x)$ is zero, a contradiction, since all the zero of Φ lie outside the support of μ. Whence S_n has exact degree n. The orthogonality of S_n to all polynomials of degree $< n$ with respect to $\Phi\, d\mu$ follows from (2.7.2) and the orthogonality of P_n to all polynomials of degree $< n$ with respect to $d\mu$. The case when some of the x_k are multiple zeros is similarly treated and the theorem follows. □

Observe that the special case $m = 1$ of (2.7.2) shows that the kernel polynomials $\{K_n(x, c)\}$, see Definition 2.2.1, are orthogonal with respect to $(x - c)\, d\mu(x)$.

Formula (2.7.2) is useful when m is small but, in general, it establishes the fact that $\Phi(x)S_n(x)$ is a linear combination of $P_n(x), \ldots, P_{n+m}(x)$ and it may by possible to evaluate the coefficients in a different way, for example by equating coefficients of $x^{n+m}, \ldots, x^n$.

Given a measure μ define

$$d\nu(x) = \frac{\prod\limits_{i=1}^{m} (x - x_i)}{\prod\limits_{j=1}^{k} (x - y_j)}\, d\mu(x), \tag{2.7.4}$$

where the products $\prod\limits_{i=1}^{m} (x - x_i)$, and $\prod\limits_{j=1}^{k} (x - y_j)$ are positive for x in the support of μ. We now construct the polynomials orthogonal with respect to ν.

Lemma 2.7.2 *Let $y_1, \ldots, y_k$ be distinct complex numbers and for $s = 0, 1, \ldots, k-1$, let*

$$\frac{1}{\prod\limits_{j=s+1}^{k} (x - y_j)} = \sum_{j=s+1}^{k} \frac{u_j(s)}{x - y_j} \tag{2.7.5}$$

Then we have, for $0 \leq \ell \leq k - s - 1$,

$$\sum_{j=s+1}^{k} u_j(s)\, y_j^{\ell} = \delta_{\ell, k-s-1}. \tag{2.7.6}$$

Proof The case $\ell = 0$ follows by multiplying (2.7.5) by x and let $x \to \infty$. By induction and repeated calculations of the residue at $x = \infty$, we establish (2.7.6). □

Theorem 2.7.3 (Uvarov) *Let ν be as in* (2.7.4) *and assume that $\{P_n(x;m,k)\}$ are orthogonal with respect to ν. Set*

$$\tilde{Q}_n(x) := \int_{\mathbb{R}} \frac{P_n(y)}{x-y}\, d\mu(y). \tag{2.7.7}$$

Then for $n \geq k$ we have

$$\begin{aligned}&\left[\prod_{i=1}^{m}(x-x_i)\right] P_n(x;m,k)\\ &= \begin{vmatrix} P_{n-k}(x_1) & P_{n-k+1}(x_1) & \cdots & P_{n+m}(x_1)\\ \vdots & \vdots & \vdots & \vdots \\ P_{n-k}(x_m) & P_{n-k+1}(x_m) & \cdots & P_{n+m}(x_m)\\ \tilde{Q}_{n-k}(y_1) & \tilde{Q}_{n-k+1}(y_1) & \cdots & \tilde{Q}_{n+m}(y_1)\\ \vdots & \vdots & \vdots & \vdots \\ \tilde{Q}_{n-k}(y_k) & \tilde{Q}_{n-k+1}(y_k) & \cdots & \tilde{Q}_{n+m}(y_k)\\ P_{n-k}(x) & P_{n+1}(x) & \cdots & P_{n+m}(x)\end{vmatrix}.\end{aligned} \tag{2.7.8}$$

If $n < k$ then

$$\begin{aligned}&\left[\prod_{i=1}^{m}(x-x_i)\right] P_n(x;m,k)\\ &= \begin{vmatrix} a_{1,1} & \cdots & a_{1,k-n} & P_0(x_1) & \cdots & P_{n+m}(x_1)\\ \vdots & \vdots & \vdots & \vdots & & \\ a_{m,1} & \cdots & a_{m,k-n} & P_0(x_m) & \cdots & P_{n+m}(x_m)\\ b_{1,1} & \cdots & b_{1,k-n} & \hat{Q}_0(y_1) & \cdots & \tilde{Q}_{n+m}(y_1)\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{k,1} & \cdots & b_{k,k-n} & \tilde{Q}_0(y_k) & \cdots & \tilde{Q}_{n+m}(y_k)\\ c_1 & \cdots & c_{k-n} & P_0(x) & \ldots & P_{n+m}(x)\end{vmatrix},\end{aligned} \tag{2.7.9}$$

where

$$b_{ij} = y_i^{j-1}, \quad 1 \leq i \leq k, \quad 1 < j \leq k-n,$$

$a_{ij} = 0$*;* $1 \leq i \leq m$*,* $1 \leq j \leq k-n$*,* $c_j = 0$.

If an x_j (or y_l is repeated r times, then the corresponding r rows will contain $P_s(x_j), \ldots, P_s^{(r-1)}(x_j)$ ($\tilde{Q}_s(x_j), \ldots, \tilde{Q}_s^{(r-1)}(x_j)$), respectively.

Uvarov proved this result in a brief announcement (Uvarov, 1959) and later gave the details in (Uvarov, 1969). The proof given below is a slight modification of Uvarov's original proof.

Proof of Theorem 2.7.3 Let $\pi_j(x)$ denote a generic polynomials in x of degree at most j and denote the determinant on the right-hand side of (2.7.8) by $\Delta_{k,m,n}(x)$. Clearly $\Delta_{k,m,n}(x)$ vanishes at the points $x = x_j$, with $1 \leq j \leq m$ so let $\Delta_{k,m,n}(x)$

be $S_n(x) \prod_{i=1}^{m} (x - x_i)$ with S_n of degree at most n. Moreover, $S_n(x) \not\equiv 0$, so we let

$$S_n(x) = \pi_{n-k}(x) \prod_{i=1}^{k} (x - y_i) + \pi_{k-1}(x),$$

and note the partial fraction decomposition

$$\frac{\pi_{k-1}(x)}{\prod_{i=1}^{k} (x - y_i)} = \sum_{j=1}^{k} \frac{\alpha_j}{x - y_j}.$$

With ν as in (2.7.4) we have

$$\begin{aligned}
\int_{\mathbb{R}} S_n^2(x)\, d\nu(x) &= \int_{\mathbb{R}} S_n(x) \prod_{j=1}^{m} (x - x_j) \left\{ \pi_{n-k}(x) + \frac{\pi_{k-1}(x)}{\prod_{i=1}^{k} (x - y_i)} \right\} d\mu(x) \\
&= \int_{\mathbb{R}} S_n(x) \prod_{j=1}^{m} (x - x_j)\, \pi_{n-k}(x)\, d\mu(x) \\
&\quad + \sum_{j=1}^{k} \alpha_j \int_{\mathbb{R}} \frac{\Delta_{k,m,n}(x)}{x - y_j}\, d\mu(x).
\end{aligned}$$

The term involving the sum in the last equality is zero because the last row in the integrated determinant coincides with one of the rows containing the $\tilde{Q}$ functions. If the degree of S_n is $< n$ then the first term in the last equality also vanishes because now $\pi_{n-k}(x)$ is $\pi_{n-k-1}(x)$ and the term we are concerned with is $\int_{\mathbb{R}} \Delta_{k,m,n}(x)\pi_{n-k-1}(x)\, d\mu(x)$, which obviously is zero. Thus S_n has exact degree n, and $\Delta \neq 0$, where

$$\Delta := \begin{vmatrix}
P_{n-k}(x_1) & P_{n-k+1}(x_1) & \cdots & P_{n+m}(x_1) \\
\vdots & \vdots & \vdots & \vdots \\
P_{n-k}(x_m) & P_{n-k+1}(x_m) & \cdots & P_{n+m}(x_m) \\
\tilde{Q}_{n-k}(y_1) & \tilde{Q}_{n-k+1}(y_1) & \cdots & \tilde{Q}_{n+m}(y_1) \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{Q}_{n-k}(y_k) & \tilde{Q}_{n-k+1}(y_k) & \cdots & \tilde{Q}_{n+m}(y_k)
\end{vmatrix}.$$

It is evident that from the determinant representation that $S_n(x)$ is orthogonal to any polynomial of degree $< n$ with respect to $d\nu$.

Similarly, we denote the determinant on the right-hand side of (2.7.9) by $\Delta_{k,m,n}(x)$, and it is divisible by $\prod_{j=1}^{m} (x - x_j)$, so we set $\Delta_{k,m,n}(x) = P_n(x; m, k) \prod_{i=1}^{m} (x - x_i)$. To prove that $\int_{\mathbb{R}} x^s P_n(x; m, k)\, d\nu(x) = 0$ for $0 \leq s < n$, it is sufficient to prove that

$\int_{\mathbb{R}} \prod_{j=1}^{s} (x - y_j)\, P_n(x; m, k)\, d\nu(x) = 0$, for $0 \le s < n$, that is

$$\int_{\mathbb{R}} \frac{\Delta_{k,m,n}(x)\, d\mu(x)}{\prod_{j=s+1}^{k} (x - y_j)} = 0, \quad s = 0, 1, \dots, n.$$

This reduces the problem to showing that the determinant in (2.7.9) vanishes if $P_\ell(x)$ is replaced by

$$\int_{\mathbb{R}} \frac{P_\ell(x)\, d\mu(x)}{\prod_{j=s+1}^{k} (x - y_j)}, \quad \ell = 0, 1, \dots, n + m.$$

Let D denote the determinant in (2.7.9) with P_ℓ replaced by the above integral. Hence, by expanding the reciprocal of the product as in (2.7.5) we find

$$\int_{\mathbb{R}} \frac{P_\ell(x)\, d\mu(x)}{\prod_{j=s+1}^{k} (x - y_j)} = \sum_{j=s+1}^{k} u_j(s) \tilde{Q}_\ell\,(y_j)\,.$$

By adding linear combinations of rows to the last row we can replace the last $n + m + 1$ entries in the last row of D to zero. This changes the entry in the last row and column ℓ to $-\sum_{j=s+1}^{k} u_j(s) b_{j,\ell}$, that is $-\sum_{j=s+1}^{k} u_j(s) y_j^{\ell-1}$. This last quantity is $-\delta_{\ell-1,k-s-1}$ by Lemma 2.7.2. The latter quantity is zero since $1 \le \ell \le k - n$ and $k - n < k - s$. □

2.8 Modifications of Measures: Toda

In this section we study modifying a measure of orthogonality by multiplying it by the exponential of a polynomial.

The Toda lattice equations describe the oscillations of an infinite system of points joined by spring masses, where the interaction is exponential in the distance between two spring masses (Toda, 1989). The semi-infinite Toda lattice equations in one time variable are

$$\dot{\alpha}_n(t) = \beta_n(t) - \beta_{n+1}(t), \quad n \ge 0, \tag{2.8.1}$$

$$\dot{\beta}_n(t) = \beta_n(t)\left[\alpha_{n-1}(t) - \alpha_n(t)\right], \tag{2.8.2}$$

where we followed the usual notation $\dot{f} = \dfrac{df}{dt}$. We shall show how orthogonal polynomials can be used to provide an explicit solution to (2.8.1)–(2.8.2).

Start with a system of monic polynomials $\{P_n(x)\}$ orthogonal with respect to μ and construct the recursion coefficients α_n and β_n in (2.2.1)–(2.2.2).

Theorem 2.8.1 *Let μ be a probability measure with finite moments, and let α_n and β_n be the recursion coefficients of the corresponding monic orthogonal polynomials. Let $P_n(x, t)$ be the monic polynomials orthogonal with respect to* $\exp(-xt)\, d\mu(x)$

under the additional assumption that the moments $\int_{\mathbb{R}} x^n \exp(-xt)\, d\mu(x)$ *exist for all* n*,* $n \geq 0$*. Let* $\alpha_n(t)$ *and* $\beta_n(t)$ *be the recursion coefficients for* $P_n(x,t)$*. Then* $\alpha_n(t)$ *and* $\beta_n(t)$ *solve the system* (2.8.1)–(2.8.2) *with the initial conditions* $\alpha_n(0) = \alpha_n$ *and* $\beta_n(0) = \beta_n$*.*

Proof First observe that the degree of $\dot{P}_n(x,t)$ is at most $n-1$. Let $\beta_0(t) = \int_{\mathbb{R}} e^{-xt}\, d\mu(x)$. Replace ζ_n in (2.1.5) by $\beta_0(t)\beta_1(t)\cdots\beta_n(t)$ then differentiate with respect to t to obtain

$$\beta_0(t)\beta_1(t)\cdots\beta_n(t)\sum_{k=0}^{n}\frac{\dot{\beta}_k(t)}{\beta_k(t)} = 0 - \int_{\mathbb{R}} xP_n^2(x,t)e^{-xt}\, d\mu(x).$$

Formula (2.2.1) implies

$$\alpha_n(t)\zeta_n(t) = \alpha_n(t)\beta_0(t)\beta_1(t)\cdots\beta_n(t) = \int_{\mathbb{R}} xP_n^2(x,t)e^{-xt}\, d\mu(x), \tag{2.8.3}$$

hence

$$\beta_0(t)\beta_1(t)\cdots\beta_n(t)\sum_{k=0}^{n}\frac{\dot{\beta}_k(t)}{\beta_k(t)} = -\alpha_n(t)\beta_0(t)\beta_1(t)\cdots\beta_n(t), \tag{2.8.4}$$

and (2.8.2) follows. Next differentiate (2.8.3) to find

$$\begin{aligned}
&\dot{\alpha}_n(t)\beta_0(t)\beta_1(t)\cdots\beta_n(t) + \alpha_n(t)\beta_0(t)\beta_1(t)\cdots\beta_n(t)\sum_{k=0}^{n}\frac{\dot{\beta}_k(t)}{\beta_k(t)} \\
&\quad = -\zeta_{n+1}(t) - \alpha_n^2(t)\zeta_n(t) - \beta_n^2(t)\zeta_{n-1}(t) \\
&\quad +2\beta_n(t)\int_{\mathbb{R}} \dot{P}_n(x,t)P_{n-1}(x,t)e^{-xt}\, d\mu(x).
\end{aligned}$$

The remaining integral can be evaluated by differentiating

$$0 = \int_{\mathbb{R}} P_n(x,t)P_{n-1}(x,t)e^{-xt}\, d\mu(x),$$

which implies

$$\int_{\mathbb{R}} \dot{P}_n(x,t)P_{n-1}(x,t)e^{-xt}\, d\mu(x) = \int_{\mathbb{R}} xP_n(x,t)P_{n-1}(x,t)e^{-xt}\, d\mu(x).$$

This yields

$$\int_{\mathbb{R}} \dot{P}_n(x,t)P_{n-1}(x,t)e^{-xt}\, d\mu(x) = \beta_n(t)\zeta_{n-1}(t).$$

Combining the above calculations with (2.8.4) we establish (2.8.1). □

The multitime Toda lattice equations can be written in the form

$$\partial_{t_j} Q = \left[Q, \left(Q^j\right)_+\right], \quad j = 1, \ldots, M, \tag{2.8.5}$$

where Q is the tridiagonal matrix with entries (q_{ij}), $q_{ii} = \alpha_i$, $q_{i,i+1} = 1$, $q_{i+1,i} = \beta_{i+1}$, $i = 0, 1, 2, \ldots$, and $q_{ij} = 0$ if $|i-j| > 1$. For a matrix A, $(A)_+$ means replace all the entries below the main diagonal by zeros.

We start with a tridiagonal matrix Q formed by the initial values of α_n and β_n and find a measure of the orthogonal polynomials. We form a new probability measure according to

$$d\mu\,(x;\mathbf{t}) = \frac{1}{\zeta_0(\mathbf{t})} \exp\left(-\sum_{s=1}^{M} t_s x^s\right) d\mu(x), \tag{2.8.6}$$

where $\mathbf{t}$ stands for $(t_1, \ldots, t_M)$. Let the corresponding monic orthogonal polynomials be $\{P_n\,(x;\mathbf{t})\}$, and $\{\alpha_n\,(\mathbf{t})\}$ and $\{\beta_n\,(\mathbf{t})\}$ be their recursion coefficients. Then the matrix $Q\,(\mathbf{t})$ be formed by the new recursion coefficients solves (2.8.5).

The partition function is

$$\begin{aligned} Z_n(\mathbf{t}) := \frac{1}{\zeta_0^n(\mathbf{t})} \int\limits_{\mathbb{R}^n} \exp\left(-\sum_{j=1}^{n}\sum_{s=1}^{M} t_s x_j^s\right) \\ \times \prod_{1\le i<j\le n} (x_i - x_j)^2\, d\mu\,(x_1)\cdots d\mu\,(x_n) \end{aligned} \tag{2.8.7}$$

and the tau function is

$$\tau_n(\mathbf{t}) = Z_n(\mathbf{t})/n! \tag{2.8.8}$$

Formulas (2.1.9) and (2.1.13) establish

$$\tau_{n+1}(\mathbf{t}) = D_n = \prod_{j=1}^{n} \beta_j^{n-j+1}. \tag{2.8.9}$$

2.9 Modification by Adding Finite Discrete Parts

We consider modification of a measure by adding a finite discrete part and express the polynomials orthogonal with respect to the new measure in terms of the polynomials orthogonal with respect to the old measure. This material is from Uvarov's very interesting paper (Uvarov, 1969).

Let

$$\nu(x) = \mu(x) + \sum_{j=1}^{r} m_j \epsilon_{x_j}, \tag{2.9.1}$$

where ϵ_u is an atomic measure concentrated at $x = u$. Let $\{P_n(x)\}$ and $\{R_n(x)\}$ be monic polynomials orthogonal with respect to μ and ν, respectively, with $\zeta_n = \int\limits_{\mathbb{R}} P_n^2\, d\mu$. Set

$$R_n(x) = \sum_{s=0}^{n} C_s P_s(x), \quad C_n = 1. \tag{2.9.2}$$

Since $\int_{\mathbb{R}} R_n P_s \, d\nu = 0$ for $s < n$, we get

$$\zeta_s C_s + \sum_{j=1}^{r} m_j R_n\left(x_j\right) P_s\left(x_j\right) = 0, \quad 0 \le s < n.$$

Substituting this evaluation of C_s into (2.9.2) and using $C_n = 1$ we obtain

$$R_n(x) = P_n(x) - \sum_{j=1}^{r} m_j R_n\left(x_j\right) a_j(x), \tag{2.9.3}$$

with

$$a_j(x) = \sum_{s=0}^{n-1} \frac{P_s(x) P_s\left(x_j\right)}{\zeta_s},$$

so that

$$\begin{aligned} a_j(x) &= \frac{P_n(x) P_{n-1}\left(x_j\right) - P_n\left(x_j\right) P_{n-1}(x)}{\zeta_{n-1}\left(x - x_j\right)}, \quad x \ne x_j \\ a_j\left(x_j\right) &= \frac{P_n'\left(x_j\right) P_{n-1}\left(x_j\right) - P_n\left(x_j\right) P_{n-1}'\left(x_j\right)}{\zeta_{n-1}} \end{aligned} \tag{2.9.4}$$

upon the use of the Christoffel–Darboux formula (2.2.4). Set

$$a_{ij} = a_j\left(x_i\right). \tag{2.9.5}$$

The system of equations (2.9.3) can be written as

$$R_n\left(x_i\right) + \sum_{j=1}^{r} a_{ij} m_j R_n\left(x_j\right) = P_n\left(x_i\right), \quad 1 \le i \le r, \tag{2.9.6}$$

and the values $R_n\left(x_j\right)$, $1 \le j \le r$ can be found in terms of the evaluations $\left\{P_n\left(x_j\right)\right\}_{j=1}^{r}$. Indeed (2.9.6) is

$$\begin{pmatrix} R_n\left(x_1\right) \\ \vdots \\ R_n\left(x_r\right) \end{pmatrix} = (I + A)^{-1} \begin{pmatrix} P_n\left(x_1\right) \\ \vdots \\ P_n\left(x_r\right) \end{pmatrix}$$

with

$$A = \begin{pmatrix} m_1 a_1\left(x_1\right) & m_2 a_2\left(x_1\right) & \cdots & m_r a_r\left(x_1\right) \\ m_1 a_1\left(x_2\right) & m_2 a_2\left(x_2\right) & \cdots & m_r a_r\left(x_2\right) \\ \vdots & \vdots & \vdots & \vdots \\ m_1 a_1\left(x_r\right) & m_2 a_2\left(x_r\right) & \cdots & m_r a_r\left(x_r\right) \end{pmatrix}.$$

Thus $R_n(x)$ is given by

$$R_n(x) = P_n(x) - \sum_{j=1}^{r} m_j R_n\left(x_j\right) a_j(x). \tag{2.9.7}$$

As an example consider the Jacobi polynomials where $d\mu$ is $(1-x)^{\alpha}(1+x)^{\beta}\, dx$, see Chapter 4. Let

$$A d\nu(x) = (1-x)^{\alpha}(1+x)^{\beta}\, dx + M_1 \epsilon_1 + M_2 \epsilon_{-1},$$

and A is a normalization constant, as

$$A = 2^{\alpha+\beta+1}\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)} + M_1 + M_2.$$

The monic Jacobi polynomials are (4.1.1) and (4.1.4)

$$\begin{aligned} P_n(x) &= \frac{(\alpha+1)_n 2^n}{(\alpha+\beta+n+1)_n}\, {}_2F_1\left(\left.\begin{matrix} -n, \alpha+\beta+n+1 \\ \alpha+1 \end{matrix}\,\right|\, \frac{1-x}{2}\right) \\ &= \frac{(-2)^n(\beta+1)_n}{(\alpha+\beta+n+1)_n}\, {}_2F_1\left(\left.\begin{matrix} -n, \alpha+\beta+n+1 \\ \beta+1 \end{matrix}\,\right|\, \frac{1+x}{2}\right). \end{aligned}$$

Hence

$$P_n(1) = \frac{2^n(\alpha+1)_n}{(\alpha+\beta+n+1)_n}, \quad P(-1) = \frac{(-2)^n(\beta+1)_n}{(\alpha+\beta+n+1)_n}.$$

One can then find an explicit formula for $R_n(x)$.

The above example was studied in (Koornwinder, 1984) without the use of Uvarov's theorem.

2.10 Modifications of Recursion Coefficients

We saw in §§2.7–2.9 different examples of modifying a measure of orthogonality through multiplication by a rational function or an exponential function or by adding a finite discrete part to the measure of orthogonality. The question we address here is the effect of changing the recursion coefficients on the orthogonal polynomials.

The following theorem was proved in (Wendroff, 1961), but it was probably known to Geronimus in the 1950s, see (Geronimus, 1977). It is stated without proof as a footnote in (Geronimus, 1946) where the corresponding result for orthogonal polynomials on the circle is proved.

Theorem 2.10.1 (Wendroff) *Given sequences* $\{\alpha_n : n \ge N\}$, $\{\beta_n : n \ge N\}$, $\alpha_n \in \mathbb{R}$, $\beta_n > 0$*; and two finite sequences* $x_1 > x_2 > \cdots > x_N$, $y_1 > y_2 > \cdots > y_{N-1}$ *such that* $x_{k-1} > y_{k-1} > x_k$, $1 < k \le N$, *then there is a sequence of monic orthogonal polynomials* $\{P_n(x) : n \ge 0\}$ *such that*

$$P_N(x) = \prod_{j=1}^{N}(x - x_j), \quad P_{N-1}(x) = \prod_{j=1}^{N-1}(x-y_j), \tag{2.10.1}$$

and

$$xP_n(x) = P_{n+1}(x) + \tilde{\alpha}_n P_n(x) + \tilde{\beta}_n P_{n-1}(x), \quad n > 0, \tag{2.10.2}$$

and $\tilde{\alpha}_n = \alpha_n$, $\tilde{\beta}_n = \beta_n$, *for* $n \ge N$.

Proof Use (2.10.2) for $n \ge N$ to define $P_{N+j}(x)$, $j > 0$, hence P_n has precise degree n for $n > N$. Define $\tilde{\alpha}_{N-1}$ by demanding that φ_{N-2},

$$\varphi_{N-2}(x) := (x - \tilde{\alpha}_{N-1})\,P_{N-1}(x) - P_N(x),$$

has degree at most $N-2$. Clearly $\operatorname{sgn}\varphi_{N-2}(y_j) = -\operatorname{sgn}P_N(y_j) = (-1)^{j-1}$, hence $\varphi_{N-2}(x)$ has at least $N-1$ sign changes, so it must have degree $N-2$ and its zeros interlace with the zeros of $P_{N-1}(x)$. Choose $\tilde{\beta}_{N-1}$ so that $\varphi_{N-2}(x)/\tilde{\beta}_{N-1}$ is monic. Hence $\tilde{\beta}_{N-1} > 0$. By continuing this process we generate all the remaining polynomials and the orthogonality follows from the spectral theorem. □

Remark 2.10.1 *It is clear that Theorem* 2.10.1 *can be stated in terms of eigenvalues of tridiagonal matrices instead of zeros of P_N and P_{N-1}, see* (2.2.7). *This was the subject of (Drew et al., 2000), (Gray & Wilson, 1976), (Elsner & Hershkowitz, 2003) and (Elsner et al., 2003), whose authors were not aware of Wendroff's theorem or the connection between tridiagonal matrices and orthogonal polynomials.*

Now start with a recursion relation of the form (2.3.1) and define R_N and R_{N-1} to be monic of exact degrees N, $N-1$, respectively, and have real simple and interlacing zeros. Define $\{R_n : n > N\}$ through (2.3.1) and use Theorem 2.10.1 to generate $R_{N-2}, \ldots, R_0$ $(= 1)$. If $\{P_n(x)\}$ and $\{P_n^*(x)\}$ be as in §2.3. If the continued J-fraction corresponding to (2.3.1) converges then the continued J-fraction of $\{R_n(x)\}$ converges, and we can relate the two continued fractions because their entries differ in at most finitely different places.

It must be noted that the process of changing finitely many entries in a Jacobi matrix corresponds to finite rank perturbations in operator theory.

In general, we can define the kth associated polynomials $\{P_n(x;k)\}$ by

$$P_0(x;k) = 1, \quad P_1(x;k) = x - \alpha_k, \tag{2.10.3}$$

$$xP_n(x;k) = P_{n+1}(x;k) + \alpha_{n+k}P_n(x;k) + \beta_{n+k}P_{n-1}(x;k). \tag{2.10.4}$$

Hence, $P_{n-1}(x;1) = P_n^*(x)$. It is clear that

$$(P_n(x;k))^* = P_{n-1}(x;k+1). \tag{2.10.5}$$

When $\{P_n(x)\}$ is orthogonal with respect to a unique measure μ then the corresponding continued J-fraction, $F(x)$, say, satisfies

$$\lim_{n\to\infty} \frac{P_n^*(x)}{P_n(x)} = F(x) = \int_{\mathbb{R}} \frac{d\mu(t)}{x-t}. \tag{2.10.6}$$

The continued J-fraction $F(x;k)$ of $\{P_n(x;k)\}$ is

$$\frac{1}{x-\alpha_k-}\,\frac{\beta_{k+1}}{x-\alpha_{k+1}-}\cdots.$$

Since

$$F(x) = \frac{1}{x-\alpha_0-}\,\frac{\beta_1}{x-\alpha_1-}\cdots\frac{\beta_k}{x-\alpha_k-}\cdots,$$

then

$$F(x) = \frac{1}{x-\alpha_0-}\,\frac{\beta_1}{x-\alpha_1-}\cdots\frac{\beta_k}{x-\alpha_{k-1}-\beta_k F(x;k)}. \tag{2.10.7}$$

Thus (2.10.7) evaluates $F(x;k)$, from which we can recover the spectral measure of $\{P_n(x;k)\}$.

An interesting problem is to consider $\{P_n(x;k)\}$ when k is not an integer. This is meaningful when the coefficients in (2.10.3) and (2.10.4) are well-defined. The problem of finding $\mu(x;k)$, the measure of orthogonality of $\{P_n(x;k)\}$, becomes highly nontrivial when $0 < k < 1$. If this measure is found, then the Stieltjes transforms of $\mu(x;k)$ for $k > 1$ can be found from the corresponding continued fraction. The parameter k is called the association parameter.

The interested reader may consult the survey article (Rahman, 2001) for a detailed account of the recent developments and a complete bibliography.

The measures and explicit forms of two families of associated Jacobi polynomials are in (Wimp, 1987) and (Ismail & Masson, 1991). The associated Laguerre polynomials are in (Askey & Wimp, 1984). Two families of associated Laguerre and Meixner polynomials are in (Ismail et al., 1988). The most general associated classical orthogonal polynomials are the Ismail–Rahman polynomials which arise as associated polynomials of the Askey–Wilson polynomial system. See (Ismail & Rahman, 1991) for details. The weight functions for a general class of polynomials orthogonal on $[-1,1]$ containing the associated Jacobi polynomials have been studied by Pollaczek in his memoir (Pollaczek, 1956). Pollaczek's techniques have been instrumental in finding orthogonality measures for polynomials defined by three-term recurrence relations, as we shall see in Chapter 5.

In §5.6 we shall treat the associated Laguerre and Hermite polynomials, and §5.7 contains a brief account of the two families of associated Jacobi polynomials. The Ismail–Rahman work will be mentioned in §15.10.

2.11 Dual Systems

A discrete orthogonality relation of a system of polynomials induces an orthogonality relation for the dual system where the role of the variable and the degree are interchanged. The next theorem states this in a precise fashion.

Theorem 2.11.1 (Dual orthogonality) *Assume that the coefficients $\{\beta_n\}$ in (2.2.1) are bounded and that the moment problem has a unique solution. If the orthogonality measure μ has isolated point masses at α and β (β may $= \alpha$) then the dual orthogonality relation*

$$\sum_{n=0}^{\infty} P_n(\alpha)P_n(\beta)/\zeta_n = \frac{1}{\mu\{\alpha\}}\,\delta_{\alpha,\beta}, \tag{2.11.1}$$

holds.

Proof The case $\alpha = \beta$ is Theorem 2.5.6. If $\alpha \neq \beta$, then the Christoffel–Darboux formula yields

$$\sum_{n=0}^{\infty} P_n(\alpha)P_n(\beta)/\zeta_n = \lim_{N\to\infty} \frac{P_{N+1}(\alpha)P_N(\beta) - P_{N+1}(\beta)P_N(\alpha)}{\zeta_N\,(\alpha-\beta)}$$

which implies the result since $\lim\limits_{N\to\infty} P_N(x)/\sqrt{\zeta_N} = 0$, $x = \alpha, \beta$. ☐

My friend Christian Berg sent me the following related theorem and its proof.

Theorem 2.11.2 *Let $\mu = \sum_{\lambda\in\Lambda} a_\lambda \epsilon_\lambda$ be a discrete probability measure and assume that $\{p_n(x)\}$ is a polynomial system orthonormal with respect to μ and complete in $L^2(\mu;\mathbb{R})$. Then the dual system $\left\{p_n(\lambda)\sqrt{a_\lambda} : \lambda \in \Lambda\right\}$ is a dual orthonormal basis for ℓ^2, that is*

$$\sum_{n=0}^{\infty} p_n\left(\lambda_1\right) p_n\left(\lambda_2\right) = \delta_{\lambda_1,\lambda_2}/a_{\lambda_1}, \forall \lambda_1, \lambda_2 \in \Lambda.$$

The proof will appear elsewhere.

Dual orthogonality arises in a natural way when our system of polynomials has finitely many members. This happens if β_n in (2.2.1) is positive for $0 \le n < N$ but $\beta_{N+1} \le 0$. Examples of such systems will be given in §6.2 and §15.6.

Let $\{\phi_n(x) : 0 \le n \le N\}$ be a finite system of orthogonal polynomials satisfying

$$\sum_{x=0}^{N} \phi_m(x)\phi_n(x)w(x) = \delta_{m,n}/h_n. \tag{2.11.2}$$

Then

$$\sum_{n=0}^{N} \phi_n(x)\phi_n(y)h_n = \delta_{x,y}/w(x) \tag{2.11.3}$$

holds for $x, y = 0, 1, \ldots, N$.

de Boor and Saff introduced another concept of duality which we shall refer to as the deB–S duality (de Boor & Saff, 1986). Given a sequence of polynomials satisfying (2.2.1)–(2.2.2), define a deB–S dual system $\{Q_n(x) : 0 \le n \le N\}$ by

$$Q_0(x) = 1, \quad Q_1(x) = x - \alpha_{N-1}, \tag{2.11.4}$$

$$Q_{n+1}(x) = \left(x - \alpha_{N-n-1}\right) Q_n(x) - \beta_{N-n} Q_{n-1}(x), \tag{2.11.5}$$

for $0 \le n < N$. From (2.2.1)–(2.2.2) and (2.11.4)–(2.11.5) it follows that

$$Q_n(x) = P_n(x; N - n).$$

The material below is from (Vinet & Zhedanov, 2004).

Clearly the mapping $\{P_n\} \to \{Q_n\}$ is an involution in the sense that the deB–S dual of $\{Q_n\}$ is $\{P_n\}$. By induction we can show that

$$Q_{N-n-1}(x)P_{n+1}(x) - \beta_n Q_{N-n-2}(x)P_n(x) = P_N(x), \tag{2.11.6}$$

$0 \le n < N$. As in the proof of Theorem 2.6.2, we apply (2.5.1) to see that $\{P_n(x) : 0 \le n < N\}$ are orthogonal on $\{x_{N,j} : 1 \le j \le N\}$ with respect to a discrete measure with masses $\rho\left(x_{N,j}\right)$ at $x_{N,j}$, $1 \le j \le N$. Moreover

$$\frac{P_N^*(x)}{P_N(x)} = \sum_{k=1}^{N} \frac{\rho\left(x_{N,j}\right)}{x - x_{N,j}}.$$

Using the fact that $P_N(x) = Q_N(x)$ and $Q_N^*(x) = P_{N-1}(x)$ it follows that the masses $\rho_Q\left(x_{N,j}\right)$ defined by

$$\frac{P_{N-1}(x)}{P_N(x)} = \frac{Q_N^*(x)}{Q_N(x)} = \sum \frac{\rho_Q\left(x_{N,j}\right)}{x - x_{N,j}},$$

so that the numbers $\rho_Q(x_{N,j}) = P_{N-1}(x_{N,j})/P'_N(x_{N,j})$, have the property

$$\sum_{j=1}^{N} \rho_Q(x_{N,j}) = 1, \quad \sum_{j=1}^{N} \rho_Q(x_{N,j}) Q_m(x_{N,j}) Q_n(x_{N,j}) = \zeta_n(Q)\delta_{m,n}, \tag{2.11.7}$$

and

$$\zeta_n(Q) = \beta_{N-1}\beta_{N-2}\cdots\beta_{N-n}, \quad \text{or} \quad \zeta_n(Q) = \zeta_{N-1}/\zeta_{N-n-1}. \tag{2.11.8}$$

Vinet and Zhedanov refer to $\rho_Q(x_{N,j})$ as the dual weights.

Theorem 2.11.3 *The Jacobi, Hermite and Laguerre polynomials are the only orthogonal polynomials where $\rho_Q(x_{N,j}) = \pi(x_{N,j})/c_N$ for all N, where π is a polynomial of degree at most 2 and c_N is a constant.*

For a proof, see (Vinet & Zhedanov, 2004).

A sequence of orthogonal polynomials is called **semiclassical** if it is orthogonal with respect to an absolutely continuous measure μ, where $\mu'(x)$ satisfies a differential equation $y'(x) = r(x)\,y(x)$, and r is a rational function.

Theorem 2.11.4 ((Vinet & Zhedanov, 2004)) *Assume that $\{Q_n(x)\}$ is the deB–S dual to $\{P_n(x)\}$ and let $\{x_{n,i}\}$ be the zeros of $P_n(x)$. Then $\{P_n(x)\}$ are semiclassical if and only if, for every $n > 0$, the polynomials $\{Q_i(x) : 0 \le i < n\}$ are orthogonal on $\{x_{n,i} : 1 \le i \le n\}$ with respect to the weights*

$$\frac{q(x_{n,i})}{\tau(x_{n,i}, n)},$$

where $q(x)$ is a polynomial of fixed degree and its coefficients do not depend on n, but $\tau(x,n)$ is a polynomial of fixed degree whose coefficients may depend on n.

Exercises

2.1 Let $\{P_n(x)\}$ satisfy (2.2.1)–(2.2.2). Prove that the coefficient of x^{n-2} in $P_n(x)$ is

$$\sum_{0\le i<j<n-1} \alpha_i\alpha_j - \sum_{k=1}^{n-1} \beta_k.$$

2.2 Assume that $\alpha_n = 0$, $n \ge 0$, $\beta_n < 0$ for $n > 0$. If $\{P_n(x)\}$ satisfies (2.2.1)–(2.2.2), prove that $\{i^{-n}P_n(ix)\}$ is a real sequence of orthogonal polynomials.

2.3 Consider an absolutely continuous measure μ on $\mathbb{R}$ with $\mu'(x) = e^{-x^2}dx/\sqrt{\pi}$. Prove that $\mu_{2n+1} = 0$ and $\mu_{2n} = (1/2)_n$. Evaluate the Hankel determinants D_n, $n > 0$ and find an explicit formula for the orthonormal polynomials.

2.4 Repeat Exercise 2.3 for μ supported on $[0,\infty)$ with

$$\mu'(x) = x^\alpha e^{-x}dx/\Gamma(\alpha+1).$$

2.5 Assume that $\{P_n(x)\}$ satisfies (2.2.1)–(2.2.2) and $Q_n(x) = a^{-n} P_n(ax + b)$. Show that if $\{P_n(x)\}$ are orthogonal with respect to a probability measure with moments $\{\mu_n\}$ then $\{Q_n(x)\}$ are orthogonal with respect to a probability measure whose moments m_n are given by

$$m_n = a^{-n} \sum_{k=0}^{n} \binom{n}{k} (-b)^{n-k} \mu_k.$$

2.6 Suppose that $P_n(x)$ satisfies the three-term recurrence relation

$$P_{n+1}(x) = (x - \alpha_n)\, P_n(x) - \beta_n P_{n-1}(x), \quad n > 0,$$
$$P_0(x) = 1, \quad P_{-1}(x) = 0.$$

Assume that $\beta_n > 0$, and that L is the positive definite linear functional for which P_n is orthogonal and $L(1) = 1$, that is $L\left(p_m(x) p_n(x)\right) = 0$ if $m \neq n$. Suppose that S is another linear functional such that $S\left(P_k P_\ell\right) = 0$ for $n \geq k > \ell \geq 0$, $S(1) = 1$. Show that S has the same jth moments as L, for j at most $2n - 1$.

2.7 Explicitly renormalize the polynomials $P_n(x)$ to $P_n^*(x)$ in Exercise 2.6 so that the eigenvalues and eigenvectors of the related real symmetric tridiagonal n by n matrix are explicitly given by the zeros $x_{n,k}$ of $P_n(x)$ and the vectors $\left(P_0^*\left(x_{n,k}\right), \ldots, P_{n-1}^*\left(x_{n,k}\right)\right)$. Find these results. Next, using orthogonality of the eigenvectors, rederive the Gaussian quadrature formula

$$L_n(p) = \sum_{k=1}^{n} \Lambda_{nk}\, p\left(x_{nk}\right).$$

Then use Exercise 2.6 to conclude that L_n and L have the identical moments up to $2n - 1$.

2.8 Let $\Delta(x) = \prod\limits_{k>j} \left(x_k - x_j\right)$ be the Vandermonde determinant in n variables $x_1, \ldots, x_n$. Let $w(x) = x^a (1-x)^b$ on $[-1, 1]$, where $a, b > -1$. Evaluate

(a) $\int\limits_{[0,1]^n} [\Delta(x)]^2 w\left(x_1\right) w\left(x_2\right) \ldots w\left(x_n\right) dx_1 \ldots dx_n$

(b) $\int\limits_{[0,1]^n} [\Delta(x)]^2 w\left(x_1\right) w\left(x_2\right) \ldots w\left(x_n\right) e_k\left(x_1, \ldots, x_n\right) dx_1 \ldots dx_n$, where e_k is the elementary symmetric function of degree k.

2.9 Let $\{r_n(x)\}$ and $\{s_n(x)\}$ be orthonormal with respect to $\rho(x)\, dx$ and $\sigma(x)\, dx$, respectively, and assume that $\int\limits_{\mathbb{R}} \rho(x)\, dx = \int\limits_{\mathbb{R}} w(x)\, dx = 1$. Prove that if

$$r_n(x) = \sum_{k=0}^{n} c_{n,k}\, s_k(x),$$

then

$$\sigma(x) s_k(x) = \sum_{n=k}^{\infty} c_{n,k}\, \rho(x) r_n(x).$$

The second equality holds in $L^2(\mathbb{R})$.

2.10 Assume that $y \in \mathbb{R}$ and does not belong to the true interval of orthogonality of $\{P_n(x)\}$. Prove that the kernel polynomials $\{K_n(x, y)\}$ are orthogonal with respect to $|x - y|\, d\mu(x)$, μ being the orthogonality measure of $\{P_n(x)\}$.

3

Differential Equations, Discriminants and Electrostatics

In this chapter we derive linear second order differential equations satisfied by general orthogonal polynomials with absolutely continuous measures of orthogonality. We also give a closed form expression of the discriminant of such orthogonal polynomial in terms of the recursion coefficients. These results are then applied to electrostatic models of a system of interacting charged particles.

3.1 Preliminaries

Assume that μ is absolutely continuous and let

$$d\mu(x) = w(x)\,dx, \quad x \in (a,b), \tag{3.1.1}$$

where $[a,b]$ is not necessarily bounded. We shall write

$$w(x) = \exp(-v(x)), \tag{3.1.2}$$

require v to be twice differentiable and assume that the integrals

$$\int_a^b y^n \frac{v'(x)-v'(y)}{x-y}\, w(y)\,dy, \quad n = 0,1,\dots, \tag{3.1.3}$$

exist. We shall also use the orthonormal form $\{p_n(x)\}$ of the polynomials $\{P_n(x)\}$, that is

$$p_n(x) = P_n(x)/\sqrt{\zeta_n}. \tag{3.1.4}$$

Rewrite (2.2.1) and (2.2.2) in terms of the p_ns. The result is

$$p_0(x) = 1, \quad p_1(x) = (x-\alpha_0)/a_1, \tag{3.1.5}$$

$$xp_n(x) = a_{n+1}p_{n+1}(x) + \alpha_n p_n(x) + a_n p_{n-1}(x), \quad n > 0. \tag{3.1.6}$$

The discriminant D of a polynomial g,

$$g(x) := \gamma \prod_{j=1}^{n} (x - x_j) \tag{3.1.7}$$

is defined by, (Dickson, 1939),

$$D(g) := \gamma^{2n-2} \prod_{1 \le j < k \le n} (x_j - x_k)^2 . \tag{3.1.8}$$

An alternate representation for the discriminant is

$$D(g) := (-1)^{n(n-1)/2} \gamma^{n-2} \prod_{j=1}^{n} g'(x_j), \tag{3.1.9}$$

as can be seen by direct substitution of g in (3.1.9) and using (3.1.7), see (Dickson, 1939). The resultant of g and a polynomial f of degree m is

$$\operatorname{Res}\{g, f\} = \gamma^m \prod_{j=1}^{n} f(x_i). \tag{3.1.10}$$

Clearly

$$D(g) = (-1)^{\binom{n}{2}} \gamma^{-1} \operatorname{Res}\{g, g'\}. \tag{3.1.11}$$

We shall consider the case of logarithmic potential on the line. In this model the potential energy at x of a point charge e located at c is $-e \ln |x - c|$. Consider the system of n movable unit charged particles in the presence of the external potential $V(x)$. The particles are restricted to lie in $[a, b]$. Let

$$\mathbf{x} := (x_1, x_2, \dots, x_n), \tag{3.1.12}$$

where $x_1, \dots, x_n$ are the positions of the particles arranged in decreasing order. The total energy of the system is

$$E(\mathbf{x}) = \sum_{k=1}^{n} V(x_k) - 2 \sum_{1 \le j < k \le n} \ln |x_j - x_k| . \tag{3.1.13}$$

Let

$$T(\mathbf{x}) := \exp(-E(\mathbf{x})). \tag{3.1.14}$$

In §3.5 we shall describe how the zeros of p_n describe the equilibrium position of the particles in such a system.

3.2 Differential Equations

Define $A_n(x)$ and $B_n(x)$ via

$$\frac{A_n(x)}{a_n} = \left. \frac{w(y)\, p_n^2(y)}{y - x} \right|_a^b + \int_a^b \frac{v'(x) - v'(y)}{x - y}\, p_n^2(y)\, w(y)\, dy, \tag{3.2.1}$$

$$\begin{aligned} \frac{B_n(x)}{a_n} &= \left. \frac{w(y)\, p_n(y)\, p_{n-1}(y)}{y - x} \right|_a^b \\ &\quad + \int_a^b \frac{v'(x) - v'(y)}{x - y}\, p_n(y) p_{n-1}(y)\, w(y)\, dy. \end{aligned} \tag{3.2.2}$$