Tutorials, Schools, and Workshops in the
Mathematical Sciences

Mama Foupouagnigni
Wolfram Koepf
Editors

Orthogonal Polynomials

2nd AIMS-Volkswagen Stiftung Workshop, Douala, Cameroon, 5-12 October, 2018

Birkhäuser

Tutorials, Schools, and Workshops in the Mathematical Sciences

Birkhäuser

This series will serve as a resource for the publication of results and developments presented at summer or winter schools, workshops, tutorials, and seminars. Written in an informal and accessible style, they present important and emerging topics in scientific research for PhD students and researchers. Filling a gap between traditional lecture notes, proceedings, and standard textbooks, the titles included in TSWMS present material from the forefront of research.

More information about this series at http://www.springer.com/series/15641

Mama Foupouagnigni • Wolfram Koepf
Editors

Orthogonal Polynomials

2nd AIMS-Volkswagen Stiftung Workshop, Douala, Cameroon, 5-12 October, 2018

Birkhäuser

Editors
Mama Foupouagnigni
University of Yaoundé I
Yaoundé, Cameroon

African Institute for Mathematical Sciences
Limbe, Cameroon

Wolfram Koepf
Institute for Mathematics
University of Kassel
Kassel, Germany

ISSN 2522-0969 ISSN 2522-0977 (electronic)
Tutorials, Schools, and Workshops in the Mathematical Sciences
ISBN 978-3-030-36743-5 ISBN 978-3-030-36744-2 (eBook)
https://doi.org/10.1007/978-3-030-36744-2

Mathematics Subject Classification (2010): 33C, 33D, 33F

This book is published under the imprint Birkhäuser, www.birkhauser-science.com, by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

In July 2015, following an invitation from the African Institute for Mathematical Sciences (AIMS) Global Secretariat, the two authors of this foreword decided to apply for funding to the *Volkswagen Foundation's Symposia and Summer Schools* initiative to support the organization of two workshops planned to take place in Cameroon.

AIMS is a pan-African network of centres of excellence for post-graduate training, research and public engagement in mathematical sciences. AIMS enables Africa's brightest students to become innovators that propel scientific, educational and economic self-sufficiency.

AIMS-Cameroon is the fourth Center of Excellence of the AIMS Network created in 2013 after AIMS South Africa, AIMS Senegal and AIMS Ghana, followed by the creation of AIMS Tanzania and AIMS Rwanda. It is located in Limbe in Cameroon, a country from the Central Africa sub-region, well known as Africa in miniature due to its diverse landscapes that represent the continent's major climatic zones. Among these, we can mention that white and black beaches, mountainous areas, tropical rain forests, savanna grasslands and sparse deserts are found in this country. As illustration, the Mount Cameroon, located in Buea near Limbe, is the highest point in central and west Africa while Debundscha also located near Limbe is the sixth wettest place in the world.

Of course, it took quite a long time to develop this idea and make it concrete: We had to propose programmes for the two planned meetings on *Introduction to Computer Algebra and Applications* and on *Introduction to Orthogonal Polynomials and Applications*. We had also to invite renowned international experts to present plenary lectures on the current state of the art in their domains, and we were very happy about their acceptance. We had to write the corresponding proposal and to get approval from the reviewers and finally from Volkswagen Foundation. In the first round, some reviewers of our proposal had recommended us to add some additional actual major research fields in orthogonal polynomials which we did.

In March 2017, we fortunately received the positive grant letter from Volkswagen Foundation and could start the final planning. The first workshop on *Introduction to Computer Algebra and Applications* took place on October 6–13, 2017, whereas the second workshop on *Introduction to Orthogonal Polynomials and Applications* was scheduled for October 5–12, 2018, both taking place in the Hotel Prince de Galles in Douala, the economic capital of Cameroon. They were hosted by the African Institute for Mathematical Sciences, Cameroon (https://www.aims-cameroon.org/), and were co-organized together with the University of Kassel in Germany (http://www.mathematik.uni-kassel.de/~koepf/). All details about the two workshops can be found on the web domain http://www.aims-volkswagen-workshops.org/ of the meetings which is still active.

These Proceedings contain the results of the second workshop. This workshop on *Introduction to Orthogonal Polynomials and Applications* was aimed globally at promoting capacity building in terms of research and training in orthogonal polynomials and applications, discussions and development of new ideas, development and enhancement of networking including south–south cooperation.

The workshop brought together 60 participants from 18 African, 7 European and 2 North American countries including 19 plenary speakers and trainers who are all experts in their various domains. In total, about 50 plenary talks, tutorials, training sessions and contributed talks were delivered during the preliminary workshop (October 5–7, 2018) and the workshop (October 8–12, 2018).

To announce the workshop, we designed a nice and informative web page at http://www.aims-volkswagen-workshops.org/ and a link to the first workshop web site was included. Also, a Facebook event page was installed. On this web page, information about the objectives and expectations of the workshop, the organizers and the funding partners, the plenary speakers (CV and photo), programme schedule and abstracts of lectures and tutorials could be downloaded, also by those interested researchers that we could not invite. Also, we have put on the web site and spread by email all over Africa the call for application. Those interested to apply had to fill in the online application form from our web site. As a result, we have received 130 applications within about 1 month. Many of these applications were excellent so that we could finally invite 25 Africans from outside Cameroon from 18 different countries and 16 Cameroonian researchers. 10 of our African participants were female. The 6 African participants with the best proposals for a talk were selected to present their research at the workshop.

The workshop evaluation by the participants was very positive, and the workshop which enabled active interactions between the participants was a great success, enabling the achievement of the stated objectives!

Since we did not expect prior knowledge about the workshop topic by the African participants, the preliminary workshop was giving a formal introduction to the field of orthogonal polynomials. This was possible through the great help of a group of five former Cameroonian PhD students who all wrote their dissertations—supervised by the two workshop organizers—in the field of Orthogonal Polynomials and Special Functions and whom we are very grateful for their help: Maurice Kenfack Nangho, Salifou Mboutngam, Merlin Mouafo Wouodjie, Patrick Njionou Sadjang and Daniel D. Tcheutia. Finally, the talks given by the two organizers and the one of Aletta Jooste from the University of Pretoria in South Africa (who could not attend due to last minute health constraints) given by Daniel D. Tcheutia, complemented the lecturers of the preliminary workshop. All those contributions are contained in these Proceedings.

The main workshop was aimed at introducing concepts of modern and actual research topics in orthogonal polynomials: Multiple Orthogonal Polynomials, Orthogonal Polynomials and Painlevé Equations, Orthogonal Polynomials and Random Matrices, Orthogonal Polynomials in Sobolev Spaces, Matrix Polynomials, Zeros of Orthogonal Polynomials, Computer Algebra and Orthogonal Polynomials, Multivariate Orthogonal Polynomials, Askey–Wilson Scheme.

Based on the concept of the workshop, we have divided these Proceedings into two parts. Part I gives an *Introduction to Orthogonal Polynomials* based on the talks given in the preliminary workshop. They are organized in their logical structure. Part II presents the remaining lectures on *Actual Research Topics in Orthogonal Polynomials and Applications*. They are organized in alphabetical order of their authors. An ordering by their topics was not possible since some authors preferred to write one article about several topics.

We hope that these Proceedings will not only give a very good introduction into the state of the actual research in orthogonal polynomials and applications, but also help those interested in orthogonal polynomials without prior knowledge to embark into this interesting field of research.

Acknowledgements

We would like to thank Volkswagen Foundation for their great support! Without this generous grant (Symposia and Summer Schools, Wissen für Morgen, AZ 93 000, http://portal.volkswagenstiftung.de/search/projectDetails.do?ref=93000) this workshop clearly would not have been possible. Furthermore, we would like to thank Clemens Heine from Birkhäuser who had the idea to collect this volume after visiting our web site. Our thanks go also to the plenary speakers, the trainers and Ms Nathalie Diane Wandji Nanda for their decisive contributions to the success of this workshop. We are also delighted to acknowledge and value the tremendous efforts of the AIMS Global Network in building the capacity of the next generation of African scientists and innovators. We would also like to thank the Alexander von Humboldt Foundation for continuously supporting the academic cooperation between the two authors. This collaboration has made possible the organization of the two Volkswagen workshops from where the current proceedings emerged.

In the same line, Mama Foupouagnigni is very pleased to acknowledge here with big thanks the valuable and decisive contribution of Wolfram Koepf to the capacity building of the Cameroon mathematical community by means of supervision, co-supervision or mentoring, which has led to the successful completion of at least 7 PhD theses in mathematics (Mama Foupouagnigni, Etienne Nana Chiadjeu and the above listed five former Cameroonian PhD students, Bertrand Teguia Tabuguia (PhD thesis almost completed)); and at least three German Habilitation theses in mathematics (Mama Foupouagnigni, Jean Sire Eyebe Fouda and Daniel Duviol

Tcheutia, Patrick Sadjang Njionou (Habilitation thesis almost completed)). His wife Angelika Wolf is also acknowledged for having continuously supported this academic cooperation which first started in 1997 as student–supervisor relationship, then migrated to collaboration among two Humboldtians. It has now developed into an active institutional academic cooperation combined with sustained friendship and extended family ties, connecting Cameroon and Germany. Such academic cooperation and cultural dialogue are in line with the objectives of the Alexander von Humboldt Foundation which is to promote academic cooperation between excellent scientists and scholars from abroad and from Germany, as an intermediary organization for German foreign cultural and educational policy promoting international cultural dialogue and academic exchange.

Limbe, Cameroon
Kassel, Germany
August 2019

Mama Foupouagnigni
Wolfram Koepf

Contents

Part I
Introduction to Orthogonal Polynomials

An Introduction to Orthogonal Polynomials

Mama Foupouagnigni

Abstract In this introductory talk, we first revisit with proof for illustration purposes some basic properties of a specific system of orthogonal polynomials, namely the Chebyshev polynomials of the first kind. Then we define the notion of orthogonal polynomials and provide with proof some basic properties such as: The uniqueness of a family of orthogonal polynomials with respect to a weight (up to a multiplicative factor), the matrix representation, the three-term recurrence relation, the Christoffel-Darboux formula and some of its consequences such as the separation of zeros and the Gauss quadrature rules.

Keywords Orthogonal polynomials · Differential equations · Chebyshev polynomials · Zeros

Mathematics Subject Classification (2000) Primary 33C45; Secondary 42C05

The research of the author was partially supported by the AIMS-Cameroon Research Allowance 2018–2019.

M. Foupouagnigni (✉)
Department of Mathematics, Higher Teachers' Training College, University of Yaounde I, Yaounde, Cameroon

The African Institute for Mathematical Sciences, Limbe, Cameroon
e-mail: mfoupouagnigni@aims-cameroon.org

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_1

1 Introduction: An Example of a Family of Orthogonal Polynomials

Univariate orthogonal polynomials (or orthogonal polynomials for short) are systems of polynomials $(p_n)_n$ with $\deg(p_n) = n$, satisfying a certain orthogonality relation. They are very useful in practice in various domains of mathematics, physics, engineering and so on, because of the many properties and relations they satisfy. As examples of areas where orthogonal polynomials play important roles, I could list approximation theory (see [5, 23]) and also numerical analysis (see for example [9, 10]). Among those relations, we can mention the following, with the first seven valid for all families of orthogonal polynomials. The last three are in general valid for some specific families of orthogonal polynomials, the so-called classical orthogonal polynomials (see [1–3, 6, 7, 12, 14]) and the preliminary training given by S. Mboutngam, M. Kenfack Nangho and P. Njionou Sadjang of these proceedings):

- Orthogonality relation
- Matrix representation
- Three-term recurrence relation
- Christoffel-Darboux formula
- Separation of zeros
- Gauss quadrature
- Generating functions
- Second-order holonomic differential, difference or q-difference equation
- Rodrigues formula
- Expansion of functions which are continuous differentiable and square integrable, in terms of Fourier series of OP.

Before going into details and for illustration purposes, let us give a concrete example of a family of orthogonal polynomials, then state and prove some of its properties most of which are common to any family of orthogonal polynomials.

Theorem 1.1 (Chebyshev Polynomials of the First Kind [17, 21]) *The polynomial family $(T_n)_n$ defined by (and called Chebyshev polynomials of the first kind or Chebyshev polynomials for short as we will study only this family in this article)*

$$T_n(x) = \cos(n\theta),\ x = \cos\theta,\ 0 < \theta < \pi,\ n \in \mathbb{N}, \tag{1.1}$$

fulfills the following properties:

1. *T_n is a polynomial of degree n in x with leading coefficient $a_n = 2^{n-1}$, satisfying the following recurrence relation (called three-term recurrence relation)*

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x),\ n \geq 1,\ T_0(x) = 1,\ T_1(x) = x; \tag{1.2}$$

2. *$(T_n)_n$ satisfies the following relation (called orthogonality relation)*

$$\int_0^{\pi} \cos(n\theta)\cos(m\theta)d\theta = k_n\delta_{n,m} = \int_{-1}^{1} T_n(x)\, T_m(x)\frac{dx}{\sqrt{1-x^2}}, \tag{1.3}$$

with $k_0 = \pi,\ k_n = \frac{\pi}{2},\ n \geq 1.$

3. *T_n satisfies the second-order holonomic differential equation:*

$$(1-x^2)\, T_n''(x) - x\, T_n'(x) + n^2\, T_n(x) = 0,\ n \geq 0. \tag{1.4}$$

4. *For any $n \geq 1$, T_n has exactly n zeros, all belonging to the interval of orthogonality* $(-1,\ 1)$. *Those zeros, ranked in increasing order, are given by:*

$$x_{n,k} = \cos\left(\frac{2(n-k)+1}{2n}\pi\right),\ 1 \leq k \leq n,\ n \geq 1. \tag{1.5}$$

5. *The zeros $x_{n,k}$ of T_n satisfy*

$$x_{n,j} \neq x_{n+1,k},\ \forall n \geq 1,\ 1 \leq j \leq n,\ 1 \leq k \leq n+1; \tag{1.6}$$

$$x_{n+1,k} < x_{n,k} < x_{n+1,k+1},\ 1 \leq k \leq n. \tag{1.7}$$

6. *The monic Chebyshev polynomial of degree $n \geq 1$ is the polynomial deviating least from zero on* $[-1, 1]$ *among all monic polynomials of degree n:*

$$\min\left\{\max_{-1\leq x\leq 1} |q_n(x)|,\ q_n \in \mathbb{R}[x], q_n(x) = x^n + \ldots\right\} = \max_{-1\leq x\leq 1}\left|\frac{T_n(x)}{2^{n-1}}\right| = \frac{1}{2^{n-1}}, \tag{1.8}$$

where $\mathbb{R}[x]$ *is the ring of polynomials with real coefficients.*

7. *The following property, which is called the Gauss quadrature formula for the specific case of the Chebyshev polynomials, is valid*

$$\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}dx = \frac{\pi}{n}\sum_{k=1}^{n} f(x_{n,k}),\ \forall f \in \mathbb{R}_{2n-1}[x],\ n \geq 1, \tag{1.9}$$

$\mathbb{R}_{2n-1}[x]$ *is the ring of polynomials of degree at most $2n-1$, with real coefficients. In addition, the integral of any function continuous on the compact interval* $[-1,\ 1]$ *can be approximated by the previous formula:*

$$\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}dx = \lim_{n\to\infty}\frac{\pi}{n}\sum_{k=1}^{n} f(x_{n,k}),\ \forall f \in \mathcal{C}[-1,\ 1], \tag{1.10}$$

where $\mathcal{C}[-1,\ 1]$ *is the set of continuous functions on the interval* $[-1,\ 1]$.

Proof Let us provide a quick proof of the first six above properties.

Proof of Property 1 Equation (1.2) is obtained by direct computation:

$$T_0(x) = \cos(0) = 1, \; T_1(x) = \cos\theta = x,$$

and

$$T_{n+1}(x) + T_{n-1}(x) = \cos(n+1)\theta + \cos(n-1)\theta = 2\,\cos\theta\,\cos(n\theta) = 2\,x\,T_n(x),$$

using the cosine addition formula $\cos(a+b) = \cos a\,\cos b - \sin a\,\sin b$.

Next, we now prove by induction that T_n is a polynomial of degree n in the variable x with 2^{n-1} as leading coefficient, that is

$$T_n(x) = 2^{n-1}\,x^n + \text{lower degree terms},\; n \geq 1. \tag{1.11}$$

For $n = 1$, Eq. (1.11) is satisfied as $T_1(x) = x = 2^{1-1}\,x$ and its degree is 1. By assuming that Eq. (1.11) is satisfied for a fixed integer $n \geq 1$, we can then write T_n as $T_n(x) = 2^{n-1}\,x^n + A_{n-1}(x)$ where A_{n-1} is a polynomial of degree at most $n-1$ in the variable x. We complete the proof by using relation (1.2) to obtain that

$$T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x) = 2x\,(2^{n-1}\,x^n + A_{n-1}(x)) - T_{n-1}(x) = 2^n\,x^{n+1} + \tilde{A}_n(x),$$

where $\tilde{A}_n$ is a polynomial of degree at most n in x. Therefore, T_n is a polynomial of degree n in the variable x with 2^{n-1} as leading coefficient.

From the three-term recurrence relation (1.2), one can generate any T_n; and in particular, the first 10 Chebyshev polynomials are given by:

$$\begin{aligned}
T_0(x) &= 1,\\
T_1(x) &= x;\\
T_2(x) &= 2\,x^2 - 1,\\
T_3(x) &= 4\,x^3 - 3\,x,\\
T_4(x) &= 8\,x^4 - 8\,x^2 + 1,\\
T_5(x) &= 16\,x^5 - 20\,x^3 + 5\,x,\\
T_6(x) &= 32\,x^6 - 48\,x^4 + 18\,x^2 - 1,\\
T_7(x) &= 64\,x^7 - 112\,x^5 + 56\,x^3 - 7\,x,\\
T_8(x) &= 128\,x^8 - 256\,x^6 + 160\,x^4 - 32\,x^2 + 1,\\
T_9(x) &= 256\,x^9 - 576\,x^7 + 432\,x^5 - 120\,x^3 + 9\,x.
\end{aligned} \tag{1.12}$$

Proof of Property 2 Relation (1.3) is proved by direct computation using again the addition formula

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

and the fact that $x = \cos\theta,\ 0 < \theta < \pi \Longrightarrow dx = -\sin\theta\, d\theta = -\sqrt{1-\cos^2\theta}d\theta.$

Proof of Property 3 Relation (1.4) is also proved by direct computation. In fact

$$\begin{aligned}
T_n'(x) &= \frac{d}{dx}T_n(x) = \frac{d\theta}{dx}\frac{d}{d\theta}T_n(x) = \frac{-1}{\sin\theta}\frac{d}{d\theta}\cos(n\theta) = \frac{n\ \sin(n\theta)}{\sin\theta},\ n \geq 1,\\
T_n''(x) &= \frac{d}{dx}\frac{d}{dx}T_n(x)\\
&= \frac{d\theta}{dx}\frac{d}{d\theta}\left(\frac{d\theta}{dx}\frac{d}{d\theta}T_n(x)\right)\\
&= \frac{-1}{\sin\theta}\frac{d}{d\theta}\left(\frac{-1}{\sin\theta}\frac{d}{d\theta}\cos(n\theta)\right)\\
&= \frac{n\cos\theta\ \sin(n\theta)}{\sin\theta\ \sin^2\theta} + \frac{-n^2\cos(n\theta)}{\sin^2\theta}\\
&= \frac{x\ T_n'(x)}{1-x^2} + \frac{-n^2\,T_n(x)}{1-x^2},\ n \geq 1.
\end{aligned}$$

Proof of Property 4 To obtain the zeros of T_n, we solve the following equation for a fixed $n \geq 1,\ x \in (-1,\ 1)$ and $\theta \in (-\pi\ \pi)$.

$$T_n(x) = 0 \Longleftrightarrow \cos(n\theta) = 0 \Longleftrightarrow n\theta = \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}.$$

Since $0 < \theta < \pi$, then $0 \leq k \leq n-1$. Therefore, T_n has exactly n zeros which are $\cos\left(\frac{(2k+1)\pi}{2n}\right),\ 0 \leq k \leq n-1$. But since those zeros are ranked by decreasing order for the function $\theta \to \cos\theta$ is decreasing on $(-\pi,\ \pi)$ and the sequence $k \to \frac{(2k+1)\pi}{2n}$ is increasing, there is a need to reverse the order. This is done by replacing k by $n-k$. Therefore we obtain the following zeros ranked by increasing order

$$x_{n,k} = \cos\theta_{n,k},\ \text{with } \theta_{n,k} = \frac{2(n-k)+1}{2n}\pi,\ 1 \leq k \leq n.$$

The zeros also belong to the interval of orthogonality $(-1,\ 1)$.

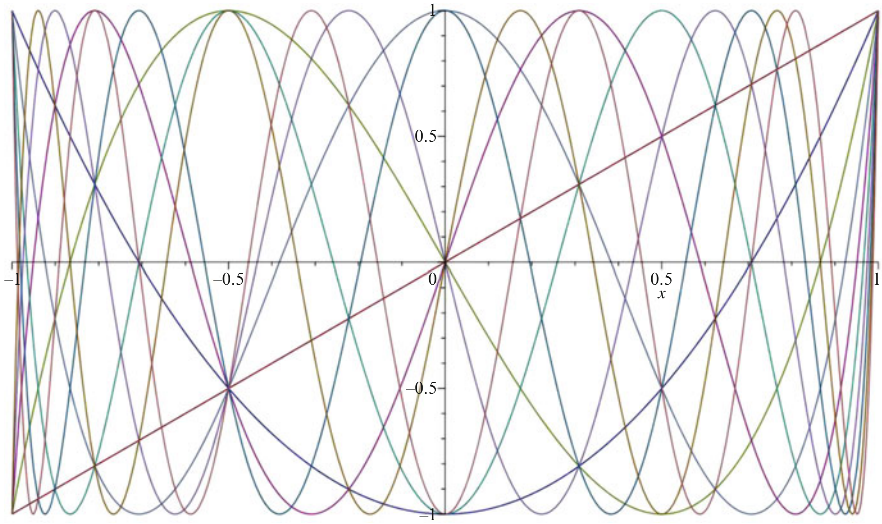


Fig. 1 The first 10 Chebyshev polynomials

Proof of Property 5 Equation (1.6) is satisfied since the cosine function is a bijection from $(0,\ \pi)$ into $(-1,\ 1)$ and

$$\theta_{n,j} \neq \theta_{n+1,k},\ \forall n \geq 1,\ \ 1 \leq j \leq n,\ \ 1 \leq k \leq n+1.$$

The inequalities (1.7) are deduced using the fact that the cosine function is strictly decreasing in $(-\pi,\ \pi)$ combined with the following inequalities which can be obtained by a direct and quick computation

$$\theta_{n+1,k+1} < \theta_{n,k} < \theta_{n+1,k},\ \ 1 \leq k \leq n.$$

The interlacing properties of the zeros of the Chebyshev polynomials can be observed on the above graph of the first ten Chebyshev polynomials (Fig. 1).

Proof of Property 6 Let us first denote the monic Chebyshev polynomial of degree n by t_n: $t_n(x) = \frac{T_n(x)}{2^{n-1}},\ n \geq 1,\ t_0(x) = T_0(x) = 1$. Next, we define the set of monic polynomials of degree n, $\mathcal{P}_n$, the sup-norm $||.||_{max}$ and the subset I of the set of real numbers $\mathbb{R}$, respectively, by

$$\begin{aligned} \mathcal{P}_n &= \left\{q_n \in \mathbb{R}_n[x],\, q_n(x) = x^n + \text{lower degree terms}\right\} \\ ||p||_{\max} &= \max_{-1 \leq x \leq 1} |p(x)|, \\ I &= \{||p||_{\max},\ p \in \mathcal{P}_n\}. \end{aligned}$$

To prove that $\min I = ||t_n||_{\max} = \frac{1}{2^{n-1}}$, for a fixed but arbitrary integer $n \geq 1$, we proceed as follows:

– In the first step, we derive the extrema for the function t_n:

$$t_n'(x) = 0 \Longleftrightarrow \frac{n\sin(n\theta)}{\sin(\theta)} = 0 \Longleftrightarrow \sin(n\theta) = 0,\ \sin\theta \neq 0.$$

Since $0 \leq \theta \leq \pi$, we get $\theta = \frac{k\pi}{n}$, $1 \leq k \leq n-1$. We have excluded $k = 0$ and $k = n$ to make sure that $\sin\theta \neq 0$. The extrema for t_n are therefore

$$z_{n,k} = \cos\left(\frac{k\pi}{n}\right),\ 1 \leq k \leq n-1.$$

– In the second step, we study the sign of $t_n(x)$ on the extrema. Before this, we remark that for $\theta = 0$, $x = 1 := z_{n,0}$ and for $\theta = \pi$, $x = -1 := z_{n,n}$, enabling us to get the following information on the action of t_n on $z_{n,k}$:

$$t_n(-1) = \frac{\cos(n\,\pi)}{2^{n-1}} = \frac{(-1)^n}{2^{n-1}} = t_n(z_{n,n}), \tag{1.13}$$

$$t_n(1) = \frac{\cos(n\,0)}{2^{n-1}} = \frac{1}{2^{n-1}} = t_n(z_{n,0}), \tag{1.14}$$

$$t_n(z_{n,k}) = \frac{(-1)^k}{2^{n-1}},\ 1 \leq k \leq n-1. \tag{1.15}$$

Equations (1.13) and (1.14) confirm that Eq. (1.15) which was initially valid for $1 \leq k \leq n-1$ is also valid for $k = 0,\ n$ and can then be written as

$$t_n(z_{n,k}) = \frac{(-1)^k}{2^{n-1}},\ 0 \leq k \leq n.$$

The previous equation, combined with the fact that $t_n(x) = \frac{\cos(n\theta)}{2^{n-1}}$, allows us to deduce that

$$||t_n||_{\max} = \frac{1}{2^{n-1}}.$$

1. In the third step, we remark that the set I is not empty since it contains $||t_n||_{\max}$. In addition, it has zero as a lower bound. Let us assume that $||t_n||_{\max}$ is not the minimum element of I. Then there exists a polynomial q belonging to $\mathcal{P}_n$ such that

$$-\frac{1}{2^{n-1}} < q(x) < \frac{1}{2^{n-1}},\ -1 \leq x \leq 1.$$

We next set $P_{n-1}(x) = t_n(x) - q(x)$ and observe, taking into account the previous inequalities, that P_{n-1} which is a polynomial of degree at most $n-1$ fulfills the following properties:

$$P_{n-1}(z_{n,2j}) = t_n(z_{n,2j}) - q(z_{n,2j}) = \frac{1}{2^{n-1}} - q(z_{n,2j}) > 0,$$

$$P_{n-1}(z_{n,2j+1}) = t_n(z_{n,2j+1}) - q(z_{n,2j+1}) = \frac{-1}{2^{n-1}} - q(z_{n,2j+1}) < 0,$$

for any integer j such that $0 \le 2j+1 \le n$. We obtain a contradiction to the fact that the polynomial P_{n-1} which is of degree at most $n-1$ will have n zeros for it will change its sign n times in the intervals $(z_{n,k},\ z_{n,k+1})$, $k = 0 \ldots n-1$. We therefore conclude that $||t_n||_{\max} = \frac{1}{2^{n-1}}$ is the minimum of I.

Proof Illustration of Property 7 The Gauss formula (1.9) is given in the general case in the paper by A. S. Jooste in these proceedings (see also [3, 6, 12, 22], but one would need to proceed with additional careful computations to verify that the Christoffel number $\lambda_{n,k}$ in the general Gauss quadrature formula is given by $\lambda_{n,k} = \frac{\pi}{n}$ for the specific case of the Chebyshev polynomials T_n (see also [15], Theorem 8.4, where the Christoffel numbers have been given explicitly for Chebyshev polynomials of the first, second, third and fourth kinds). We refer to [6], page 33 and also to [3], page 252 for the proof of relation (1.10) and other approximation formulas.

□

2 Construction of a System of Orthogonal Polynomials

In this section, after having provided a concrete example of a family of orthogonal polynomials with proof of some of its nice properties—some of which are common for any family of orthogonal polynomials—, we will now show how to construct a family of orthogonal polynomials from a scalar product and then relate this with the definition of orthogonal polynomials.

Let us consider a scalar product $(\ ,\ )$ defined on $\mathbb{R}[x] \times \mathbb{R}[x]$ in terms of a Stieltjes integral as

$$(p, q) = \int_a^b p(x)\, q(x)\, d\alpha(x), \tag{2.1}$$

where $\mathbb{R}[x]$ is the ring of polynomials with a real variable and $d\alpha$ is a non-negative Borel measure supported in the interval (a, b). As scalar product, it fulfills the following properties:

$$
\begin{aligned}
(p, p) &\geq 0,\ \forall p \in \mathbb{R}[x],\ \text{and } (p, p) = 0 \Longrightarrow p = 0, \\
(p, q) &= (q, p),\ \forall p,\ q \in \mathbb{R}[x], \\
(\lambda p, q) &= \lambda (p, q),\ \forall \lambda \in \mathbb{R},\ \forall p,\ q \in \mathbb{R}[x], \\
(p + q, r) &= (p, r) + (q, r),\ \forall p,\ q,\ r \in \mathbb{R}[x].
\end{aligned}
$$

As an example of scalar product on $\mathbb{R}[x]$ with connection to known systems of orthogonal polynomials, we mention:

$$(p, q) = \int_{-1}^{1} p(x)\, q(x) \frac{dx}{\sqrt{1 - x^2}}, \tag{2.2}$$

which yields the Chebyshev orthogonal polynomials.

The following theorem provides a method for construction of a family of polynomials, orthogonal with respect to a given scalar product. It is called Gram-Schmidt orthogonalisation process.

Theorem 2.1 (Gram-Schmidt Orthogonalisation Process [6, 12, 22]) *The polynomial systems $(q_n)_n$ and $(p_n)_n$ defined recurrently by the relations*

$$q_0 = 1,\ q_n = x^n - \sum_{k=0}^{n-1} \frac{(x^n, q_k)}{(q_k, q_k)} q_k,\ n \geq 1,\ p_k = \frac{q_k}{\sqrt{(q_k, q_k)}},\ k \geq 0, \tag{2.3}$$

satisfy the relations

$$\deg(q_n) = \deg(p_n) = n, \forall n \geq 0, \tag{2.4}$$

$$(q_n, q_m) = 0,\ n \neq m,\ (q_n, q_n) \neq 0,\ \forall n \geq 0, \tag{2.5}$$

$$(p_n, p_m) = 0,\ n \neq m,\ (p_n, p_n) = 1,\ \forall n \geq 0. \tag{2.6}$$

The polynomials $(q_n)_n$ and $(p_n)_n$ are said to be orthogonal or orthonormal with respect to the scalar product $(\, ,\,)$, respectively. In fact, they represent the same polynomial system with different normalisation: $(q_n)_n$ is monic*—to say the coefficient of the leading monomial is equal to 1; while $(p_n)_n$ is orthonormal—to say $(p_n, p_n) = 1$ or the corresponding norm of p_n is equal to 1.*

Proof Equation (2.4) is obvious while Eq. (2.6) is a direct consequence of Eq. (2.5). We will prove Eq. (2.5) by induction. Because of the properties of the scalar product, we just need to prove the following:

$$(q_n, q_m) = 0,\ \forall n \geq 1,\ 0 \leq m \leq n - 1. \tag{2.7}$$

For $n = 1$, we have, using relations (2.3)

$$(q_1, q_0) = \left(x - \frac{(x, q_0)}{(q_0, q_0)}\, q_0, q_0\right) = (x, q_0) - \frac{(x, q_0)}{(q_0, q_0)}\,(q_0, q_0) = 0.$$

We now assume that relation (2.7) is satisfied up to a given $n \geq 1$. Let $m \in \mathbb{N}$, $0 \leq m \leq n$.

$$\begin{aligned}(q_{n+1}, q_m) &= \left(x^{n+1} - \sum_{k=0}^{n} \frac{(x^{n+1}, q_k)}{(q_k, q_k)}\, q_k, q_m\right)\\ &= (x^{n+1}, q_m) - \sum_{k=0}^{n} \frac{(x^{n+1}, q_k)}{(q_k, q_k)}\,(q_k, q_m)\\ &= (x^{n+1}, q_m) - \frac{(x^{n+1}, q_m)}{(q_m, q_m)}\,(q_m, q_m) = 0,\end{aligned}$$

since from the induction hypothesis (Eq. (2.7)) and the symmetry of the scalar product, $(q_k, q_m) = 0,\ 0 \leq k \neq m \leq n-1$. □

Definition 2.2 (Orthogonal Polynomials [6]) Any sequence of polynomials $(p_n)_n$ satisfying Eqs. (2.4) and (2.5) (rewritten as follows with q_n replaced by p_n)

$$\deg(p_n) = n, \tag{2.8}$$

$$\int_a^b p_n(x)\, p_m(x)\, d\alpha(x) = 0,\ n \neq m, \tag{2.9}$$

$$\int_a^b p_n(x)\, p_n(x)\, d\alpha(x) \neq 0,\ \forall n \geq 0, \tag{2.10}$$

is said to be orthogonal with respect to the measure $d\alpha$, and called an orthogonal polynomial system or an orthogonal polynomial for short.

Definition 2.3 (Orthogonal Polynomials w.r.t. to a Weight Function [3, 6, 12, 17, 18, 21]) When the measure $d\alpha$ is absolutely continuous, that is $d\alpha(x) = \rho(x)\, dx$ where ρ is an appropriate function—called *weight function*, then relations (2.8)–(2.10) read

$$\deg(p_n) = n, \tag{2.11}$$

$$\int_a^b p_n(x)\, p_m(x)\, \rho(x)\, dx = 0,\ n \neq m, \tag{2.12}$$

$$\int_a^b p_n(x)\, p_n(x)\, \rho(x)\, dx \neq 0,\ \forall n \geq 0. \tag{2.13}$$

The polynomial system $(p_n)_n$ is said to be orthogonal with respect to the *weight function* ρ. Because of the form of the orthogonality relation, the variable here is continuous. We therefore obtain orthogonal polynomials of a continuous variable.

Definition 2.4 (Orthogonal Polynomials of a Discrete Variable [12, 18]) When the measure $d\alpha$ is discrete and supported in $\mathbb{N}$, that is, $\alpha = \rho$ on $\mathbb{N}$, then the relations (2.8), (2.9) and (2.10) become

$$\deg p_n = n,\ n \geq 0,$$

$$\sum_{k=0}^{N} \rho(k)\, p_n(k)\, p_m(k) = 0,\ \forall n,\ m \in \mathbb{N},\ n \neq m,$$

$$\sum_{k=0}^{N} \rho(k)\, p_n(k)\, p_n(k) \neq 0,\ n \geq 0,$$

where the parameter N belongs to $\mathbb{N} \cup \{\infty\}$. $(p_n)_n$ is said to be orthogonal with respect to the discrete weight ρ. It is also called a sequence of orthogonal polynomials of a discrete variable.

Notice that if N is finite, then there exist only a finite number of orthogonal polynomials, this because the bilinear application defined in (2.1) is positive definite not on the entire $\mathbb{R}[x]$ but rather on its linear subspace, $\mathbb{R}_l[x]$, for an appropriate choice of the positive integer l.

Definition 2.5 (Orthogonal Polynomials of a q-Discrete Variable [8, 11, 13, 18]) When the measure $d\alpha$ is q-discrete and supported in $q^{\mathbb{Z}}$, that is, $\alpha = \rho$ on $q^{\mathbb{Z}}$, where $\mathbb{Z}$ is the set of integers, then the relations (2.8), (2.9) and (2.10) become

$$\deg p_n = n,\ n \geq 0,$$

$$\sum_{k=0}^{N} \rho(q^k)\, p_n(q^k)\, p_m(q^k) = 0,\ \forall n,\ m \in \mathbb{N},\ n \neq m,$$

$$\sum_{k=0}^{N} \rho(q^k)\, p_n(q^k)\, p_n(q^k) \neq 0,\ n \geq 0,$$

where the parameter N belongs to $\mathbb{N} \cup \{\infty\}$. $(p_n)_n$ is said to be orthogonal with respect to the q-discrete weight ρ. It is also called a sequence of orthogonal polynomials of a q-discrete variable.

Remark 2.6 When the measure $d\alpha$ is discrete or q-discrete supported on a quadratic or a q-quadratic lattice, this gives the orthogonal polynomials of a quadratic or a q-quadratic variable. As examples of such polynomials, we mention the Wilson and the Askey-Wilson polynomials [4, 8, 12, 13].

3 Basic Properties of Orthogonal Polynomials

3.1 The Uniqueness of a Family of Orthogonal Polynomials

Before stating the result about the uniqueness of a family of orthogonal polynomials, let us start with the following remarks:

1. If $(p_n)_n$ is a family of orthogonal polynomials, then due to the fact that the degree of each p_n is equal to n, any subset of $\{p_n, n \in \mathbb{N}\}$ is a linearly independent subset of the linear space $\mathbb{R}[x]$.
2. Moreover, for any $n \geq 1$, the set $\{p_k, 0 \leq k \leq n\}$, like the canonical basis of monomials $\{x^k, 0 \leq k \leq n\}$, constitutes a basis of the linear space $\mathbb{R}_n[x]$ of polynomials of degree at most n.

The following result states an equivalent orthogonality relation.

Lemma 3.1 *Let $(p_n)_n$ be a sequence of polynomials with* $\deg(p_n) = n$, $n \geq 0$. *Then Eqs. (2.9) and (2.10) are equivalent to the two following equations*

$$\int_a^b p_n(x)\, x^m\, d\alpha(x) = 0, \;\forall n \geq 1,\; 0 \leq m \leq n-1, \tag{3.1}$$

$$\int_a^b p_n(x)\, x^n\, d\alpha(x) \neq 0, \;\forall n \geq 0. \tag{3.2}$$

Proof The proof is obtained by combining the orthogonality relations (2.9) and (2.10) or (respectively (3.1) and (3.2)) with the expansion of the polynomial p_m in the canonical basis of monomials (respectively the expansion of x^m in the basis $\{p_k, 0 \leq k \leq n\}$). □

The uniqueness of a family of polynomials orthogonal with respect to a measure $d\alpha$ can then be stated as follows.

Theorem 3.2 (Uniqueness of a Family of Orthogonal Polynomials) *To the measure $d\alpha$ corresponds a unique (up to a multiplicative factor) family of orthogonal polynomials. Or equivalently, if $(p_n)_n$ and $(q_n)_n$ are two families of polynomials satisfying relations (2.8)–(2.10), then they are proportional, to say that there exists a sequence $(b_n)_n$ such that $p_n = b_n\, q_n$, $n \geq 0$, with $b_n \neq 0$, $n \geq 0$.*

Proof The proof is obtained by expanding the polynomial q_n in the basis $\{p_k, 0 \leq k \leq n\}$ of $\mathbb{R}_n[x]$ and using the orthogonality relations (2.9) and (2.10) to show that the other coefficients, except the leading one, are equal to zero. □

3.2 The Matrix Representation

The following results give information about the Hankel determinant and a matrix representation of a given family of orthogonal polynomials. Before that, let us define what we mean by linear functional and orthogonality with respect to a linear functional.

Definition 3.3 (Linear Functional) Linear functional here means any linear mapping from $\mathbb{R}[x]$ to $\mathbb{R}$.

The sequence of polynomials $(p_n)_n$ will be said to be orthogonal with respect to the linear functional $\mathcal{U}$ if $\deg(p_n) = n$ and

$$\langle \mathcal{U}, x^m p_n \rangle = 0,\ n \geq 0,\ 0 \leq m \leq n, \tag{3.3}$$

$$\langle \mathcal{U}, x^n p_n \rangle \neq 0,\ \forall n \geq 0. \tag{3.4}$$

In this case, the linear functional $\mathcal{U}$ is said to be quasi-definite, to say that there exists a family of polynomials orthogonal with respect to $\mathcal{U}$.

As example, we define a linear functional $\mathcal{L}$ by

$$\langle \mathcal{L}, p \rangle = \int_{-1}^{1} \frac{p(x)}{\sqrt{1-x^2}} dx,$$

corresponding to the Chebyshev orthogonal polynomials $(T_n)_n$. The definition of orthogonality by means of a linear functional is very useful in practice because it enables an elegant proof of equivalent properties of standard orthogonal polynomials [6, 7, 12, 16] in addition to providing the proof of the so-called Favard Theorem [6] stating that any sequence of polynomials satisfying a three-term recurrence relation with some specific restriction on one of its coefficients is orthogonal with respect to a quasi-definite functional.

Theorem 3.4 ([21]) *Let $(p_n)_n$ be a sequence of polynomials with $\deg(p_n) = n$, $n \geq 0$ and satisfying the orthogonality conditions (2.9) and (2.10).*

1. *Then for any integer $n \geq 0$, the following relation holds*

$$\Delta_n > 0,\ n \geq 0, \tag{3.5}$$

where Δ_n is the Hankel determinant defined by

$$\Delta_n = \det(\mu_{k+j})_{0\leq k,j\leq n} = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_1 & \cdots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} & \mu_{2n} \end{vmatrix},\ n \geq 1,\ \Delta_0 := \mu_0. \tag{3.6}$$

The number μ_n which is given by

$$\mu_n = \int_a^b x^n \, d\alpha(x), \;\; n \geq 0,$$

denotes the canonical moment with respect to the measure $d\alpha$.

2. *For any positive integer n, the polynomial p_n has the following matrix representation*

$$p_n(x) = \frac{a_{n,n}}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_1 & \cdots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ 1 & x & \cdots & x^{n-1} & x^n \end{vmatrix}, \tag{3.7}$$

where $a_{n,n}$ is the leading coefficient of p_n.

3. *Conversely, given any sequence of real numbers $(\mu_n)_n$ satisfying relation (3.5), then since $\Delta_n \neq 0$, $n \geq 0$, there exists a sequence of polynomials orthogonal with respect to the quasi-definite linear functional $\mathcal{U}$ defined on the canonical basis of monomials by*

$$\langle \mathcal{U}, x^n \rangle = \mu_n, \; n \geq 0.$$

In addition, from (3.5) the linear functional $\mathcal{U}$ is positive-definite and as a consequence (see [6]) there exists a positive Borel measure associated with it.
The corresponding family is given explicitly by (3.7).

Proof

1. For the proof of the first property, let $(p_n)_n$ be a sequence of polynomials with $\deg(p_n) = n$, $n \geq 0$ and satisfying the orthogonality conditions (2.9) and (2.10) which are equivalent to orthogonality conditions (3.1) and (3.2). Writing for a fixed integer $n \geq 1$

$$p_n(x) = \sum_{k=0}^{n} a_{n,k} \, x^k,$$

in the orthogonality relation (3.1) for the integers $m = 0 \ldots n$ and then for orthogonality relation (3.2), we obtain the following system of linear equations for the unknowns $(a_{n,k})_k$ whose matrix form is given by

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_1 & \cdots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} & \mu_{2n} \end{pmatrix} \begin{pmatrix} a_{n,0} \\ a_{n,1} \\ \vdots \\ a_{n,n-1} \\ a_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ k_n \end{pmatrix}, \tag{3.8}$$

where $k_n = \int_a^b p_n(x)\, x^n d\alpha \neq 0$. Since the polynomial sequence $(p_n)_n$ not only exists and is uniquely determined by fixing k_n, then necessarily, the Hankel determinant is different from zero. The positiveness of the Hankel's determinant will be deduced in the following paragraph.

2. To prove the second property, we first use (3.7) to obtain for $0 \leq m \leq n$ that

$$\int_a^b p_n(x)\, x^m\, d\alpha(x) = \frac{a_{n,n}}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_1 & \cdots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ \mu_m & \mu_{m+1} & \cdots & \mu_{m+n-1} & \mu_{m+n} \end{vmatrix}. \tag{3.9}$$

The previous relation reads

$$\int_a^b p_n(x)\, x^m\, d\alpha(x) = 0,\ 0 \leq m \leq n-1 \tag{3.10}$$

since the $m+1$-th row and the last row of the determinant will be identical. Also, use of (3.9) for $m = n$ taking combined with the following relation

$$\int_a^b p_n(x)\, x^n\, d\alpha(x) = \frac{1}{a_{n,n}} \int_a^b p_n(x)\, p_n(x)\, d\alpha(x) = \frac{d_n^2}{a_{n,n}}$$

obtained using orthogonality, leads to

$$\int_a^b p_n(x)\, x^n\, d\alpha(x) = a_{n,n} \frac{\Delta_n}{\Delta_{n-1}} = \frac{d_n^2}{a_{n,n}} \neq 0,\ n \geq 1. \tag{3.11}$$

We then deduce from Eqs. (3.10) and (3.11) combined with (3.1) and (3.2) that $(p_n)_n$ is orthogonal with respect to $d\alpha(x)$.

The positiveness of Δ_n is seen from the relation

$$\Delta_n = \Delta_0 \prod_{k=1}^{n} \frac{d_k^2}{a_{k,k}^2} = \mu_0 \prod_{k=1}^{n} \frac{d_k^2}{a_{k,k}^2} > 0,$$

deduced from (3.11).

3. The third property is proved by showing, in a similar way as done in the proof of Property 2 above, that the polynomial sequence given by (3.7) satisfies orthogonality relations (3.3) and (3.4).

□

3.3 The Three-Term Recurrence Relation

Theorem 3.5 (Three-Term Recurrence Relation [6, 21, 24]) *Any polynomial sequence $(p_n)_n$, orthogonal with respect to the measure $d\alpha$ or fulfilling the orthogonality relations (2.8)–(2.10), satisfies the following relation called three-term recurrence relation*

$$x\,p_n(x) = \frac{a_n}{a_{n+1}}p_{n+1} + \left(\frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}\right)p_n + \frac{a_{n-1}}{a_n}\,\frac{d_n^2}{d_{n-1}^2}p_{n-1},\ \ p_{-1}=0,\ p_0=1, \tag{3.12}$$

with

$$p_n = a_n\,x^n + b_n\,x^{n-1} + \textit{ lower degree terms},\quad \textit{and } d_n^2 = (p_n, p_n). \tag{3.13}$$

When (p_n) is monic (i.e. $a_n = 1$) or orthonormal (ie. $d_n = 1$), then Eq. (3.12) can be written in the following forms, respectively:

$$p_{n+1} = (x-\beta_n)\,p_n - \gamma_n\,p_{n-1},\ \ p_{-1}=0,\ p_0=1, \tag{3.14}$$

with $\beta_n = b_n - b_{n+1},\ \gamma_n = \frac{d_n^2}{d_{n-1}^2}$, and

$$x\,p_n = \alpha_{n+1}\,p_{n+1} + \eta_n\,p_n + \alpha_n\,p_{n-1},\ \ p_{-1}=0,\ p_0=1, \tag{3.15}$$

with $\alpha_n = \frac{a_{n-1}}{a_n},\ \eta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}$.

Also, the recurrence coefficients of the monic and orthonormal forms of the orthogonal polynomial system are connected by

$$\eta_n = \beta_n,\ \gamma_n = \alpha_n^2. \tag{3.16}$$

Proof For fixed $n \geq 0$, we expand $x\,p_n$ in the basis $\{p_0,\ p_1,\ldots,\ p_{n+1}\}$

$$x\,p_n = \sum_{k=0}^{n+1} c_{k,n}\,p_k,$$

and then use orthogonality to obtain

$$c_{k,n} = \frac{\int_a^b x p_n(x)\,p_k(x)\,d\alpha(x)}{\int_a^b p_k(x)\,p_k(x)\,d\alpha(x)} = \frac{\int_a^b p_n(x)\,x p_k(x)\,d\alpha(x)}{\int_a^b p_k(x)\,p_k(x)\,d\alpha(x)} = 0,\ \text{for } 0 \leq k \leq n-2.$$

Hence

$$x\,p_n = c_{n+1,n}\,p_{n+1} + c_{n,n}\,p_n + c_{n-1,n}\,p_{n-1}. \tag{3.17}$$

Substituting (3.13) into (3.12) and identifying the leading coefficients of the monomials x^{n+1} and x^n yields

$$c_{n+1,n} = \frac{a_n}{a_{n+1}}, \; c_{n,n} = \left(\frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}\right). \tag{3.18}$$

Using (3.17) twice combined with the orthogonality properties (2.9) and (2.10) gives

$$\begin{aligned} c_{n-1,n}\, d_{n-1}^2 &= \int_a^b p_n(x)(x)\, x p_{n-1}(x)\, d\alpha(x) \\ &= \int_a^b p_n(x)(x)\, [c_{n,n-1}\, p_n + c_{n-1,n-1}\, p_{n-1} + c_{n-2,n-1}\, p_{n-2}] d\alpha(x) \\ &= c_{n,n-1}\, d_n^2, \end{aligned}$$

from which we deduce using (3.18) that

$$c_{n-1,n} = c_{n,n-1} \frac{d_n^2}{d_{n-1}^2} = \frac{a_{n-1}}{a_n} \frac{d_n^2}{d_{n-1}^2}.$$

Equations (3.16) are obtained by identifying the coefficients of Eq. (3.14) with those of the monic form of Eq. (3.15). □

3.4 *The Christoffel-Darboux Formula*

The following formulas are consequences of the three-term recurrence relation.

Theorem 3.6 (Christoffel-Darboux Formula [6, 21]) *Any system of orthogonal polynomials satisfying the three-term recurrence relation (3.12), satisfies also a so-called Christoffel-Darboux formula given, respectively, in its initial and confluent forms as*

$$\sum_{k=0}^{n} \frac{p_k(x) p_k(y)}{d_k^2} = \frac{a_n}{a_{n+1}} \frac{1}{d_n^2} \frac{p_{n+1}(x)\, p_n(y) - p_{n+1}(y)\, p_n(x)}{x-y}, \; x \neq y, \tag{3.19}$$

$$\sum_{k=0}^{n} \frac{p_k(x) p_k(x)}{d_k^2} = \frac{a_n}{a_{n+1}} \frac{1}{d_n^2} \left(p_{n+1}'(x)\, p_n(x) - p_{n+1}(x)\, p_n'(x)\right). \tag{3.20}$$

Proof For the proof of Eq. (3.19), we multiply by $p_k(y)$, Eq. (3.17) in which n is replaced by k, to obtain

$$x\, p_k(x)\, p_k(y) = c_{k+1,k}\, p_{k+1}(x)\, p_k(y) + c_{k,k}\, p_k(x)\, p_k(y) + c_{k-1,k}\, p_{k-1}(x)\, p_k(y).$$

Interchanging the role of x and y in the previous equation, we obtain

$$y\, p_k(x)\, p_k(y) = c_{k+1,k}\, p_{k+1}(y)\, p_k(x) + c_{k,k}\, p_k(x)\, p_k(y) + c_{k-1,k}\, p_{k-1}(y)\, p_k(x).$$

Subtracting the last two equations from each other, we obtain that

$$\frac{p_k(x)\, p_k(y)}{d_k^2} = \frac{A_k(x, y) - A_{k-1}(x, y)}{x - y},$$

where

$$A_k(x, y) = \frac{c_{k+1,k}}{d_k^2}\,(p_{k+1}(x)\, p_k(y) - p_{k+1}(y)\, p_k(x)),$$

after taking into account the relation $\frac{c_{k+1,k}}{d_k^2} = \frac{a_k}{a_{k+1}}\frac{1}{d_k^2} = \frac{c_{k,k+1}}{d_{k+1}^2}$. Equation (3.19) is obtained by summing the previous equation for k from 0 to n, taking into account that $A_{-1}(x, y) = 0$ as $p_{-1}(x) = 0$. Equation (3.20) is obtained by taking the limit of (3.19) when y tends to x. □

3.5 The Interlacing Properties of the Zeros

The following properties of the zeros of orthogonal polynomials are direct consequences of the confluent form of the Christoffel-Darboux formula (3.20). Their proof is given in the lecture notes of A. Jooste in these proceedings.

Theorem 3.7 (On the Zeros of Orthogonal Polynomials [3, 12, 21, 22]) *If $(p_n)_n$ is a polynomial system, orthogonal with respect to the positive Borel measure $d\alpha$ supported on the interval (a, b), then we have the following properties:*

1. *p_n has n simple real zeros $x_{n,k}$ satisfying $a < x_{n,k} < b,\ 1 \le k \le n$.*
2. *p_n and p_{n+1} have no common zero. The same applies for P_n and P_n';*
3. *if $x_{n,1} < x_{n,2} < \cdots < x_{n,n}$ are the n zeros of p_n, then*

$$x_{n+1,k} < x_{n,k} < x_{n+1,k+1},\ \ 1 \le k \le n.$$

Remark 3.8 It should be noticed that the three-term recurrence relation in the current section yields a matrix representation of the multiplication operator, called the Jacobi matrix (see for instance [25]). From this fact one can deduce that the zeros of the n-th orthogonal polynomials are the eigenvalues of the leading principal submatrix of size $n \times n$ of such a Jacobi matrix. This provides a method to find in an efficient numerical way such zeros even in the quasi-definite case.

3.6 Solution to the $L^2(\alpha)$ Extremal Problem

Theorem 3.9 (Minimal Property) *Let $(p_n)_n$ be a sequence of monic polynomials orthogonal with respect to a positive Borel measure $d\alpha(x)$ supported on the real line. For any fixed positive integer n, p_n is the minimal polynomial with respect to the L^2-norm*

$$||p||_\alpha = \sqrt{\int p^2(x)\, d\alpha(x)}$$

associated with the corresponding orthogonality measure:

$$\min\left\{\int q_n^2(x)\, d\alpha(x),\ q_n \in \mathbb{R}[x], q_n(x) = x^n + \textit{lower degree terms}\right\} = \int p_n^2(x)\, d\alpha(x) \tag{3.21}$$

where $\mathbb{R}[x]$ is the ring of polynomials with real coefficients.

Proof Let $n \geq 1$ and q_n be a monic polynomial of degree n. Combining the expansion of q_n in terms of the $(p_k)_k$

$$q_n(x) = \sum_{k=0}^{n} a_{n,k}\, p_k(x),$$

with the orthogonality give

$$\int q_n^2(x)\, d\alpha(x) = \sum_{k=0}^{n} a_{n,k}^2\, d_k^2.$$

Therefore,

$$\int q_n^2(x)\, d\alpha(x) \geq a_{n,n}^2\, d_n^2 = d_n^2 = \int p_n^2(x)\, d\alpha(x).$$

In addition, there is equality if and only if $a_{n,k} = 0,\ 0 \leq k \leq n-1$. □

It should be noticed that relation (3.21) which is valid for any sequence of polynomial orthogonal to the positive Borel measure $d\alpha(x)$, is similar to relation (1.8), given for the specific case of the Chebyshev polynomials of the first kind, with the Sup-norm (instead of the corresponding L^2-norm).

3.7 Gauss Quadrature Formula

The following property, which is called the Gauss quadrature formula is valid for any sequence of polynomials orthogonal with respect to the weight function ρ.

Theorem 3.10 ([3, 12, 21, 22]) *Let $(p_n)_n$ be a family of polynomials satisfying orthogonality relations (2.11)–(2.13). Then there exists a sequence of positive real numbers $(\lambda_{n,k})_n$, called Christoffel numbers, such that*

$$\int_a^b \rho(x)\, f(x)dx = \sum_{k=1}^{n} \lambda_{n,k}\, f(x_{n,k}),\ \forall f \in \mathbb{R}_{2n-1}[x],\ n \geq 1, \tag{3.22}$$

where the $x_{n,k}$, $1 \leq k \leq n$ are the zeros of p_n ranked by increasing order. In addition, the integral of any function continuous on the compact interval $[a, b]$ can be approximated by the previous formula:

$$\int_a^b \rho(x)\, f(x)dx = \lim_{n\to\infty} \sum_{k=1}^{n} \lambda_{n,k}\, f(x_{n,k}),\ \forall f \in \mathcal{C}[a, b]. \tag{3.23}$$

Proof The proof of Eq. (3.22) which generalises property number 7 of the first theorem, is given in the paper by A. Jooste. It can also be found in [3, 12, 21, 22]. The proof of Eq. (3.23) is given in [3, 6]. □

3.8 Concluding Remarks

We would like to complete this paper with the following information and remark which will help to connect this lecture notes with the forthcoming ones, especially with those involved with classical and semi-classical orthogonal polynomials, as well as orthogonal polynomials of the Sobolev type:

1. Among the classes of orthogonal polynomials, we mention the classical orthogonal polynomials and the semi-classical orthogonal polynomials. The first class is contained in the second one.
2. Classical orthogonal polynomials of a continuous, discrete, q-discrete, quadratic and q-quadratic variable, respectively, are those orthogonal with respect to a weight function satisfying a so-called Pearson equation which is a first-order linear homogeneous differential, difference, q-difference, divided-difference or q-divided-difference equation with polynomial coefficients of degree one and at most 2, respectively, with some boundary conditions at the ends of the interval. Depending on the type of the variable, we get classical orthogonal polynomials of continuous, a discrete, a q-discrete, a quadratic and a q-quadratic variable.

   Semi-classical orthogonal polynomial are defined in the same way like the classical ones, but with less restriction on the degree of the polynomial coefficients of the Pearson equation which can take higher values.

3. The properties such as the uniqueness of a family of polynomials orthogonal with respect to a measure, the matrix representation, the three-term recurrence relation, the Christoffel-Darboux formula and its confluent form, the interlacing properties of the zeros and the Gauss quadrature formula are valid for any family of orthogonal polynomials. In addition, it should also be noticed that Theorems 3.4 (Matrix representation), 3.5 (Three-term recurrence relation), and 3.6 (Christoffel-Darboux formula) are also valid if we replace the positive Borel measure by a quasi-definite linear functional. In this case and for Theorem 3.4, the positiveness of the Hankel's determinant is to be replaced by the fact that this determinant does not vanish.
4. The Chebyshev polynomials of the first, second, third and fourth kinds are up to now the only known families of orthogonal polynomials for which the zeros are explicitly known. In addition to the Chebyshev polynomials of the first kind which have been studied here, the three other families are, respectively, given for $z = \cos\theta,\ 0 < \theta < \pi$, by [15, 21]

$$U_n(z) = \frac{\sin((n+1)\theta)}{\sin\theta},\ V_n(z) = \frac{\cos((n+\frac{1}{2})\theta)}{\cos(\frac{\theta}{2})},\ W_n(z) = \frac{\sin((n+\frac{1}{2})\theta)}{\sin(\frac{\theta}{2})}.$$

The zeros of $U_n(z)$, $V_n(z)$ and $W_n(z)$ are given in increasing order, respectively, by

$$\begin{aligned} z_{n,k} &= \cos\theta_{n,k},\ \text{with}\ \theta_{n,k} = \frac{n+1-k}{n+1}\pi,\ k = 1,\ 2, \ldots, n, \\ z_{n,k} &= \cos\theta_{n,k},\ \text{with}\ \theta_{n,k} = \frac{2(n-k)+1}{2n+1}\pi,\ k = 1, 2, \ldots, n, \\ z_{n,k} &= \cos\theta_{n,k},\ \text{with}\ \theta_{n,k} = \frac{2(n+1-k)}{2n+1}\pi,\ k = 1, 2, \ldots, n. \end{aligned}$$

5. Additional information on general orthogonal polynomials can be found for example in [5, 6, 12, 19–21].

Acknowledgements The author would like to thank the anonymous reviewer for his careful reading of the manuscript and his many insightful comments and suggestions.

References

1. W.A. Al-Salam, Characterization theorems for orthogonal polynomials, in *Orthogonal Polynomials: Theory and Practice*, ed. by P. Nevai. NATO ASI Series C, vol. 294 (Kluwer Academic Publishers, Dordrecht, 1990), pp. 1–24
2. G.E. Andrews, R. Askey, Classical orthogonal polynomials, in *Polynômes orthogonaux et applications*, ed. by C. Brezinski et al. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 36–62

3. G.E. Andrews, R. Askey, R. Roy, *Special Functions. Encyclopedia of Mathematics and Its Applications 71* (Cambridge University Press, Cambridge, 1999)
4. N.M. Atakishiyev, M. Rahman, S.K. Suslov, On classical orthogonal polynomials. Constr. Approx. **11**, 181–223 (1995)
5. C. Brezinski, *Padé-type Approximation and General Orthogonal Polynomials*. International Series of Numerical Mathematics, vol. 50 (Birkhäuser Verlag, Basel, 1980)
6. T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Gordon and Breach, New York, 1978)
7. A.G. García, F. Marcellán, L. Salto, A distributional study of discrete classical orthogonal polynomials. J. Comput. Appl. Math. **57**, 147–162 (1995)
8. G. Gasper, M. Rahman, *Basic Hypergeometric Series*. Encyclopedia of Mathematics and its Applications, vol. 35 (Cambridge University Press, Cambridge, 1990)
9. W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*. Numerical Mathematics and Scientific Computation (Oxford Science Publications, Oxford University Press, New York, 2004)
10. W. Gautschi, Orthogonal polynomials, quadrature, and approximation: computational methods and software (in Matlab), in *Orthogonal Polynomials and Special Functions*. Lecture Notes in Mathematics, vol. 1883 (Springer, Berlin, 2006), pp. 1–77
11. W. Hahn, Über Orthogonalpolynome, die q-Differenzengleichungen genügen. Math. Nachr. **2**, 4–34 (1949)
12. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
13. R. Koekoek, P.A. Lesky, R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer-Verlag, Berlin, 2010)
14. F. Marcellàn, A. Branquinho, J. Petronilho, Classical orthogonal polynomials: a functional approach. Acta Appl. Math. **34**, 283–303 (1994)
15. J.C. Mason, D.C. Handscomb, *Chebyshev Polynomials* (Chapman and Hall/CRC, New York, 2003)
16. J.C. Medem, On the q-polynomials: a distributional study. J. Comput. Appl. Math. **135**, 157–196 (2001)
17. A.F. Nikiforov, V.B. Uvarov, *Special Functions of Mathematical Physics* (Birkhäuser, Boston, 1984)
18. A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable* (Springer, Berlin, 1991)
19. H. Stahl, V. Totik, General orthogonal polynomials, in *Encyclopedia of Mathematics and Its Applications*, vol. 43 (Cambridge University Press, Cambridge, 1992)
20. P.K. Suetin, *Orthogonal Polynomials in Two Variables* (Gordon and Breach Science Publishers, Amsterdam, 1999)
21. G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, 4th edn. (American Mathematical Society, Providence, 1975)
22. N.M. Temme, *Special Functions: An Introduction to the Classical Functions of Mathematical Physics* (Wiley, New York, 1996)
23. L.N. Trefethen, *Approximation Theory and Approximation Practice* (Society for Industrial and Applied Mathematics, Philadelphia, 2013)
24. F. Tricomi, *Vorlesungen über Orthogonalreihen* (Springer, Berlin, 1955)
25. W. Van Assche, Non-symmetric linear difference equations for multiple orthogonal polynomials, in *CRM Proceedings and Lecture Notes*, vol. 25 (American Mathematical Society, Providence, 2000), pp. 391–405

Classical Continuous Orthogonal Polynomials

Salifou Mboutngam

Abstract Classical orthogonal polynomials (Hermite, Laguerre, Jacobi and Bessel) constitute the most important families of orthogonal polynomials. They appear in mathematical physics when Sturn-Liouville problems for hypergeometric differential equation are studied. These families of orthogonal polynomials have specific properties. Our main aim is to:

1. recall the definition of classical continuous orthogonal polynomials;
2. prove the orthogonality of the sequence of the derivatives;
3. prove that each element of the classical orthogonal polynomial sequence satisfies a second-order linear homogeneous differential equation;
4. give the Rodrigues formula.

Keywords Classical orthogonal polynomials · Rodrigues formula · Differential equation · Pearson type equation

Mathematics Subject Classification (2000) 33C45, 33D45

1 Definitions

Activity 1.1 ([3]) *Let n be a non-negative integer, $x \in (-1, 1)$. We consider:*

$$T_n(x) = \cos(n Arccos x). \tag{1.1}$$

S. Mboutngam (✉)
Higher Teachers' Training College, The University of Maroua, Maroua, Cameroon
e-mail: mbsalif@gmail.com

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_2

1. *Evaluate in terms of the integers n and m the quantity:*

$$I_{n,m} = \int_{-1}^{1} T_n(x)T_m(x)\rho(x)dx, \quad with \;\; \rho(x) = \frac{1}{\sqrt{1-x^2}}. \tag{1.2}$$

2. *Prove that*

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x). \tag{1.3}$$

3. *Evaluate* $\big((1-x^2)\rho(x)\big)' + x\rho(x)$.

Solution 1.2 (1.)

$$T_n(x) = \cos(nArccosx)$$

1. *By direct computation, we get*

$$I_{n,m} = \begin{cases} \pi \;\; if \;\; n = m = 0 \\ \frac{\pi}{2} \;\; if \;\; n = m \neq 0 \\ \quad 0 \;\; if \;\; n \neq m. \end{cases}$$

2.

$$\begin{aligned} T_{n+1}(x) &= \cos((n+1)Arccosx) \\ &= \cos(Arccosx)\cos(nArccosx) - \sin(Arccosx)\sin(nArccosx), \\ T_{n-1}(x) &= \cos((n-1)Arccosx) \\ &= \cos(Arccosx)\cos(nArccosx) + \sin(Arccosx)\sin(nArccosx). \end{aligned}$$

Then, we have

$$\begin{aligned} T_{n+1}(x) + T_{n-1}(x) &= 2\cos(Arccosx)\cos(nArccosx) \\ &= 2xT_n(x). \end{aligned}$$

Using this relation and the initial conditions $T_0(x) = 1, \; T_1(x) = x$*, we can prove that* $(T_n(x))_n$ *is a sequence of polynomials and for all non-negative integer* n*,* T_n *is of degree exactly* n*.*

3.

$$\begin{aligned}\Big((1-x^2)\rho(x)\Big)' + x\rho(x) &= (\sqrt{1-x^2})' + \frac{x}{\sqrt{1-x^2}}\\ &= \frac{-2x}{2\sqrt{1-x^2}} + \frac{x}{\sqrt{1-x^2}}\\ &= 0.\end{aligned}$$

Definition 1.3 Let ρ be a positive continuous function defined on an interval (a, b). ρ is called the weight function if there exists $(\mu_n)_n$ a sequence of complex numbers such that

$$\int_a^b x^n \rho(x)dx = \mu_n, \quad \forall n \in \mathbb{N}_0.$$

μ_n is called the nth moment associated to the weight function ρ.

Definition 1.4 Let ρ be a weight function defined on an interval (a, b). A sequence $(P_n)_n$ of polynomials is said to be an orthogonal polynomial sequence (OPS) with respect to ρ if, for each non-negative integer n, $\deg(P_n) = n$ and

$$\int_a^b P_n(x)P_m(x)\rho(x)dx = k_n\delta_{nm} \quad (k_n \neq 0), \quad m, n = 0, 1 \ldots .$$

Definition 1.5 ([7]) Let ρ be a weight function defined on an interval (a, b), $(P_n)_n$ an OPS with respect to ρ. $(P_n)_n$ is said to be a classical orthogonal polynomial sequence if there exist two polynomials ϕ of degree at most two and ψ of degree exactly one such that the weight function ρ satisfies the following differential equation called Pearson type equation:

$$(\phi(x)\rho(x))' = \psi(x)\rho(x) \tag{1.4}$$

and the boundary condition

$$\lim_{x\to a} x^n\phi(x)\rho(x) = 0 \text{ and } \lim_{x\to b} x^n\phi(x)\rho(x) = 0. \tag{1.5}$$

2 Orthogonality of the Derivatives

Activity 2.1 *Let n be a non-negative integer, $x \in (-1, 1)$. We consider the sequence $(T_n)_n$ defined by (1.1). Since $(T_n)_n$ is an OPS, T'_{n+1} is a polynomial of degree exactly n.*

Evaluate in terms of the integers n and m the quantity:

$$J_{n,m}=\int_{-1}^{1} T'_{n+1}(x)T'_{m+1}(x)\phi(x)\rho(x)dx, \quad \textit{with} \;\; \phi(x)=1-x^2, \; \rho(x)=\frac{1}{\sqrt{1-x^2}}. \tag{2.1}$$

Solution 2.2

$$T_n(x) = \cos(nArccosx),$$

$$J_{n,m} = \int_{-1}^{1} T'_{n+1}(x)T'_{m+1}(x)\phi(x)\rho(x)dx = \begin{cases} \frac{\pi}{2}(n+1)^2 & \textit{if} \;\; n=m, \\ 0 & \textit{if} \;\; n \neq m. \end{cases}$$

The sequence of the derivatives $\left(T'_{n+1}\right)_n$ *is an OPS with respect to the weight function*

$$\rho_1(x) = \phi(x)\rho(x).$$

Lemma 2.3 ([3]) *Let* ρ *be a weight function on the interval* (a, b) *and let* $(P_n)_n$ *be a sequence of polynomials. Then the following properties are equivalent:*

1. $(P_n)_n$ *is an orthogonal polynomial sequence with respect to the weight function* ρ.
2. $\int_a^b \pi(x)P_n(x)\rho(x)dx = 0$ *for every polynomial* π *of degree* $m < n$ *while* $\int_a^b \pi(x)P_n(x)\rho(x)dx \neq 0$ *if* $m = n$.
3. $\int_a^b x^m P_n(x)\rho(x)dx = K_n\delta_{mn}$ *where* $K_n \neq 0, \quad m = 0, 1, \cdots, n$.

Proposition 2.4 *If* $(P_n)_n$ *is a classical orthogonal polynomial sequence such that the corresponding weight function* ρ *satisfies the Pearson type equation (1.4), the sequence* $(P'_{n+1})_n$ *is orthogonal with respect to the weight function* $\rho_1(x) = \phi(x)\rho(x)$.

Proof Let $(P_n)_n$ be a classical orthogonal polynomial sequence such that the corresponding weight function ρ satisfy the Pearson type equation (1.4). Let n and m be two non-negative integers such that $m < n$.

The fact that ψ is a polynomial of degree exactly one implies that $x^{m-1}\psi(x)$ is of degree m and according to Lemma 2.3, we have

$$0 = \int_a^b P_n(x)x^{m-1}\psi(x)\rho(x)dx.$$

Using the Pearson type equation, we have

$$\begin{aligned}0=\int_a^b P_n(x)x^{m-1}\psi(x)\rho(x)dx &= \int_a^b P_n(x)x^{m-1}\left(\phi(x)\rho(x)\right)'dx\\ &= P_n(x)x^{m-1}\phi(x)\rho(x)\Big|_a^b-\int_a^b\left(P_n(x)x^{m-1}\right)'\phi(x)\rho(x)dx.\end{aligned}$$

The boundary condition (1.5) implies that $P_n(x)x^{m-1}\phi(x)\rho(x)\Big|_a^b=0$ and we have:

$$\begin{aligned}0&=-\int_a^b\left(P_n(x)x^{m-1}\right)'\phi(x)\rho(x)dx\\ &=-\int_a^b P_n'(x)x^{m-1}\phi(x)\rho(x)dx-(m-1)\int_a^b P_n(x)x^{m-2}\phi(x)\rho(x)dx.\end{aligned}$$

The fact that ϕ is a polynomial of degree at most two implies that $x^{m-2}\phi(x)$ is of degree at most m and according to Lemma 2.3, we have

$$\int_a^b P_n'(x)x^{m-1}\phi(x)\rho(x)dx=0,\quad m<n.$$

ψ is a polynomial of degree exactly one and according to Lemma 2.3, we have

$$\int_a^b P_n(x)x^{n-1}\psi(x)\rho(x)dx\neq 0.$$

Using the Pearson type equation, we have

$$\begin{aligned}0\neq\int_a^b P_n(x)x^{n-1}\psi(x)\rho(x)dx &= \int_a^b P_n(x)x^{n-1}\left(\phi(x)\rho(x)\right)'dx\\ &= P_n(x)x^{n-1}\phi(x)\rho(x)\Big|_a^b-\int_a^b\left(P_n(x)x^{n-1}\right)'\phi(x)\rho(x)dx.\end{aligned}$$

The boundary condition (1.5) implies that $P_n(x)x^{n-1}\phi(x)\rho(x)\Big|_a^b=0$ and we have:

$$\int_a^b\left(P_n(x)x^{n-1}\right)'\phi(x)\rho(x)dx=\int_a^b P_n'(x)x^{n-1}\phi(x)\rho(x)dx+(n-1)\int_a^b P_n(x)x^{n-2}\phi(x)\rho(x)dx.$$

The fact that ϕ is a polynomial of degree at most two implies that $x^{n-2}\phi(x)$ is of degree at most n and according to Lemma 2.3, we have

$$\int_a^b P_n'(x)x^{n-1}\phi(x)\rho(x)dx\neq 0.$$

Then according to Lemma 2.3, $(P'_{n+1})_n$ is an orthogonal polynomial sequence with respect to the weight function $\rho_1(x) = \phi(x)\rho(x)$ in the interval (a, b). $\square$

In addition, we have

$$\begin{aligned}[\phi(x)\rho_1(x)]' &= \phi'(x)\rho_1(x) + \phi(x)\rho_1'(x) \\ &= \phi'(x)\rho_1(x) + \phi(x)(\phi(x)\rho(x))' \\ &= (\phi'(x) + \psi(x))\rho_1(x).\end{aligned}$$

ρ_1 satisfies a differential equation of type (1.4). This implies that the sequence $(P'_n)_n$ is also classical.

Theorem 2.3 ([3]) *Let m be a fixed non-negative integer and $(P_n)_n$ be a classical orthogonal polynomial sequence with respect to a weight function ρ which satisfies the Pearson type equation (1.4). $\left(\frac{d^m}{dx^m} P_{n+m}\right)_n$ is a classical orthogonal polynomial sequence with respect to the weight function*

$$\rho_m(x) = (\phi(x))^m \rho(x).$$

We have also

$$(\phi(x)\rho_m(x))' = \psi_m(x)\rho_m(x),$$

where

$$\psi_m(x) = \psi(x) + m\phi'(x).$$

3 Second-Order Differential Equation

Activity 3.1 *Let n be a non-negative integer, $x \in (-1, 1)$. We consider:*

$$T_n(x) = \cos(nArccosx).$$

Compute $T_n'(x)$, $T_n''(x)$ and evaluate the expression

$$(1 - x^2)T_n''(x) - xT_n''(x).$$

Solution 3.2

$$T_n'(x) = \frac{n\sin(nArccosx)}{\sqrt{1-x^2}},$$

$$T_n''(x) = -\frac{n^2\cos(nArccosx)}{1-x^2} + \frac{nx\sin(nArccosx)}{(1-x^2)\sqrt{1-x^2}}.$$

Therefore

$$(1-x^2)T_n''(x) - xT_n'(x) = -n^2 T_n(x),$$

that is,

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0. \tag{3.1}$$

Each term of the sequence $(T_n)_n$ *satisfies the second order differential equation (3.1).*

Lemma 3.3 *If* $(P_n)_n$ *is an OPS with respect to the weight function* ρ *on the interval* (a, b)*, each* P_n *is uniquely determined up to an arbitrary non-zero factor. That is, if* $(Q_n)_n$ *is also an OPS with respect to* ρ*, then there are constants* $c_n \neq 0$ *such that*

$$Q_n(x) = c_n P_n(x), \quad n = 0, 1, 2, \ldots.$$

Theorem 3.4 *Let* $(P_n)_n$ *be a classical orthogonal polynomial sequence such that the corresponding weight function* ρ *satisfies the Pearson type equation (1.4) in the interval* (a, b)*. For all non-negative integer* n*, we have*

$$\phi(x)P_n''(x) + \psi(x)P_n'(x) + \lambda_n P_n(x) = 0, \tag{3.2}$$

with

$$\lambda_n = -n\psi' - \frac{n(n-1)}{2}\phi''. \tag{3.3}$$

Proof Let $(P_n)_n$ be a classical orthogonal polynomial sequence such that the corresponding weight function ρ satisfies the Pearson type equation (1.4) in the interval (a, b). $(P_{n+1}')_n$ is an orthogonal polynomial sequence with respect to

$\rho_1(x) = \phi(x)\rho(x)$. Let n be a fixed non-negative integer. We have $\forall m < n$

$$\begin{aligned} 0 &= \int_a^b P_n'(x)(x^m)'\phi(x)\rho(x)dx \\ &= P_n'(x)x^m\phi(x)\rho(x)\Big|_a^b - \int_a^b \left[\phi(x)\rho(x)P_n'(x)\right]' x^m dx \\ &= -\int_a^b \left[\phi(x)P_n''(x) + \psi(x)P_n'(x)\right] x^m \rho(x)dx. \end{aligned}$$

Then

$$\int_a^b \left[\phi(x)P_n''(x) + \psi(x)P_n'(x)\right] x^m \rho(x)dx = 0, \quad m < n.$$

Since $deg(\phi P_n'' + \psi P_n') = n$, then $(\phi P_n'' + \psi P_n')_n$ is an orthogonal polynomial sequence with respect to the weight function ρ, and according to Lemma 3.3, there exists a constant λ_n such that

$$\phi(x)P_n''(x) + \psi(x)P_n'(x) + \lambda_n P_n(x) = 0. \tag{3.4}$$

The coefficient of x^n in (3.4) is

$$\lambda_n a_n + n\psi' a_n + \frac{n(n-1)}{2}\phi'' a_n = 0$$

where a_n is the leading coefficient of P_n. Then, we conclude that if (P_n) is a classical orthogonal polynomial sequence, then for all non-negative integer n, P_n satisfies a second order differential equation of the form (3.2) where the constant λ_n is given by (3.3). □

4 Rodrigues' Formula

Let $(P_n)_n$ be a classical orthogonal polynomial sequence such that the corresponding weight function ρ satisfies the Pearson type equation (1.4) in an interval (a, b). $\forall n \in \mathbb{N}$, P_n satisfies the second order differential equation (3.2). Multiplying (3.2) by $\rho(x)$ and using Pearson's equation, we can rewrite (3.2) as

$$\left[\phi(x)\rho(x)P_n'(x)\right]' + \lambda_n \rho(x)P_n(x) = 0.$$

Thus

$$-\lambda_n \rho(x)P_n(x) = \left[\phi(x)\rho(x)P_n'(x)\right]'. \tag{4.1}$$

Since $(P_n')_n$ is also a classical orthogonal polynomial sequence with respect to $\rho_1(x)$, we also get using the same process the equation

$$-\mu_{n,1}\rho_1(x)P_n'(x) = \left[\phi(x)\rho_1(x)P_n''(x)\right]', \tag{4.2}$$

with $\mu_{n,1} = \psi' + \lambda_n$. (4.2) in (4.1) yields

$$\rho(x)P_n(x) = \frac{1}{\lambda_n\mu_{n,1}}\left[\phi^2(x)\rho(x)P_n''(x)\right]''.$$

Continuing the process, we obtain

$$\rho(x)P_n(x) = \frac{(-1)^n}{\lambda_n\mu_{n,1}\cdots\mu_{n,n-1}}\left[\phi^n(x)\rho(x)\frac{d^n}{dx^n}P_n(x)\right]^{(n)};$$

where

$$\mu_{n,k+1} = \psi_k' + \mu_{n,k},\ \ k \geq 0,\quad \mu_{n,0} = \lambda_n,$$

with

$$\psi_{k+1}(x) = \phi'(x) + \psi_k(x),\quad \psi_0(x) = \psi(x).$$

The fact that P_n is a polynomial of degree n implies that $\frac{d^n}{dx^n}P_n(x) = n!a_n$ where a_n is the leading coefficient of P_n. Thus we have

$$P_n(x) = \frac{A_n}{\rho(x)}\left[\phi^n(x)\rho(x)\right]^{(n)}.$$

5 Classification of Classical Orthogonal Polynomials of a Continuous Variable

A classical orthogonal polynomial sequence is such that the corresponding weight function ρ satisfies the Pearson type equation

$$(\rho(x)\phi(x))' = \psi(x)\rho(x), \tag{5.1}$$

where ϕ is a polynomial of degree at most two and ψ is a polynomial of degree one. Equation (5.1) is equivalent to

$$\frac{\rho'(x)}{\rho(x)} = \frac{\psi(x) + \phi'(x)}{\phi(x)}.$$

The solution of this differential equation is given by

$$\rho(x) = \exp\left(\int \frac{\psi(x) + \phi'(x)}{\phi(x)} dx\right). \tag{5.2}$$

We will give the classification in terms of the degree of the polynomial ϕ.

5.1 Classical Orthogonal Polynomials Obtained if $deg(\phi) = 2$

We suppose that $deg(\phi) = 2$, ϕ has two distinct zeros x_0 and x_1 and the boundary condition (1.5) is satisfied. We take $x_0 = a$, $x_1 = b$ and then, there exist α and β such that

$$\frac{\psi(x) + \phi'(x)}{\phi(x)} = \frac{\alpha}{b - x} + \frac{\beta}{a - x}.$$

Using the transformation $t = \frac{2x-a-b}{b-a}$, $\phi(t) = (1 - t^2)$, the interval of orthogonality (a, b) becomes $(-1, 1)$, and from (5.2), we get $\rho(t) = (1 - t)^\alpha(1 + t)^\beta$. Using the Pearson equation, we deduce that $\psi(t) = -(\alpha + \beta + 2)t + \beta - \alpha$. The fact that $\int_{-1}^{1} t^n \rho(t)dt$ converges for all non-negative integer n requires that $\alpha, \beta > -1$.

The data $\phi(x) = 1 - x^2$, $x \in (-1, 1)$ and

$$\rho(x) = (1 - x)^\alpha(1 + x)^\beta,\ \alpha > -1,\ \beta > -1,$$

correspond to the Jacobi polynomials denoted by $(P_n^{(\alpha,\beta)}(x))_n$.

Second-Order Differential Equation of the Jacobi Polynomials Since $\phi(x) = 1 - x^2$, $\psi(x) = \beta - \alpha + (\alpha + \beta + 2)x$ and $\lambda_n = n(n + \alpha + \beta + 1)$, the second-order differential equation satisfied by the Jacobi polynomials is given by [6, 7]

$$(1-x^2)y''(x) + (\beta - \alpha + (\alpha + \beta + 2)x)\, y'(x) + n(n+\alpha+\beta+1)y(x) = 0,\ \ y(x) = P_n^{(\alpha,\beta)}(x).$$

Rodrigues' Formula It is given for the Jacobi polynomials by

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

5.2 Classical Orthogonal Polynomials Obtained if $deg(\phi) = 1$

If $deg(\phi) = 1$, there exist a, b and α such that

$$\frac{\psi(x) + \phi'(x)}{\phi(x)} = \frac{\alpha}{x - a} + b.$$

If we use the change of variable $t = \begin{cases} -x + a \\ x - a \end{cases}$, then $\phi(t) = t$ and the interval of orthogonality becomes $(0, +\infty)$. From (5.2), we get $\rho(t) = t^{\alpha} e^{-t}$, and we deduce from the Pearson equation that $\psi(t) = -t + \alpha + 1$. The fact that $\int_0^{+\infty} t^n \rho(t)dt$ converges for all non-negative integer n requires that $\alpha > -1$.

The data $\phi(x) = x$, $x \in (0, +\infty)$ and

$$\rho(x) = x^{\alpha} e^{-x}, \ \alpha > -1,$$

correspond to the Laguerre polynomials denoted by $(L_n^{(\alpha)}(x))_n$.

Second-Order Differential Equation of the Laguerre Polynomials Since $\phi(x) = x$, $\psi(x) = \alpha + 1 - x$ and $\lambda_n = n$, the second-order differential equation satisfied by the Laguerre polynomials is given by [6, 7]

$$xy''(x) + (\alpha + 1 - x)\, y'(x) + ny(x) = 0, \ \ y(x) = L_n^{(\alpha)}(x).$$

Rodrigues Formula

$$L_n^{(\alpha)}(x) = \frac{1}{n! x^{\alpha} e^{-x}} \left(x^{\alpha+n} e^{-x} \right)^{(n)}.$$

5.3 Classical Orthogonal Polynomials Obtained if $deg(\phi) = 0$

If $deg(\phi) = 0$, then there exist a and b such that

$$\frac{\psi(x) + \phi'(x)}{\phi(x)} = ax + b.$$

If we use the change of variable $t = -\psi(x)$, then $\phi(t) = 1$ and the interval of orthogonality becomes $(-\infty, +\infty)$. From (5.2), we get $\rho(t) = e^{-t^2}$, and we obtain $\psi(t) = -2t$ using Pearson's equation.

The data $\phi(x) = 1$, $(a, b) = \mathbb{R}$ and $\rho(x) = e^{-x^2}$, correspond to the Hermite polynomials denoted by $(H_n(x))_n$.

Second-Order Differential Equation of the Hermite Polynomials Since $\phi(x) = 1$, $\psi(x) = -2x$ and $\lambda_n = 2n$, the second-order differential equation satisfied by the Hermite polynomials is given by [6, 7]

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

Rodrigues Formula For the Hermite polynomials, we have the formula

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

6 Characterization Theorem of Classical Orthogonal Polynomials

The following theorem gives some characterization properties of classical continuous orthogonal polynomials.

Theorem 6.1 ([1, 2, 4, 5, 7]) *Let $(P_n)_n$ be a monic orthogonal polynomial sequence with respect to a weight function ρ on an interval (a, b) and $Q_{n,m}$ the monic polynomial of degree n defined by*

$$Q_{n,m}(x) = \frac{n!}{(n+m)!}\frac{d^m}{dx^m}P_{n+m}(x).$$

The following properties are equivalent:

1. *There exist two polynomials, ϕ of degree at most two and ψ of degree one, such that*

$$(\phi(x)\rho(x))' = \psi(x)\rho(x).$$

2. *There exist two polynomials, ϕ of degree at most two and ψ of degree one, such that for any non-negative integer m,*

$$\left[(\phi(x))^{m+1}\rho(x)\right]' = (\psi(x) + m\phi'(x))(\phi(x))^m\rho(x),$$

$$\int_a^b Q_{j,m}(x)Q_{n,m}(x)(\phi(x))^m\rho(x)dx = k_n\delta_{j,n}, \quad k_n \neq 0; \quad \forall j, n \in \mathbb{N}_0.$$

3. *There exist two polynomials, ϕ of degree at most two and ψ of degree one, such that for any non-negative integer m, the following second-order differential equation holds:*

$$\phi(x)Q_{n,m}''(x) + (\psi(x) + m\phi'(x))Q_{n,m}'(x) + \mu_{n,m}Q_{n,m}(x) = 0,$$

with the constant $\mu_{n,m}$ given by

$$\mu_{n,m} = -n\left[\psi' + (2m+n-1)\frac{\phi''}{2}\right].$$

4. *There exist two polynomials, ϕ of degree at most two and ψ of degree one, such that for any non-negative integer m, the following relation holds*

$$Q_{n,m}(x) = \frac{A_{n,m}}{(\phi(x))^m \rho(x)} \frac{d^n}{dx^n}\left((\phi(x))^{m+n}\rho(x)\right).$$

7 Tutorial

Exercise 7.1 *We consider the sequence (U_n) defined by*

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad x = \cos\theta.$$

1. *Compute $U_0(x)$, $U_1(x)$, $U_2(x)$.*
2. *Prove that for all integer n, we have*

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

3. *Deduce that U_n is a polynomial of degree exactly n.*
4. *Evaluate*

$$\int_0^\pi \sin(n\theta)\sin(m\theta)d\theta, \quad m, n \in \mathbb{N}.$$

5. *Prove that*

$$\int_{-1}^{1} U_n(x)U_m(x)(1-x^2)^{\frac{1}{2}}dx = k_n\delta_{n,m}.$$

6. *Deduce that $(U_n)_{n\geq 0}$ is an orthogonal polynomial sequence on $(-1, 1)$ with respect to a weight function ρ which is to determine.*
7. *Give the Rodrigues formula of $(U_n)_{n\geq 0}$.*
8. *Give a second-order differential equation satisfied by U_n.*
9. *Give the Pearson type equation satisfied by ρ.*

Solution 7.2 *We consider the sequence $(U_n)_{n\geq 0}$ defined by*

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad x = \cos\theta.$$

1. *Compute $U_0(x)$, $U_1(x)$, $U_2(x)$.*
   $U_0(x) = 1,\ U_1(x) = \frac{\sin(2\theta)}{\sin\theta} = 2\cos\theta = 2x.$

$$\begin{aligned} U_2(x) = \frac{\sin(3\theta)}{\sin\theta} &= \frac{\sin(2\theta+\theta)}{\sin\theta} \\ &= \frac{\sin 2\theta\cos\theta + \cos 2\theta\sin\theta}{\sin\theta} \\ &= 2\cos^2\theta + \cos 2\theta \\ &= 2\cos^2\theta + 2\cos^2\theta - 1 \\ &= 4x^2 - 1. \end{aligned}$$

2. *We prove that for all integer n, we have*

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

   From the expansion

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

   and

$$U_{n-1}(x) = \frac{\sin(n+1)\theta\cos\theta}{\sin\theta} - \frac{\cos(n+1)\theta\sin\theta}{\sin\theta},$$

   we get $U_{n+1}(x) + U_{n-1}(x) = 2x\frac{\sin(n+1)\theta}{\sin\theta} = 2xU_n(x)$.
3. *Deduce that U_n is a polynomial of degree exactly n.*
   U_1 is a polynomial of degree 1. Fixed $n \in \mathbb{N}$ and assume that U_k is a polynomial of degree k, $\forall k \in \mathbb{N}, k \leq n$. We want to prove that U_{n+1} is a polynomial of degree $n+1$.

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

   Since U_n is a polynomial of degree n, $2xU_n(x)$ is a polynomial of degree $n+1$. The addition of two polynomials A and B gives a polynomial of degree $\max\{\deg(A), \deg(B)\}$. *It follows that U_{n+1} is a polynomial of degree $n+1$.*
   We conclude that U_n is a polynomial of degree n, $\forall n \in \mathbb{N}$.

4. *Evaluate*

$$\int_0^{\pi} \sin(n\theta)\sin(m\theta)d\theta, \quad m, n \in \mathbb{N}.$$

From the expansions $\cos(a+b) = \cos a \cos b - \sin a \sin b$; $\cos(a-b) = \cos a \cos b + \sin a \sin b$, *we get*
$\sin a \sin b = \frac{1}{2}(\cos(a-b) - \cos(a+b))$.

$$\text{If } m \neq n, \int_0^{\pi} \sin n\theta \sin m\theta d\theta = \frac{1}{2}\int_0^{\pi} \cos(n-m)\theta d\theta - \frac{1}{2}\int_0^{\pi} \cos(n+m)\theta d\theta$$
$$= 0.$$

If $n = m$, $\sin^2 n\theta = \frac{1}{2}(1 - \cos(2n\theta))$ *and then* $\int_0^{\pi} \sin^2(n\theta)d\theta = \frac{\pi}{2}$.
In conclusion,

$$\int_0^{\pi} \sin(n\theta)\sin(m\theta)d\theta = \begin{cases} 0 \text{ if } n \neq m \\ \\ \frac{\pi}{2} \text{ if } n = m. \end{cases}$$

5. *Let us compute*

$$I_{n,m} = \int_{-1}^{1} U_n(x)U_m(x)(1-x^2)^{\frac{1}{2}}dx.$$

$I_{n,m} = \int_{-1}^{1} \frac{\sin[(n+1)\arccos x]\sin[(n+1)\arccos x]}{\sin^2(\arccos x)}(1-x^2)^{\frac{1}{2}}dx.$
$x = cos\theta, \quad x = -1 \Rightarrow \theta = \pi; \quad x = 1 \Rightarrow \theta = 0 \;, dx = -\sin\theta d\theta.$

$$\begin{aligned} I_{n,m} &= \int_{\pi}^{0} \frac{\sin(n+1)\theta \sin(m+1)\theta(-\sin^2\theta)}{\sin^2\theta}d\theta \\ &= \int_0^{\pi} \sin(n+1)\theta \sin(m+1)\theta d\theta \\ &= \begin{cases} 0 & \text{if } n \neq m \\ \frac{\pi}{2} & \text{if } n = m. \end{cases} \end{aligned}$$

6. *We deduce from the last question 5 that the family* $(U_n)_{n\geq 0}$ *is orthogonal on* $(-1, 1)$ *with respect to the weight function* $\rho(x) = \sqrt{1-x^2}$.
7. *From the relation*

$$U_1(x) = \frac{A_1}{\rho(x)}[\phi(x)\rho(x)]',$$

we get

$$\phi(x) = \frac{1}{A_1\rho(x)}\int U_1(x)\rho(x)dx = \frac{-2}{3A_1}(1-x^2).$$

Taking $A_1 = -\frac{3}{2}$ *yields* $\phi(x) = 1 - x^2$ *and the Rodrigues formula of* (U_n) *is given by*

$$U_n(x) = \frac{A_n}{\sqrt{1-x^2}}\frac{d^n}{dx^n}\left[(1-x^2)^{n+\frac{1}{2}}\right].$$

8. *Give the second-order differential equation satisfied by* U_n.
   We have

$$U_1(x) = \frac{A_1}{\rho(x)}[\phi(x)\rho(x)]' = A_1\psi(x),$$

then $\psi(x) = -3x$. *Since* $\phi(x) = 1 - x^2$, *we get* $\lambda_n = -n\left((n-1)\frac{\phi''}{2} + \psi'\right) = n(n+2)$. *Therefore,* U_n *is solution of the second-order differential equation*

$$(1-x^2)y'' - 3xy' + n(n+2)y = 0.$$

9. *The Pearson type equation satisfied by* ρ *is given by*

$$((1-x^2)\rho(x))' = -3x\rho(x).$$

Exercise 7.3 *We consider* $(P_n)_{n\geq 0}$, *an orthogonal polynomial sequence defined for all integer n by the following Rodrigues formula:*

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

1. *Determine the weight function* ρ *associated to the family* $(P_n)_{n\geq 0}$.
2. *Give a second-order differential equation satisfied by* P_n.
3. *Let S be a polynomial. We assume that for all integer n,*

$$\int_{-1}^{1} S(x)P_n(x)dx = \frac{(-1)^n}{2^n n!}\int_{-1}^{1}\frac{d^n S(x)}{dx^n}(x^2-1)^n dx.$$

Prove that

$$\int_{-1}^{1} P_m(x)P_n(x)dx = \begin{cases} 0 \quad \textit{if} \quad n \neq m, \\ \frac{(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(1-x^2)^n dx \quad \textit{if} \quad n = m. \end{cases}$$

4. *We set*

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

Verify that

$$I_n = 2^{2n+1} \int_0^1 v^n (1 - v)^n dv,$$

where $1 - x = 2v$.

5. *Deduce that*

$$\int_{-1}^{1} P_m(x) P_n(x) dx = k_n \delta_{n,m}.$$

Solution 7.4

(1) We know that

$$\begin{aligned} P_n(x) &= \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[(x^2 - 1)^n \right] \\ &= \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[(x^2 - 1)^n \right] \\ &= \frac{A_n}{\rho(x)} \frac{d^n}{dx^n} \left[\phi^n(x) \rho(x) \right]. \end{aligned}$$

We deduce by identification

$$\rho(x) = 1, \phi(x) = 1 - x^2.$$

(2) The second order differential equation satisfied by P_n.
Since $\phi(x) = 1 - x^2$, $\rho(x) = 1$*, it follows from the Pearson equation*

$$(\phi(x)\rho(x))' = \psi(x)\rho(x)$$

that

$$\psi(x) = -2x.$$

By definition,

$$\begin{aligned}\lambda_n &= -n\left[\psi^{'} + \frac{(n-1)}{2}\phi^{''}\right]\\ &= -n\left[-2 + \frac{(n-1)}{2}(-2)\right]\\ &= n(2+n-1)\\ &= n(n+1).\end{aligned}$$

The second order differential equation satisfied by P_n is:

$$(1-x^2)P_n^{''} - 2xP_n^{'} + n(n+1)P_n = 0.$$

(3) We know that

$$\int_{-1}^{1} S(x)P_n(x)dx = \frac{(-1)^n}{2^n n!}\int_{-1}^{1}\frac{d^n}{dx^n}S(x)(x^2-1)^n dx,$$

and we want to prove that:

$$\int_{-1}^{1} P_m(x)P_n(x)dx = \begin{cases} 0 & \textit{if}\, n \neq m \\ \frac{(2n)!}{(2^n n!)^2}\int_{-1}^{1}(x^2-1)^n dx & \textit{if}\, n = m. \end{cases}$$

If $m \neq n$, then without loss of generality, we assume that $m < n$.

$$\begin{aligned}\int_{-1}^{1} P_m(x)P_n(x)dx &= \frac{(-1)^n}{2^n n!}\int_{-1}^{1}\frac{d^n}{dx^n}P_m(x)(x^2-1)^n dx\\ &= 0,\end{aligned}$$

since $\deg(P_m) = m < n$ *implies* $\frac{d^n}{dx^n}P_m(x) = 0$.

For $m = n$, $(x^2-1)^n$ is a monic polynomial of degree $2n$ and therefore

$$\begin{aligned}\int_{-1}^{1} P_n^2(x)dx &= \frac{(-1)^n}{2^n n!}\int_{-1}^{1}\frac{d^n}{dx^n}P_n(x)(x^2-1)^n dx\\ &= \frac{(-1)^n}{2^{2n}(n!)^2}\int_{-1}^{1}\frac{d^{2n}}{dx^{2n}}\left[(x^2-1)^n\right](x^2-1)^n dx\\ &= \frac{(-1)^n}{(2^n n!)^2}(2n)!\int_{-1}^{1}(x^2-1)^n dx\\ &= \frac{(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(x^2-1)^n dx.\end{aligned}$$

(4) We rewrite I_n as

$$I_n = \int_{-1}^{1} (1-x^2)^n dx = \int_{-1}^{1} (1-x)^n (1+x)^n dx.$$

By setting $1-x = v$*, we get* $dx = -2dv$, $\quad x = -1 \Rightarrow v = 1 \ ; x = 1 \Rightarrow v = 0.$ *This yields*

$$I_n = \int_{1}^{0} 2^{2n} v^n (1-v)^n (-2dv) = 2^{2n+1} \int_{0}^{1} v^n (1-v)^n dv.$$

(5) Using questions (3) and (4), we get

$$\int_{-1}^{1} P_m(x) P_n(x) dx = 0 \quad \textit{if} \quad m \neq n,$$

and

$$\int_{-1}^{1} P_n(x) P_n(x) dx = 2\frac{(2n)!}{(n!)^2} \int_{0}^{1} v^n (1-v)^n dv = \frac{2(2n)!}{(n!)^2} B(n+1, n+1) = \frac{2}{2n+1},$$

where

$$B(x, y) = \int_{0}^{1} t^{x-1} (1-t)^{y-1} dt = \frac{(x-1)!(y-1)!}{(x+y-1)!}, \ \textit{if}\, x, y \in \mathbb{N}.$$

References

1. W.A. Al-Salam, Characterization theorems for orthogonal polynomials, in *Orthogonal Polynomials: Theory and Practice*, ed. by P. Nevai. NATO ASI Series C: Mathematical and Physical Sciences, vol. 294 (Kluwer, Dordrecht, 1990), pp. 1–24
2. W.A. Al-Salam, T.S. Chihara, Another characterization of the classical orthogonal polynomials. SIAM J. Math. Anal. **3**, 65–70 (1972)
3. T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Gordon and Breach, New York, 1978), p. 259
4. M. Foupouagnigni, On difference equations for orthogonal polynomials on nonuniform lattices. J. Diff. Eqn. Appl. **14**, 127–174 (2008)
5. A.G. García, F. Marcellán, L. Salto, A distributional study of discrete classical orthogonal polynomials. J. Comput. Appl. Math. **5**(7), 147–162 (1995)
6. R. Koekoek, P.A. Lesky, R.F. Swarttouw, Hypergeometric orthogonal polynomials and their q-analogues, in *Springer Monographs in Mathematics* (Springer, Berlin, 2010), p. 578
7. F. Marcellán, A. Branquinho, J. Petronilho, Classical orthogonal polynomials: a functional approach. Acta Appl. Math. **34**, 283–303 (1994)

Generating Functions and Hypergeometric Representations of Classical Continuous Orthogonal Polynomials

Maurice Kenfack Nangho

Abstract The aim of this work is to show how to obtain generating functions for classical orthogonal polynomials and derive their hypergeometric representations.

Keywords Classical continuous orthogonal polynomials · Generating functions · Hypergeometric representations

Mathematics Subject Classification (2000) 33C45

1 Introduction

Definition 1.1 ([6]) Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of polynomials. The bivariate function $G(x,t)$ defined by

$$G(x,t) = \sum_{n=0}^{\infty} c_n p_n(x) t^n, \tag{1.1}$$

where $\{c_n\}_{n=0}^{\infty}$ is a sequence of real or complex numbers is a generating function of $\{p_n(x)\}_{n=0}^{\infty}$.

This work was completed with the support of the AIMS-Volkswagen Stiftung Workshop 2018.

M. Kenfack Nangho (✉)
University of Dschang, Department of Mathematics and Computer Science, Dschang, Cameroon
e-mail: maurice.kenfack@univ-dschang.org

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_3

As example, the generating function of the sequence of polynomials $\{x^n\}_{n=0}^{\infty}$, with $c_n = \frac{1}{n!}$ is

$$\exp(xt) = \sum_{n=0}^{\infty} \frac{x^n}{n!} t^n.$$

Generating functions appear as important tools in various domains of mathematics (Algebra, Statistics, Analysis,...). For instance they can be used to [7]

1. Find recurrence relation satisfied by the sequence;
2. Find other statistical properties of a sequence;
3. Find asymptotic formulae of a sequence;
4. Find exact formulae for a sequence.

Various methods have been developed for obtaining generating functions. Among them we have Rainville's, Weisner's, Truesdell's Method (cf. [3, 6]) as well as the one presented by Nikiforov et al. in their book [5] entitled "Classical Orthogonal Polynomials of a Discrete Variable". For a sequence of classical continuous orthogonal polynomials $\{p_n(x)\}_{n=0}^{\infty}$, orthogonal with respect to a weight function ρ on a real interval (a, b), Nikiforov et al. used the Rodrigues formula [1, 5]

$$p_n(x) = \frac{A_n}{\rho(x)} \frac{d^n}{dx^n} [\phi(x)^n \rho(x)],$$

where A_n depends on n and ϕ is a polynomial of degree at most two, to provide a generating function of the sequence $\{p_n(x)\}_{n=0}^{\infty}$. This work aims to present this method, where generating functions of classical orthogonal polynomials (Hermite, Laguerre and Jacobi) will be obtained as applications. Using the generating functions, the hypergeometric representations of continuous classical orthogonal polynomials will be derived.

2 Generating Functions of Classical Continuous Orthogonal Polynomials

We show how to obtain by means of Rodrigues formula, generating functions of classical continuous orthogonal polynomials and give explicit formulae for the Hermite, the Laguerre and the Jacobi polynomials.

Definition 2.1 ([1]) A sequence $\{p_n(x)\}_{n=0}^{\infty}$ of polynomials of a continuous variable, orthogonal with respect to a weight function ρ on a real interval (a, b) is classical if and only if there exists a polynomial ϕ of degree at most two and a sequence $\{A_n\}_{n=0}^{\infty}$ of numbers such that

$$p_n(x) = \frac{A_n}{\rho(x)} \frac{d^n}{dx^n} [\phi(x)^n \rho(x)]. \tag{2.1}$$

Definition 2.2 ([5]) A Generating function of a classical continuous orthogonal polynomial family $\{p_n(x)\}_{n=0}^{\infty}$ is a function $G(x,t)$ for which the series expansion in a neighborhood of $t=0$ is

$$G(x,t)=\sum_{n=0}^{\infty}\frac{p_n(x)}{A_n n!}t^n, \tag{2.2}$$

where A_n is the coefficient in the Rodrigues formula (2.1).

Theorem 2.3 ([5]) *Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of classical continuous orthogonal polynomials that satisfies the Rodrigues formula (2.1). The function*

$$G(x,t)=\frac{\rho(z)}{\rho(x)}\left.\frac{1}{1-\phi'(z)t}\right|_{z=\zeta(x,t)} \tag{2.3}$$

is a generating function of $\{p_n(x)\}_{n=0}^{\infty}$, where $\zeta(x,t)$ is the zero of $z-x-\phi(z)t$, satisfying $\lim_{t\to 0}\zeta(x,t)=x$.

Proof Since p_n, $n=0,1,2,\ldots$, is continuous and satisfies the Rodrigues formula

$$p_n(x)=\frac{A_n}{\rho(x)}\frac{d^n}{dx^n}[\phi(x)^n\rho(x)],$$

the function given by

$$G(x,t)=\sum_{n=0}^{\infty}\frac{p_n(x)}{A_n n!}t^n$$

is a generating function of $\{p_n(x)\}_{n=0}^{\infty}$. Let x be in the interval of orthogonality of $p_n(x)$ and let $\mathcal{C}$ be a circle surrounding x. We obtain from Cauchy's formula

$$\frac{d^n}{dx^n}[\phi(x)^n\rho(x)]=\frac{n!}{2i\pi}\int_{\mathcal{C}}\frac{\phi(z)^n\rho(z)dz}{(z-x)^{n+1}}.$$

So, the Rodrigues formula of $p_n(x)$ becomes

$$p_n(x)=\frac{A_n}{\rho(x)}\frac{n!}{2i\pi}\int_{\mathcal{C}}\frac{\phi(z)^n\rho(z)dz}{(z-x)^{n+1}}$$

and the generating function of $\{p_n(x)\}_{n=0}^{\infty}$ reads as

$$G(x,t)=\sum_{n=0}^{\infty}\frac{1}{2i\pi\rho(x)}\int_{\mathcal{C}}\frac{(\phi(z)t)^n\rho(z)dz}{(z-x)^{n+1}}.$$

The function $f : z \mapsto \frac{\phi(z)}{z-x}$ is bounded on the compact set $\mathcal{C}$, since it is continuous on $\mathcal{C}$. So, for $|t| < \frac{1}{3M}$, where M is an upper bound of f on $\mathcal{C}$, $\left|\frac{\phi(z)t}{z-x}\right|^n < \frac{1}{3^n}$ for all $z \in \mathcal{C}$. Since the series $\sum_{n=0}^{\infty}\frac{1}{3^n}$ converges, the series functions $\left\{\left(\frac{\phi(z)t}{z-x}\right)^n\right\}_{n=0}^{\infty}$ converges uniformly. Therefore, we can interchange the summation and the integral and then use the geometric series to obtain

$$\begin{aligned} G(x,t) &= \frac{1}{2i\pi\rho(x)}\int_{\mathcal{C}}\sum_{n=0}^{\infty}\frac{(\phi(z)t)^n\rho(z)dz}{(z-x)^{n+1}} \\ &= \frac{1}{2i\pi\rho(x)}\int_{\mathcal{C}}\frac{\rho(z)dz}{z-x-\phi(z)t}. \end{aligned} \tag{2.4}$$

To end the proof, we evaluate the above integral by using the residue theorem.
If ϕ is a polynomial of degree two, one of the zeros of the denominator

$$p(z) = z - x - \phi(z)t$$

in the integrand tends to ∞ when t tends to 0 and the other one, $\zeta(x,t)$, tends to x when t tends to 0.
If the polynomial ϕ is of degree at most one, the zero of $p(z)$ tends to x when t tends to 0. So, for $|t|$ sufficiently small, the integrand in

$$G(x,t) = \frac{1}{2i\pi\rho(x)}\int_{\mathcal{C}}\frac{\rho(z)dz}{z-x-\phi(z)t}$$

has a single pole, $\zeta(x,t)$, inside the circle $\mathcal{C}$. Therefore, using the residue formula and the l'Hospital rule, we get

$$\begin{aligned} G(x,t) &= \frac{1}{\rho(x)}\lim_{z\to\zeta(x,t)}\frac{(z-\zeta(x,t))\rho(z)}{z-x-\phi(z)t}, \\ &= \frac{\rho(z)}{\rho(x)}\left.\frac{1}{1-\phi'(z)t}\right|_{z=\zeta(x,t)}. \end{aligned}$$

Generating Function of Hermite Polynomials [2, p. 251][5, p. 29] The Hermite polynomials $\{H_n(x)\}_{n=0}^{\infty}$ are orthogonal polynomials associated with the weight $\rho(x) = \exp(-x^2)$ on the real line $\mathbb{R} = (-\infty, +\infty)$. They are known to satisfy the Rodrigues formula

$$H_n(x) = (-1)^n\exp(x^2)\frac{d^n}{dx^n}\left[\exp(-x^2)\right].$$

Identifying with the Rodrigues formula (2.1), we obtain $\phi(x) = 1$, $\rho(x) = \exp(-x^2)$ and $A_n = (-1)^n$. Therefore $z - x - \phi(z)t = z - x - t$ has only one zero $z = x + t$. Hence, from (2.2) and (2.3), we have

$$\sum_{n=0}^{\infty} \frac{H_n(x)}{(-1)^n n!} t^n = \left. \frac{\rho(z)}{\rho(x)} \right|_{z=x+t} = \exp(-2xt - t^2).$$

Substituting t by $-t$, we obtain the generating function

$$\exp(2xt - t^2) = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n. \tag{2.5}$$

Generating Function of Laguerre Polynomials [2, p. 242][5, p. 28] The Laguerre polynomials $\{L_n(x)\}_{n=0}^{\infty}$ are orthogonal polynomials associated with the weight $\rho(x) = x^\alpha \exp(-x)$ on the half-line $\mathbb{R}_+ = (0, +\infty)$ $(\alpha > -1)$. They are known to satisfy the Rodrigues formula

$$L_n^{(\alpha)}(x) = \frac{1}{n!} x^{-\alpha} \exp(x) \frac{d^n}{dx^n} [x^n x^\alpha \exp(-x)],$$

which is of the form (2.1) with $\rho(x) = x^\alpha \exp(-x)$, $\phi(x) = x$ and $A_n = \frac{1}{n!}$. Therefore the polynomial $p(z) = z - x - \phi(z)t = z - x - zt$ has only one zero $\zeta(x, t) = \frac{x}{1-t}$. Hence, from (2.2) and (2.3)

$$\sum_{n=0}^{\infty} L_n^{(\alpha)}(x) t^n = \left. \frac{\rho(z)}{\rho(x)} \frac{1}{1-t} \right|_{z=\frac{x}{1-t}} = \frac{\rho(\frac{x}{1-t})}{(1-t)\rho(x)}.$$

Taking into account the fact that $\rho(x) = x^\alpha \exp(-x)$, we get the generating function

$$(1-t)^{-\alpha-1} \exp\left(\frac{xt}{t-1}\right) = \sum_{n=0}^{\infty} L_n^{(\alpha)}(x) t^n. \tag{2.6}$$

Generating Function of Jacobi Polynomials [2, p. 218] The Jacobi polynomials $\{P_n^{(\alpha,\beta)}(x)\}_{n=0}^{\infty}$ $(\alpha > -1,\ \beta > -1)$ are orthogonal with respect to the weight function $\rho(x) = (1-x)^\alpha (1+x)^\beta$ on the interval $(-1, 1)$. These polynomials are known to satisfy the Rodrigues formula

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{n! 2^n} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} [(1-x)^{\alpha+n} (1+x)^{\beta+n}], \tag{2.7}$$

which is of the form (2.1), with $\rho(x) = (1-x)^\alpha(1+x)^\beta$, $\phi(x) = (1-x^2)$ and $A_n = \frac{(-1)^n}{n!2^n}$. Therefore the polynomial $p(z) = z - x - \phi(z)t = z - x - (1-z^2)t$ is of degree two with zeros

$$\zeta_1(x,t) = \frac{-1-\sqrt{1+4tx+4t^2}}{2t} \text{ and } \zeta_2(x,t) = \frac{-1+\sqrt{1+4tx+4t^2}}{2t}.$$

Noting that $\lim_{t\to 0}\zeta_1(x,t) = \infty$ and $\lim_{t\to 0}\zeta_2(x,t) = x$, we deduce from (2.2) and (2.3) the following

$$\begin{aligned}\sum_{n=0}^{\infty}(-2)^n P_n^{(\alpha,\beta)}(x)t^n &= \frac{\rho(z)}{\rho(x)}\,\frac{1}{1+2zt}\Big|_{z=\zeta_2(x,t)},\\ &= \frac{\rho(\zeta_2(x,t))}{\rho(x)}\frac{1}{1+2\zeta_2(x,t)t}.\end{aligned}$$

Since $\zeta_2(x,t) = \frac{-1+\sqrt{1+4tx+4t^2}}{2t}$, we obtain after simplification

$$\rho(\zeta_2(x,t)) = \frac{2^{\alpha+\beta}(1-x)^\alpha(1+x)^\beta}{\left(1+2t+\sqrt{1+4tx+4t^2}\right)^\alpha\left(1-2t+\sqrt{1+4tx+4t^2}\right)^\beta},$$

$$\frac{1}{1+2t\zeta_2(x,t)} = \frac{1}{\sqrt{1+4tx+4t^2}}.$$

Therefore

$$\begin{aligned}&\sum_{n=0}^{\infty}(-2)^n P_n^{(\alpha,\beta)}(x)t^n\\ &= \frac{2^{\alpha+\beta}}{\left(1+2t+\sqrt{1+4tx+4t^2}\right)^\alpha\left(1-2t+\sqrt{1+4tx+4t^2}\right)^\beta\sqrt{1+4tx+4t^2}}.\end{aligned}$$

Substituting t by $-\frac{t}{2}$, we obtain

$$G(x,t) = \sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)t^n \tag{2.8}$$

with

$$G(x,t) = \frac{2^{\alpha+\beta}}{(1-t+R)^\alpha\,(1+t+R)^\beta\,R},\quad R = \sqrt{1-2tx+t^2}.$$

3 Hypergeometric Representations of Classical Orthogonal Polynomials

In this section, we define hypergeometric functions, recall some of their fundamental properties and derive the hypergeometric representation of classical continuous orthogonal polynomials using their generating functions.

Hypergeometric function ${}_pF_q$ (also called generalized hypergeometric function) is defined by the series [2, 6]

$$ {}_pF_q\left(\begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{matrix};z\right)=\sum_{k=0}^{\infty}\frac{(a_1)_k\cdots(a_p)_k}{(b_1)_k\cdots(b_q)_k}\frac{z^k}{k!}, \tag{3.1} $$

where $(a)_n$ is the shifted factorial or Pochhammer symbol defined as follows

$$ (a)_n=a(a+1)(a+2)\cdots(a+n-1), \quad n\geq 1, $$

and

$$ (a)_0=1. $$

The parameters must be such that the denominator factors $(b_i)_k$, $i=1,..,q$ are never zero. When one of the numerator parameters, let us say a_1, equals $-n$, where n is a nonnegative integer, this hypergeometric function is a polynomial in z

$$ {}_pF_q\left(\begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{matrix};z\right)=\sum_{k=0}^{n}\frac{(-n)_k\cdots(a_p)_k}{(b_1)_k\cdots(b_q)_k}\frac{z^k}{k!}. $$

The radius of convergence R of the hypergeometric series is given by

$$ R=\begin{cases} \infty & \text{if } p<q+1 \\ 1 & \text{if } p=q+1 \\ 0 & \text{if } p>q+1. \end{cases} $$

When $p=2$ and $q=1$ the function (3.1) becomes

$$ {}_2F_1\left(\begin{matrix} a,b \\ c \end{matrix};z\right)=\sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{z^k}{k!} \tag{3.2} $$

with $|z|<1$ and c a complex number different from $-1,-2,\ldots$. This function is known to satisfy the second order differential equation [6, p. 53]

$$ z(1-z)y''(z)+[c-(a+b+1)z]y'-aby=0 \tag{3.3} $$

with the initial conditions $y(0)=1$ and $y'(0)=\frac{ab}{c}$.

Proposition 3.1

$$\left(\frac{2}{1+\sqrt{1-z}}\right)^{\alpha} = {}_2F_1\left(\begin{matrix}\frac{\alpha+1}{2}, \frac{\alpha}{2} \\ \alpha+1\end{matrix}; z\right), \tag{3.4}$$

$$(1-z)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-z}}\right)^{\alpha} = {}_2F_1\left(\begin{matrix}\frac{\alpha+1}{2}, \frac{\alpha+2}{2} \\ \alpha+1\end{matrix}; z\right) \tag{3.5}$$

with $|z| < 1$ *and* $\alpha \neq -2, -3, \ldots$.

Proof Let us prove (3.4). Substitute $y(z)$ by $f(z) = \left(\frac{2}{1+\sqrt{1-z}}\right)^{\alpha}$ in this equation

$$(\phi_2 z^2 + \phi_1 z + \phi_0)y''(z) + (\psi_1 z + \psi_0)y'(z) + \lambda\, y(z) = 0.$$

Take $z = 2, 3, 4, 5, 6, 7$ to obtain six equations and solve the system with unknowns $\phi_2, \phi_1, \phi_0, \psi_1, \psi_0, \lambda$ to obtain $\phi_2 = -1, \phi_1 = 1, \phi_0 = 0, \psi_1 = -\frac{2\alpha+3}{2}, \psi_0 = \alpha$ and $\lambda = -\frac{\alpha(\alpha+1)}{4}$. So, the function $f(z) = \left(\frac{2}{1+\sqrt{1-z}}\right)^{\alpha}$ satisfies the second order differential equation

$$z(1-z)y''(z) + [\alpha+1-(\alpha+\frac{3}{2})z]y'(z) - \frac{\alpha(\alpha+1)}{4}y(z) = 0.$$

Therefore, $f(z)$ satisfies (3.3) with $c = \alpha+1, a+b+1 = \alpha+\frac{3}{2}$ and $ab = \frac{\alpha(\alpha+1)}{4}$. That is $(a, b) \in \{(\frac{\alpha+1}{2}, \frac{\alpha}{2}), (\frac{\alpha}{2}, \frac{\alpha+1}{2})\}$ and $c = \alpha+1$. If $(a, b) = (\frac{\alpha+1}{2}, \frac{\alpha}{2})$, then $f(z)$ is solution to (3.3) with the initial condition $\frac{ab}{c} = \frac{\alpha}{4} = f'(0)$ and $f(0) = 1$. Therefore

$$\left(\frac{2}{1+\sqrt{1-z}}\right)^{\alpha} = {}_2F_1\left(\begin{matrix}\frac{\alpha+1}{2}, \frac{\alpha}{2} \\ \alpha+1\end{matrix}; z\right).$$

When $(a, b) = (\frac{\alpha}{2}, \frac{\alpha+1}{2})$, we obtain the same result. (3.5) is obtained in a similar way.

Hypergeometric Representation of Hermite Polynomials The generating function of Hermite polynomials is

$$\exp(2xt - t^2) = \exp(2xt)\exp(-t^2) = \sum_{n=0}^{\infty}\frac{(2xt)^n}{n!}\sum_{k=0}^{\infty}\frac{(-t^2)^k}{k!}.$$

By means of the Rainville relation [6, p. 58]

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}A(k, n) = \sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}A(k, n-2k), \tag{3.6}$$

where $\lfloor \frac{n}{2} \rfloor$ denotes the greatest positive integer less than or equal to $\frac{n}{2}$, we obtain

$$\exp(2xt - t^2) = \sum_{n=0}^{\infty} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (2x)^{n-2k}}{(n-2k)!k!} t^n.$$

Comparing coefficients of t^n in this result and in

$$\exp(2xt - t^2) = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n, \tag{3.7}$$

we obtain

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

Taking $a_k = n! \frac{(-1)^k (2x)^{n-2k}}{(n-2k)!k!}$, we have $\frac{a_{k+1}}{a_k} = \left(-\frac{n}{2} + k\right) \left(-\frac{n-1}{2} + k\right) \left(-\frac{1}{x^2}\right) \frac{1}{k+1}$. Iterating this relation and substituting k by $k+1$ yields

$$a_k = \left(-\frac{n}{2}\right)_k \left(-\frac{n-1}{2}\right)_k \left(-\frac{1}{x^2}\right)^k \frac{1}{k!} a_0.$$

Since $a_0 = (2x)^n$, we get

$$\begin{aligned} H_n(x) &= (2x)^n \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \left(-\frac{n}{2}\right)_k \left(-\frac{n-1}{2}\right)_k \frac{\left(-\frac{1}{x^2}\right)^k}{k!}, \\ &= (2x)^n {}_2F_0 \left(\begin{matrix} -\frac{n}{2}, -\frac{n-1}{2} \\ - \end{matrix} ; -\frac{1}{x^2} \right). \end{aligned}$$

Hypergeometric Representation of Laguerre Polynomials The shifted factorial is an extension of the ordinary factorial since $(1)_n = n!$, $n = 0, 1, 2, \ldots$. It is particularly convenient to use the shifted factorial or Pochhammer symbol in the binomial expansion [2, 6]

$$\begin{aligned} (1-t)^{-a} &= \sum_{n=0}^{\infty} \frac{(-a)(-a-1)\ldots(-a-n+1)}{n!} (-t)^n, \\ &= \sum_{n=0}^{\infty} \frac{(a)_n}{n!} t^n, \quad |t| < 1. \end{aligned} \tag{3.8}$$

The generating function $(1-t)^{-\alpha-1}\exp\left(\frac{xt}{t-1}\right)$ of the Laguerre polynomials can be written as follows

$$(1-t)^{-\alpha-1}\exp\left(\frac{xt}{t-1}\right)=\sum_{k=0}^{\infty}\frac{(-xt)^k}{k!}(1-t)^{-1-k-\alpha}.$$

By means of the binomial expansion (3.8),

$$(1-t)^{-\alpha-1}\exp\left(\frac{xt}{t-1}\right)=\sum_{k=0}^{\infty}\frac{(-xt)^k}{k!}\sum_{n=0}^{\infty}\frac{(1+k+\alpha)_n}{n!}t^n.$$

Using the Rainville formula [6, Lemma 10, Eq.1]

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}A(k,n)=\sum_{n=0}^{\infty}\sum_{k=0}^{n}A(k,n-k), \tag{3.9}$$

we obtain

$$(1-t)^{-\alpha-1}\exp\left(\frac{xt}{t-1}\right)=\sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(-x)^k(1+k+\alpha)_{n-k}}{k!(n-k)!}t^n.$$

Comparing coefficients of t^n in this result and in (2.6), we obtain

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

Taking $a_k=\frac{(-x)^k(1+k+\alpha)_{n-k}}{k!(n-k)!}$, we obtain $\frac{a_{k+1}}{a_k}=\frac{x(k-n)}{(\alpha+k+1)(k+1)}$. Iterating this relation and substituting $k+1$ by k yields

$$a_k=\frac{(-n)_k}{(\alpha+1)_k}\frac{x^k}{k!}a_0.$$

Since $a_0=\frac{(\alpha+1)_n}{n!}$, we get

$$L_n^{(\alpha)}(x)=\frac{(\alpha+1)_n}{n!}\sum_{k=0}^{n}\frac{(-n)_k}{(\alpha+1)_k}\frac{x^k}{k!}=\frac{(\alpha+1)_n}{n!}{}_1F_1\left(\begin{matrix}-n\\ \alpha+1\end{matrix};x\right).$$

Hypergeometric Representation of Jacobi Polynomials The generating function of the Jacobi polynomials is (see (2.8))

$$G(x,t) = \frac{2^{\alpha+\beta}}{(1-t+R)^{\alpha}(1+t+R)^{\beta}R} = \sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)t^n, \quad R = \sqrt{1-2tx+t^2}. \tag{3.10}$$

Observing that

$$1-t+R = (1-t)\left(1+\sqrt{1-\frac{2t(x-1)}{(1-t)^2}}\right),$$

$$1+t+R = (1+t)\left(1+\sqrt{1-\frac{2t(x+1)}{(1+t)^2}}\right),$$

$$R = (1+t)\sqrt{1-\frac{2t(x+1)}{(1+t)^2}},$$

we obtain

$$G(x,t) = \frac{1}{(1-t)^{\alpha}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x-1)}{(1-t)^2}}}\right)^{\alpha} \frac{1}{(1+t)^{\beta+1}}\left(1-\frac{2t(x+1)}{(1+t)^2}\right)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x+1)}{(1+t)^2}}}\right)^{\beta}.$$

Using the Eq. (3.5) with $z = \frac{2t(x+1)}{(1+t)^2}$ and $\alpha = \beta$ yields

$$\left(1-\frac{2t(x+1)}{(1+t)^2}\right)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x+1)}{(1+t)^2}}}\right)^{\beta} = {}_2F_1\left(\begin{matrix}\frac{\beta+1}{2}, \frac{\beta}{2} \\ \beta+1\end{matrix}; \frac{2t(x+1)}{(1+t)^2}\right)$$

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{\beta+1}{2}\right)_k \left(\frac{\beta}{2}\right)_k}{(\beta+1)_k} \frac{(2t(x+1))^k}{(1+t)^{2k}k!}.$$

Multiplying both sides by $\frac{1}{(1+t)^{\beta+1}}$ and using the binomial theorem (see (3.8)) with $\beta+2k+1$ taken for α and $-t$ taken for t, yields

$$\frac{1}{(1+t)^{\beta+1}}\left(1-\frac{2t(x+1)}{(1+t)^2}\right)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x+1)}{(1+t)^2}}}\right)^{\beta}$$

$$= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\frac{\beta+1}{2}\right)_k \left(\frac{\beta}{2}\right)_k}{(\beta+1)_k} \frac{(2t(x+1))^k}{k!} \frac{(\beta+2k+1)_n(-t)^n}{n!}.$$

By means of the relation (3.9),

$$\begin{aligned}
&\frac{1}{(1+t)^{\beta+1}}\left(1-\frac{2t(x+1)}{(1+t)^2}\right)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x+1)}{(1+t)^2}}}\right)^{\beta}\\
&\quad=\sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{\left(\frac{\beta+1}{2}\right)_k\left(\frac{\beta}{2}\right)_k}{(\beta+1)_k}\frac{(2t(x+1))^k}{k!}\frac{(\beta+2k+1)_{n-k}(-t)^{n-k}}{(n-k)!}\\
&\quad=\sum_{n=0}^{\infty}\frac{(\beta+1)_n(-1)^n}{n!}{}_2F_1\left(\begin{matrix}-n,\beta+n+1\\ \beta+1\end{matrix};\frac{x+1}{2}\right)t^n. \qquad (3.11)
\end{aligned}$$

In a similar way, we obtain by means of the relation (3.4) and the binomial formula (3.8),

$$\frac{1}{(1-t)^{\alpha}}\left(\frac{2}{1+\sqrt{1-\frac{2t(x-1)}{(1-t)^2}}}\right)^{\alpha}=\sum_{n=0}^{\infty}\frac{(\alpha)_n}{n!}{}_2F_1\left(\begin{matrix}-n,\alpha+n\\ \alpha+1\end{matrix};\frac{1-x}{2}\right)t^n. \qquad (3.12)$$

Multiplying (3.11) and (3.12) we obtain, using (3.9) on the right hand side,

$$\begin{aligned}
G(x,t)=&\sum_{n=0}^{+\infty}\sum_{k=0}^{n}\frac{(\alpha)_{n-k}(\beta+1)_k(-1)^k}{(n-k)!k!}{}_2F_1\left(\begin{matrix}-n+k,\alpha+n-k\\ \alpha+1\end{matrix};\frac{1-x}{2}\right)\\
&\times{}_2F_1\left(\begin{matrix}-k,\beta+k+1\\ \beta+1\end{matrix};\frac{x+1}{2}\right)t^n.
\end{aligned}$$

Comparing coefficients of t^n in this result and in (3.10), we get

$$\begin{aligned}
P_n^{(\alpha,\beta)}(x)=&\sum_{k=0}^{n}\frac{(\alpha)_{n-k}(\beta+1)_k(-1)^k}{(n-k)!k!}{}_2F_1\left(\begin{matrix}-n+k,\alpha+n-k\\ \alpha+1\end{matrix};\frac{1-x}{2}\right)\\
&\times{}_2F_1\left(\begin{matrix}-k,\beta+k+1\\ \beta+1\end{matrix};\frac{x+1}{2}\right). \qquad (3.13)
\end{aligned}$$

Proposition 3.2 *The Jacobi polynomials (3.13) have the following hypergeometric representation*

$$P_n^{(\alpha,\beta)}(x)=\frac{(\alpha+1)_n}{n!}{}_2F_1\left(\begin{matrix}-n,\alpha+\beta+1+n\\ \alpha+1\end{matrix};\frac{1-x}{2}\right),n=0,1,2,\ldots.$$

Proof Since $P_n^{(\alpha,\beta)}(x)$ is a polynomial of degree n we obtain by means of Taylor's Theorem, $P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \frac{d^k}{dx^k} P_n^{(\alpha,\beta)}(1)(1-x)^k$. Substituting $x = 1$ into (3.13), we obtain

$$P_n^{(\alpha,\beta)}(1) = \sum_{k=0}^{n} \frac{(\alpha)_{n-k}(\beta+1)_k(-1)^k}{(n-k)!k!} {}_2F_1\left(\begin{matrix} -k, \beta+k+1 \\ \beta+1 \end{matrix}; 1\right).$$

The use of the Vandermonde or Chu-Vandermonde summation formula

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

with $n = k$, $b = \beta + k + 1$ and $c = \beta + 1$ yields, after simplification,

$$P_n^{(\alpha,\beta)}(1) = \sum_{k=0}^{n} \frac{(\alpha)_{n-k}(-k)_k(-1)^k}{(n-k)!k!}.$$

Using the *simplify* command of Maple (see [4]), we obtain

$$P_n^{(\alpha,\beta)}(1) = \frac{(\alpha+1)_n}{n!}. \tag{3.14}$$

From the formula (3.13), we get after straightforward computation

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

Iterating this formula, we obtain

$$\frac{d^k}{dx^k} P_n^{(\alpha,\beta)}(x) = \frac{(n+\alpha+\beta+1)_k}{2^k} P_{n-k}^{(\alpha+k,\beta+k)}(x).$$

Taking $x = 1$ and using formula (3.14) we get

$$\frac{d^k}{dx^k} P_n^{(\alpha,\beta)}(1) = \frac{(n+\alpha+\beta+1)_k}{2^k} \frac{(\alpha+k+1)_{n-k}}{(n-k)!}.$$

Therefore

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n}(-1)^k \frac{(n+\alpha+\beta+1)_k}{2^k}\frac{(\alpha+k+1)_{n-k}}{(n-k)!k!}(1-x)^k$$
$$= \frac{(\alpha+1)_n}{n!}\sum_{k=0}^{n}\frac{(-n)_k(\alpha+\beta+n+1)_k}{(\alpha+1)_k}\left(\frac{1-x}{2}\right)^k$$
$$= \frac{(\alpha+1)_n}{n!}\,{}_2F_1\left(\begin{matrix}-n,\alpha+\beta+n+1\\ \alpha+1\end{matrix};\frac{1-x}{2}\right).$$

4 Exercises

Exercise 4.1 Let

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

be a generating function of Hermite polynomials.

1. Prove that $\frac{\partial G(x,t)}{\partial x} = 2tG(x,t)$ and derive the differential difference relation

$$\frac{dH_n(x)}{dx} = 2nH_{n-1}(x).$$

2. Determine $\frac{\partial G(x,t)}{\partial t}$ and derive the three-term recurrence relation

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0,\ \ H_0(x) = 1.$$

3. Determine $H_1(x),\ H_2(x)$ and $H_3(x)$.

Exercise 4.2 The Gegenbauer polynomials $\{C_n^{(\nu)}(x)\}_{n=0}^{\infty}$, $\nu > -\frac{1}{2}$, $\nu \neq 0$, are orthogonal with respect to the weight function $\rho(x) = (1-x^2)^{\nu-\frac{1}{2}}$ on $(-1;1)$. These polynomials satisfy the Rodrigues formula

$$C_n^{(\nu)}(x) = \frac{(2\nu)_n(-1)^n}{(\nu+\frac{1}{2})_n 2^n n!\rho(x)}\frac{d^n}{dx^n}[(1-x^2)^n\rho(x)].$$

1. Prove that

$$(1-2xt+t^2)^{-\nu} = \sum_{n=0}^{\infty} C_n^{(\nu)}(x)t^n.$$

2. Derive the hypergeometric representation of $C_n^{(\nu)}(x)$.

Exercise 4.3 Let

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n} \frac{(\alpha)_{n-k}(\beta+1)_k(-1)^k}{(n-k)!k!} {}_2F_1\left(\begin{matrix} -n+k, \alpha+n-k \\ \alpha+1 \end{matrix}; \frac{1-x}{2}\right)$$
$$\times {}_2F_1\left(\begin{matrix} -k, \beta+k+1 \\ \beta+1 \end{matrix}; \frac{x+1}{2}\right), n = 0, 1, 2, \ldots$$

be the Jacobi polynomials.

1. Prove that $P_n^{(\alpha,\beta)}(-x) = (-1)^n P_n^{(\beta,\alpha)}(x)$.
2. Prove that $\frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{n+\alpha+\beta+1}{2} P_{n-1}^{(\alpha+1,\beta+1)}(x)$.

References

1. W.A. Al-Salam, Characterization theorems for orthogonal polynomials, in *Orthogonal Polynomials* (Springer, Dordrecht, 1990), pp. 1–24
2. R. Koekoek, P.A. Lesky, R.W. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues* (Springer Science and Business Media, Berlin, 2010)
3. E.B. McBride, *Obtaining Generating Functions* (Springer, Berlin, 1971)
4. M.B. Monagan, K.O. Geddes, K.M. Heal, G. Labahn, S.M. Vorkoetter, J. McCarron, P. DeMarco, *Maple 9 Introductory Programming Guide* (Maplesoft, Waterloo, 2003)
5. A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*, 1st edn. (Springer, Berlin, 1991)
6. E.D. Rainville, *Special Functions*, 1st edn. (The Macmillan Company, New York, 1960)
7. H.S. Wilf, *Generatingfunctionology*, 3rd edn. (Academic, Cambridge, 2005)

Properties and Applications of the Zeros of Classical Continuous Orthogonal Polynomials

A. S. Jooste

Abstract Suppose $\{P_n\}_{n=0}^{\infty}$ is a sequence of polynomials, orthogonal with respect to the weight function $w(x)$ on the interval $[a, b]$. In this lecture we will show that the zeros of an orthogonal polynomial are simple, that they are located in the interval of orthogonality and that the zeros of polynomials with adjacent degree, separate each other. We will also discuss the main ingredients of the Gauss quadrature formula, where the zeros of orthogonal polynomials are of decisive importance in approximating integrals.

Keywords Zeros · Orthogonality · Interlacing of zeros · Gauss quadrature

Mathematics Subject Classification (2000) 33C45

1 Introduction

We say the sequence of polynomials $\{P_n\}_{n=0}^{\infty}$, where $\deg(P_n(x)) = n$, is orthogonal with respect to the weight function $w(x) > 0$ on the interval (a, b), if

$$\int_a^b P_n(x)P_m(x)w(x)dx = h_n\delta_{mn}, h_n > 0, \tag{1.1}$$

where δ_{mn} is Kronecker's symbol,

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

A. S. Jooste (✉)
Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa
e-mail: alta.jooste@up.ac.za

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_4

Consider the Chebyshev polynomials of the first kind, defined by

$$T_n(x) = \cos(n \arccos x).$$

The sequence $\{T_n\}_{n\geq 0}$ is orthogonal on $(-1, 1)$ with respect to the weight function $w(x) = 1 - x^2$ and the zeros of $T_n(x)$ are

$$x_{n,k} = \cos\left(\frac{2(n-k)+1}{n}\right)\frac{\pi}{2},\ k = 1, 2, \ldots, n,$$

and it is clear that

$$-1 < x_{n,1} < x_{n,2} < \cdots < x_{n,n-1} < x_{n,n} < 1,$$

i.e., $T_n(x)$ has exactly n simple zeros in the interval $(-1, 1)$. In this lecture our focus is on the location and behavior of the zeros of the classical continuous orthogonal polynomials and we will also discuss how these zeros can be applied to approximate definite integrals. The work discussed here is known and is thoroughly discussed in, e.g., [2, 3, 6–8]. We will refer to the following, proved in the introductory lecture:

(1) Every sequence of orthogonal polynomials satisfies a three term recurrence equation of the form

$$P_n(x) = (A_n x + B_n)P_{n-1}(x) - C_n P_{n-2}(x), n \geq 1, \tag{1.2}$$

$P_{-1} \equiv 0$, A_n, B_n and C_n are real constants. If the highest coefficient of $P_n(x)$ is $k_n > 0$, then

$$A_n = \frac{k_n}{k_{n-1}}, C_n = \frac{A_n}{A_{n-1}} = \frac{k_n k_{n-2}}{k_{n-1}^2}.$$

(2) A sequence of orthogonal polynomials satisfies the Christoffel-Darboux formula:

$$\sum_{k=0}^{n} \frac{P_k(x)P_k(y)}{h_k} = \frac{k_n}{h_n k_{n+1}} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{x-y}, n = 0, 1, 2, \ldots, \tag{1.3}$$

and its confluent form

$$\sum_{k=0}^{n} \frac{P_k(x)^2}{h_k} = \frac{k_n}{h_n k_{n+1}}\Big(P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)\Big), n = 0, 1, 2, \ldots \tag{1.4}$$

2 The Location of Zeros of Orthogonal Polynomials

The zeros of every polynomial in an orthogonal sequence are real and simple and they lie in the interval of orthogonality.

Theorem 2.1 *If $\{P_n\}_{n=0}^{\infty}$ is a sequence of polynomials orthogonal with respect to the weight function $w(x)$ on (a, b), then the polynomial $P_n(x)$ has exactly n simple zeros in (a, b).*

Proof Since $\{P_n\}_{n=0}^{\infty}$ is a sequence of orthogonal polynomials and $\deg(P_n(x)) = n$, we know that P_n has at most n real zeros and (1.1) holds. In particular, for $n \geq 1$, taking $m = 0$, we see that

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

since $P_0(x)$ is constant. Now, $w(x) \geq 0$ for all $x \in [a, b]$ and $w(x) \neq 0$, so that, for this integral to be zero, $P_n(x)$ must have at least one zero of odd multiplicity in (a, b), i.e., $P_n(x)$ must change sign in (a, b). Let $x_1, x_2, \ldots, x_l$, $l \leq n$, be the distinct zeros (of odd multiplicity) of $P_n(x)$ in (a, b). We have just shown that $l \geq 1$. Define

$$Q(x) = (x - x_1)(x - x_2)\ldots(x - x_l),$$

i.e., $Q(x)$ is a polynomial of degree exactly l, having simple zeros (zeros with multiplicity 1) at the zeros of $P_n(x)$, which have odd multiplicity. The polynomial

$$P_n(x)Q(x) = P_n(x)(x - x_1)(x - x_2)\ldots(x - x_l)$$

does not change sign on (a, b), therefore

$$\int_a^b P_n(x)Q(x)w(x)dx \neq 0.$$

By orthogonality, this integral equals zero if $l < n$. Hence $l = n$, which implies that $P_n(x)$ has n real, distinct zeros of odd multiplicity in (a, b) and since $\deg(P_n(x)) = n$, the n zeros of $P_n(x)$ are distinct and simple. □

The zeros of different polynomials, e.g., P_n and $P_m, m \neq n$, in the same orthogonal sequence, are not just randomly positioned, but there is a specific order associated to the location of these zeros. As a start, two polynomials of adjacent degree cannot have any common zeros, and more-over, they separate each other. This will be shown in the following exercise and theorem. Furthermore, one can show that for $m < n - 1$, provided P_m and P_n have no common zeros, there exist m open intervals, with endpoints at successive zeros of P_n, each of which contains exactly one zero of P_m. This phenomenon is known as Stieltjes interlacing (cf. [8, Theorem 3.3.3] and [5]) and is discussed in detail in the lecture: Zeros of Orthogonal Polynomials, by Kerstin Jordaan.

Exercise Show that the polynomials P_n and P_{n+1} cannot have any zeros in common.

Consider the confluent form of the Christoffel-Darboux formula (1.4), from which we can deduce that

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

Suppose x^* is a common zero of P_n and P_{n+1}. By evaluating the above at x^*, we obtain

$$P'_{n+1}(x^*)P_n(x^*) - P_{n+1}(x^*)P'_n(x^*) = 0$$

and we have a contradiction.

Theorem 2.2 *If $\{P_n\}_{n=0}^{\infty}$ is a sequence of polynomials orthogonal with respect to the weight function $w(x)$ on (a, b), then the zeros of $P_n(x)$ and $P_{n+1}(x)$ separate each other.*

Proof Note that, from (1.1),

$$\int_a^b P_n(x)^2 w(x)dx = h_n > 0, n = 0, 1, 2, \ldots,$$

and it follows from (1.4), as well as our assumption that $k_n > 0$, that

$$P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x) > 0, n = 0, 1, 2, \ldots \tag{2.1}$$

Consider any $k \in \{1, 2, \ldots, n\}$. We now evaluate (2.1) at $x_{n+1,k}$ and $x_{n+1,k+1}$, two consecutive zeros of $P_{n+1}(x)$, to obtain

$$P'_{n+1}(x_{n+1,k})P_n(x_{n+1,k}) > 0 \text{ and } P'_{n+1}(x_{n+1,k+1})P_n(x_{n+1,k+1}) > 0,$$

respectively. Multiplying these two inequalities gives us

$$P_n(x_{n+1,k})P_n(x_{n+1,k+1})P'_{n+1}(x_{n+1,k})P'_{n+1}(x_{n+1,k+1}) > 0$$

and since the zeros of P_{n+1} are real and simple, it is clear that $P'_{n+1}(x_{n+1,k})P'_{n+1}(x_{n+1,k+1}) < 0$. Therefore

$$P_n(x_{n+1,k})P_n(x_{n+1,k+1}) < 0,$$

and this is true for each $k \in \{1, 2, \ldots, n\}$. P_n thus differs in sign at consecutive zeros of P_{n+1}, i.e., P_n has an odd number of zeros in each one of the n intervals

$(x_{n+1,k}, x_{n+1,k+1}), k \in \{1, 2, \ldots, n\}$. Since P_n has exactly n zeros, the result is proved. □

Exercise Consider the three term recurrence equation (1.2). What can we deduce concerning common zeros for P_n and P_{n-2}?

Suppose x^* is a common zero of P_n and P_{n-2}. We evaluate (1.2) at x^* to obtain

$$0 = (A_n x^* + B_n) P_{n-1}(x^*)$$

and since x^* cannot be a zero of P_{n-1}, it follows that $x^* = -\frac{B_n}{A_n}$.

Remark 2.3 If P_n and P_{n-2} do not have any common zeros, the n zeros of P_n interlace (in the Stieltjes sense) with the $(n-2)$ zeros of P_{n-2}. There are $(n-1)$ intervals $(x_{n,k}, x_{n,k+1}), k = 1, 2, \ldots, n-1$, with endpoints at the zeros of P_n and P_{n-2} has only $(n-2)$ zeros. In the single interval that does not contain a zero of P_{n-2} we will find the "extra-interlacing" point $x^* = -\frac{B_n}{A_n}$. See [4, 5] for interlacing results for the zeros of different sequences of Laguerre and Jacobi polynomials.

Example Consider the Jacobi polynomials:

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

orthogonal with respect to the weight function $w(x) = (1-x)^\alpha(1+x)^\beta$ on the interval $[-1, 1]$ when $\alpha, \beta > -1$. Use the program *Mathematica* to illustrate that the zeros lie in the interval of orthogonality (i.e., the interval in which the weight function is positive), as well as the result of Theorem 2.2. Note that restrictions on the parameters are necessary to ensure orthogonality and that the zeros are continuous functions of the parameters.

3 Gauss Quadrature

The Gauss quadrature formula is of use for the approximation of integrals in numerical analysis. If f is a continuous function in (a, b) and $x_1 < x_2 < \cdots < x_n$ are n distinct points in (a, b), then there exists exactly one polynomial L, with $\deg(L) \le n-1$ such that $L(x_j) = f(x_j)$ for all $j = 1, 2, \ldots, n$. This polynomial L can easily be found by using Lagrange interpolation. Define

$$p(x) = (x - x_1)(x - x_2)\ldots(x - x_n)$$

and consider the Lagrange interpolation polynomial

$$\begin{aligned} L(x) &= \sum_{k=1}^{n} f(x_k) \frac{p(x)}{(x - x_k) p'(x_k)} \\ &= \sum_{k=1}^{n} f(x_k) \frac{(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}. \end{aligned}$$

Let $\{P_n\}_{n=0}^{\infty}$ be a sequence of polynomials orthogonal on (a, b), with respect to the weight function $w(x)$, with n distinct real zeros $x_1 < x_2 < \dots < x_n$.

(a) If f is a polynomial of degree at most $2n - 1$, then $f(x) - L(x)$ is of degree $\leq 2n - 1$ with at least the zeros $x_1 < x_2 < \dots < x_n$. Now define

$$f(x) = L(x) + r(x) P_n(x),$$

where $r(x)$ is a polynomial of degree $\leq n - 1$. Then

$$f(x) = \sum_{k=1}^{n} f(x_k) \frac{P_n(x)}{(x - x_k) P_n'(x_k)} + r(x) P_n(x)$$

and

$$\int_a^b f(x)w(x)dx = \sum_{k=1}^{n} f(x_k) \int_a^b \frac{P_n(x)}{(x - x_k) P_n'(x_k)} w(x)dx + \int_a^b P_n(x)r(x)w(x)dx.$$

Since $\deg(r(x)) \leq n - 1$, the latter integral equals zero, due to orthogonality. We thus have

$$\int_a^b f(x)w(x)dx = \sum_{k=1}^{n} \lambda_{n,k} f(x_k)$$

with

$$\lambda_{n,k} := \int_a^b \frac{P_n(x)}{(x - x_k) P_n'(x_k)} w(x)dx, k = 1, 2, \dots, n.$$

(b) If f is not a polynomial of degree $\leq 2n - 1$, we can approximate the integral by

$$\int_a^b f(x)w(x)dx \approx \sum_{k=1}^{n} \lambda_{n,k} f(x_k) \tag{3.1}$$

with $\lambda_{n,k}$, $k = 1, 2, \dots, n$, defined as above.

The coefficients $\{\lambda_{n,k}\}_{k=1}^{n}$ are called the Christoffel numbers and they don't depend on the function f.

Exercise Show that the Christoffel numbers are positive.

We have

$$\lambda_{n,k} = \int_a^b l_{n,k}(x)w(x)dx, \text{ with } l_{n,k} = \frac{P_n(x)}{(x-x_k)P_n'(x_k)}, k = 1, 2, \ldots, n.$$

Then $l_{n,k}^2 - l_{n,k}$ is a polynomial, $\deg(l_{n,k}^2 - l_{n,k}) \leq 2n-2$ and it vanishes at the zeros of P_n, namely $x_1, x_2, \ldots, x_n$. We can write

$$l_{n,k}(x)^2 - l_{n,k}(x) = P_n(x)q(x),$$

where $q(x)$ is a polynomial of degree $\leq n-2$. Then

$$\int_a^b \big(l_{n,k}(x)^2 - l_{n,k}(x)\big)w(x)dx = \int_a^b P_n(x)q(x)w(x)dx = 0,$$

since $\deg(q(x)) \leq n-2$, consequently

$$\int_a^b l_{n,k}(x)^2 w(x)dx = \int_a^b l_{n,k}(x)w(x)dx > 0.$$

Example (cf. [9]) Consider the integral

$$\int_{-1}^{1} \frac{dx}{x+3} = ln2 = 0.69315.$$

Use the Gauss quadrature formula and a second degree Legendre polynomial to estimate this integral.

Consider

$$\int_{-1}^{1} \frac{dx}{x+3} = \int_a^b f(x)w(x)dx,$$

where $f(x) = \frac{1}{x+3}$, $a = -1$, $b = 1$, $w(x) = 1$. We will use the zeros of the second degree Legendre polynomial, orthogonal with respect to $w(x) = 1$ on $[-1, 1]$,

$$L_n(x) = {}_2F_1\left(\begin{matrix} -n, n+1 \\ 1 \end{matrix} \,\middle|\, \frac{1-x}{2}\right),$$

i.e.,

$$L_0(x) = 1, L_1(x) = x, L_2(x) = \frac{1}{2}\big(-1 + 3x^2\big)$$

and the zeros of $L_2(x)$ are $x_1 = \frac{1}{\sqrt{3}}$ and $x_2 = -\frac{1}{\sqrt{3}}$.

Now, using (3.1), we have

$$\int_{-1}^{1} \frac{dx}{x+3} \approx \lambda_{2,1} f(x_1) + \lambda_{2,2} f(x_2) = \frac{\lambda_{2,1}}{x_1 + 3} + \frac{\lambda_{2,2}}{x_2 + 3},$$

where

$$\lambda_{2,k} = \int_{-1}^{1} \frac{L_2(x)}{(x - x_k)L_2'(x_k)} w(x)dx = \int_{-1}^{1} \frac{\frac{1}{2}\big(-1 + 3x^2\big)}{(x - x_k)3x_k} dx, k = 1, 2,$$

i.e., $\lambda_{2,1} = \lambda_{2,2} = 1$ and

$$\int_{-1}^{1} \frac{dx}{x+3} = \frac{1}{\big(\frac{1}{\sqrt{3}} + 3\big)} + \frac{1}{\big(-\frac{1}{\sqrt{3}} + 3\big)} = 0.69231.$$

Remark 3.1 We use the Gauss quadrature rule to estimate integrals of the form $\int_a^b f(x)w(x)dx$. In the above example, i.e., for $a = -1, b = 1$ and $w(x) = 1$, a specific case (Gauss-Legendre quadrature) was used. Other choices of a, b and $w(x)$ lead to other integration rules and we refer the reader to [1, §25.4] where a summary of these rules is provided.

References

1. M. Abramovitz, I.A. Stegun (eds.), *Handbook of Mathematical Functions (with Formulas, Graphs and Mathematical Tables)* (Dover, Illinois, 1972)
2. G. Andrews, R. Askey, R. Roy, Special functions, in *Encyclopedia of Mathematics and Its Applications* (Cambridge University Press, Cambridge, 1999)
3. T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Dover Publications, New York, 2011)
4. K. Driver, K. Jordaan, Stieltjes interlacing of zeros of Laguerre polynomials from different sequences. Ind. Math. **21**, 204–211 (2011)
5. K. Driver, A. Jooste, K. Jordaan, Stieltjes interlacing of zeros of Jacobi polynomials from different sequences. Electron. Trans. Numer. Anal. **38**, 317–326 (2011)
6. M.E.H. Ismail, Classical and quantum orthogonal polynomials in one variable, in *Encyclopedia of Mathematics and Its Applications*, vol. 98 (Cambridge University Press, Cambridge, 2005)
7. E.D. Rainville, *Special Functions* (The Macmillan Company, New York, 1960)
8. G. Szegő, *Orthogonal Polynomials* (American Mathematical Society, New York, 1959)
9. N.M. Temme, *Special Functions, An Introduction to the Classical Functions of Mathematical Physics* (Wiley, New York, 1996)

Inversion, Multiplication and Connection Formulae of Classical Continuous Orthogonal Polynomials

Daniel Duviol Tcheutia

Abstract Our main objective is to establish the so-called *connection formula*,

$$p_n(x) = \sum_{k=0}^{n} C_k(n) y_k(x), \tag{0.1}$$

which for $p_n(x) = x^n$ is known as the *inversion formula*

$$x^n = \sum_{k=0}^{n} I_k(n) y_k(x),$$

for the family $y_k(x)$, where $\{p_n(x)\}_{n\in\mathbb{N}_0}$ and $\{y_n(x)\}_{n\in\mathbb{N}_0}$ are two polynomial systems. If we substitute x by ax in the left hand side of (0.1) and y_k by p_k, we get the *multiplication formula*

$$p_n(ax) = \sum_{k=0}^{n} D_k(n, a) p_k(x).$$

The coefficients $C_k(n)$, $I_k(n)$ and $D_k(n, a)$ exist and are unique since $\deg p_n = n$, $\deg y_k = k$ and the polynomials $\{p_k(x),\ k = 0, 1, \ldots, n\}$ or $\{y_k(x),\ k = 0, 1, \ldots, n\}$ are therefore linearly independent. In this session, we show how to use generating functions or the structure relations to compute the coefficients $C_k(n)$, $I_k(n)$ and $D_k(n, a)$ for classical continuous orthogonal polynomials.

Keywords Orthogonal polynomials · Inversion coefficients · Multiplication coefficients · Connection coefficients

D. D. Tcheutia (✉)
Institute of Mathematics, University of Kassel, Kassel, Germany
e-mail: duvtcheutia@yahoo.fr

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_5

Mathematics Subject Classification (2000) 33C45, 33D45, 33D15, 33F10, 68W30

1 The Classical Orthogonal Polynomials of a Continuous Variable and Their Generating Functions

We consider the following classical orthogonal polynomials of a continuous variable (see, e.g., [9, 16]):

1. The Jacobi polynomials

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

2. The Laguerre polynomials

$$L_n^{(\alpha)}(x) = \frac{(\alpha+1)_n}{n!}\,{}_1F_1\left(\left.\begin{matrix}-n\\ \alpha+1\end{matrix}\right| x\right),\ \alpha > -1,$$

3. The Hermite polynomials

$$H_n(x) = 2^n x^n\,{}_2F_0\left(\left.\begin{matrix}-n/2, -n/2+1/2\\ -\end{matrix}\right| -\frac{1}{x^2}\right),$$

4. The Bessel polynomials

$$y_n(x;\alpha) = {}_2F_0\left(\left.\begin{matrix}-n, n+\alpha+1\\ -\end{matrix}\right| -\frac{x}{2}\right),\ n = 0, 1, \ldots, N,\ \alpha < -2N-1.$$

$P_n^{(\alpha,\beta)}(x)$, $L_n^{(\alpha)}(x)$, $H_n(x)$, or $y_n(x;\alpha)$ is exactly a polynomial of degree n with respect to the continuous variable x. Generating functions [16] for the above polynomials are given for:

1. the Jacobi polynomials by

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

2. the Laguerre polynomials by

$$(1-t)^{-\alpha-1}\exp\left(\frac{xt}{t-1}\right)=\sum_{n=0}^{\infty}L_n^{(\alpha)}t^n,$$

3. the Hermite polynomials by

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

4. the Bessel polynomials by

$$(1-2xt)^{-\frac{1}{2}}\left(\frac{2}{1+\sqrt{1-2xt}}\right)^{\alpha}\exp\frac{2t}{1+\sqrt{1-2xt}}=\sum_{n=0}^{\infty}y_n(x;\alpha)\frac{t^n}{n!}.$$

2 Inversion Problem Using Generating Functions

From the definition of the Laguerre polynomials, we have

$$L_0^{(\alpha)}(x)=1\Rightarrow x^0=1=L_0^{(\alpha)}(x);$$

$$L_1^{(\alpha)}(x)=-x+\alpha+1\Rightarrow x=-L_1^{(\alpha)}(x)+(\alpha+1)L_0^{(\alpha)}(x);$$

$$\begin{aligned}L_2^{(\alpha)}(x)&=\frac{1}{2}x^2-(\alpha+2)x+\frac{1}{2}(\alpha+1)(\alpha+2)\\ &\Rightarrow x^2=2L_2^{(\alpha)}(x)-(2\alpha+4)L_1^{(\alpha)}(x)+(\alpha+1)(\alpha+2)L_0^{(\alpha)}(x).\end{aligned}$$

In general, since $\deg L_n^{(\alpha)}(x)=n$, the system $\{L_k^{(\alpha)}(x), k=0,1,2,\ldots,n\}$ is a basis of polynomials of degree at most n. Therefore, there exist coefficients $I_k(n)$ such that the inversion formula

$$x^n=\sum_{k=0}^{n}I_k(n)L_k^{(\alpha)}(x)$$

is valid. How can we compute the inversion coefficients $I_k(n)$?
One option is by using a generating function of the given polynomials (see, e.g., [16]).

A generating function of the Laguerre polynomials is given by

$$e^t {}_0F_1\left(\begin{matrix} - \\ \alpha+1 \end{matrix}\,\middle|\, -xt\right) = \sum_{n=0}^{\infty} \frac{L_n^{(\alpha)}(x)}{(\alpha+1)_n} t^n,$$

which is equivalent to

$${}_0F_1\left(\begin{matrix} - \\ \alpha+1 \end{matrix}\,\middle|\, -xt\right) = e^{-t} \sum_{n=0}^{\infty} \frac{L_n^{(\alpha)}(x)}{(\alpha+1)_n} t^n.$$

Using the series expansion of the exponential function, this yields

$$\sum_{n=0}^{\infty} \frac{(-xt)^n}{(\alpha+1)_n n!} = \left(\sum_{n=0}^{\infty} \frac{(-t)^n}{n!}\right)\left(\sum_{k=0}^{\infty} \frac{L_k^{(\alpha)}(x)}{(\alpha+1)_k} t^k\right).$$

From the relation (cf. [16])

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} A(k,n) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} A(k,n-k),$$

we get

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{(-1)^n x^n t^n}{(\alpha+1)_n n!} &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{(-1)^{n-k} t^{n-k}}{(n-k)!} \frac{L_k^{(\alpha)}(x)}{(\alpha+1)_k} t^k \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} \frac{(-1)^{n-k} L_k^{(\alpha)}(x)}{(n-k)!(\alpha+1)_k}\right) t^n.\end{aligned}$$

Equating the coefficients of t^n yields

$$x^n = \sum_{k=0}^{n} \frac{(-1)^k n! (\alpha+1)_n}{(n-k)!(\alpha+1)_k} L_k^{(\alpha)}(x).$$

Exercise Use the generating function of the Hermite polynomials and the relation (cf. [16])

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} A(k,n) = \sum_{n=0}^{\infty}\sum_{k=0}^{\lfloor n/2 \rfloor} A(k,n-2k),$$

where $\lfloor n/2 \rfloor$ is the floor of $n/2$, to show that the inversion formula

$$x^n = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n! H_{n-2k}(x)}{2^n k!(n-2k)!}$$

is valid for the Hermite polynomials.

3 Multiplication Formula Using Generating Functions (See e.g. [4, 16, 20])

We consider the following generating function of the Laguerre polynomials

$$G(x,t) := e^t {}_0F_1\left(\left.\begin{matrix} - \\ \alpha+1 \end{matrix}\right| -xt\right) = \sum_{n=0}^{\infty} \frac{L_n^{(\alpha)}(x)}{(\alpha+1)_n} t^n.$$

From this definition, we get $G(ax,t) = e^{t(1-a)} G(x,at)$ or equivalently

$$\sum_{n=0}^{\infty} \frac{L_n^{(\alpha)}(ax)}{(\alpha+1)_n} t^n = \left(\sum_{n=0}^{\infty} \frac{(1-a)^n t^n}{n!}\right)\left(\sum_{k=0}^{\infty} \frac{L_k^{(\alpha)}(x)}{(\alpha+1)_k} a^k t^k\right).$$

It follows from the Cauchy product that

$$\sum_{n=0}^{\infty} \frac{L_n^{(\alpha)}(ax)}{(\alpha+1)_n} t^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{a^k (1-a)^{n-k} L_k^{(\alpha)}(x)}{(\alpha+1)_k (n-k)!} t^n.$$

Equating the coefficients of t^n yields the multiplication formula of the laguerre polynomials

$$L_n^{(\alpha)}(ax) = \sum_{k=0}^{n} \frac{a^k (\alpha+1)_n (1-a)^{n-k}}{(\alpha+1)_k (n-k)!} L_k^{(\alpha)}(x).$$

Exercise Show that for the Hermite polynomials, the generating function $G(t,x) = \exp(2xt - t^2)$ satisfies

$$G(t,ax) = e^{(a^2-1)t^2} G(at,x).$$

Deduce that the multiplication formula

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

is valid.

4 Connection Formula Using Generating Functions (See e.g. [16])

If we rather consider the generating function

$$G(\alpha, t) := (1-t)^{-\alpha-1} \exp\left(\frac{xt}{t-1}\right) = \sum_{n=0}^{\infty} L_n^{(\alpha)}(x) t^n$$

of the Laguerre polynomials, we have the relation $G(\alpha, t) = (1-t)^{\alpha-\beta} G(\beta, t)$. This is equivalent to

$$\begin{aligned}
\sum_{n=0}^{\infty} L_n^{(\alpha)}(x) t^n &= \left(\sum_{n=0}^{\infty} \frac{(\alpha-\beta)_n}{n!} t^n\right) \left(\sum_{k=0}^{\infty} L_k^{(\beta)}(x) t^k\right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} \frac{(\alpha-\beta)_{n-m}}{(n-m)!} L_m^{(\beta)}(x)\right) t^n.
\end{aligned}$$

Equating the coefficients of t^n, we deduce the connection formula

$$L_n^{(\alpha)}(x) = \sum_{m=0}^{n} \frac{(\alpha-\beta)_{n-m}}{(n-m)!} L_m^{(\beta)}(x),$$

of the Laguerre polynomials. One application of the latter formula is the so-called parameter derivative of $L_n^{(\alpha)}(x)$ given by [11]

$$\frac{\partial}{\partial \alpha} L_n^{(\alpha)}(x) = \sum_{m=0}^{n-1} \frac{1}{n-m} L_m^{(\alpha)}(x).$$

To get this result knowing the connection relation

$$L_n^{(\alpha)}(x) = \sum_{m=0}^{n} C_m(n; \alpha, \beta) L_m^{(\beta)}(x),$$

we build the difference quotient

$$\frac{L_n^{(\alpha)}(x) - L_n^{(\beta)}(x)}{\alpha - \beta} = \sum_{m=0}^{n} \frac{C_m(n;\alpha,\beta)}{\alpha-\beta} L_m^{(\beta)}(x) - \frac{L_n^{(\beta)}(x)}{\alpha-\beta}$$
$$= \frac{C_n(n;\alpha,\beta) - 1}{\alpha-\beta} L_n^{(\beta)}(x) + \sum_{m=0}^{n-1} \frac{C_m(n;\alpha,\beta)}{\alpha-\beta} L_m^{(\beta)}(x)$$

so that with $\beta \to \alpha$

$$\frac{\partial}{\partial\alpha} L_n^{(\alpha)}(x) = \lim_{\beta\to\alpha} \frac{C_n(n;\alpha,\beta)-1}{\alpha-\beta} L_n^{(\beta)}(x) + \sum_{m=0}^{n-1} \lim_{\beta\to\alpha} \frac{C_m(n;\alpha,\beta)}{\alpha-\beta} L_m^{(\beta)}(x)$$

since the systems $L_n^{(\alpha)}(x)$ are continuous with respect to α. This gives the results.

5 Structure Relations and Applications

Every classical orthogonal polynomial sequence $\{p_n(x) = k_n x^n + \ldots\}_{n\geq 0}$ is solution of a second-order differential equation of type

$$\sigma(x)y''(x) + \tau(x)y'(x) + \lambda_n y(x) = 0,$$

where $\sigma(x) = ax^2 + bx + c$, $\tau(x) = dx + e$, $d \neq 0$, $\lambda_n = -n((n-1)a + d)$. Furthermore, a three-term recurrence relation of type

$$x p_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + c_n p_{n-1}(x), \tag{5.1}$$

is satisfied by $p_n(x)$, with [11]

$$a_n = \frac{k_n}{k_{n+1}},$$
$$b_n = -\frac{2bn(an+d-a) - e(2a-d)}{(d+2an)(d-2a+2an)},$$
$$c_n = -(n(an+d-2a)(4ac-b^2) + 4a^2c - ab^2 + ae^2 - 4acd + db^2 - bed + d^2c)$$
$$\times \frac{(an+d-2a)n}{(d-2a+2an)^2(2an-3a+d)(2an-a+d)} \frac{k_n}{k_{n-1}}.$$

Since the sequence of the derivatives $\{p_n'(x)\}_{n\geq 1}$ of $\{p_n(x)\}_{n\geq 0}$ is also an orthogonal polynomial sequence, it also satisfies a three-term recurrence relation of type

$$x p_n'(x) = \alpha_n p_{n+1}'(x) + \beta_n p_n'(x) + \gamma_n p_{n-1}'(x), \tag{5.2}$$

with [11]

$$\alpha_n = \frac{n}{n+1}\frac{k_n}{k_{n+1}},$$
$$\beta_n = \frac{-2bn(an+d-a)+d(b-e)}{(d+2an)(d-2a+2an)},$$
$$\gamma_n = -\frac{n((n-1)(an+d-a)(4ac-b^2)+ae^2+d^2c-bed)(an+d-a)}{(d-2a+2an)^2(2an-3a+d)(2an-a+d)}\frac{k_n}{k_{n-1}}.$$

The data corresponding to each classical continuous family are given in the following table:

System	$P_n^{(\alpha,\beta)}(x)$	$L_n^{(\alpha)}(x)$	$H_n(x)$	$y_n(x;\alpha)$
$\sigma(x)$	$1-x^2$	x	1	x^2
$\tau(x)$	$\beta-\alpha-(\alpha+\beta+2)x$	$\alpha+1-x$	$-2x$	$2+(\alpha+2)x$
k_n	$\frac{(\alpha+\beta+n+1)_n}{2^n n!}$	$\frac{(-1)^n}{n!}$	2^n	$\frac{(n+\alpha+1)_n}{2^n}$

We set $x^n = v_n(x)$. Therefore, we have

$$x v_n(x) = v_{n+1}(x),\ \ x v_n'(x) = \frac{n}{n+1} v_{n+1}'(x).$$

We suppose that

$$x^n = v_n(x) = \sum_{m=0}^{n} I_m(n) p_m(x), \tag{5.3}$$

which means that the coefficients $I_m(n) = 0$ for $m \neq 0, 1, \ldots, n$. The idea is to find a recurrence equation satisfied by the inversion coefficients $I_m(n)$ [1, 11, 13, 19, 22] and solve the obtained recurrence equation using Petkovšek-van-Hoeij algorithm to get its hypergeometric term solutions.

We substitute (5.3) in $x v_n(x) = v_{n+1}(x)$ to get

$$\sum_{m=0}^{n} I_m(n) x p_m(x) = \sum_{m=0}^{n+1} I_m(n+1) p_m(x).$$

Using the three-term recurrence relation (5.1), it follows that

$$\sum_{m=0}^{n} I_m(n)\Big(a_m p_{m+1}(x) + b_m p_m(x) + c_m p_{m-1}(x)\Big) = \sum_{m=0}^{n+1} I_m(n+1) p_m(x).$$

After a shift of index we get

$$\sum_{m=0}^{n+1}\Big(a_{m-1}I_{m-1}(n)+b_mI_m(n)+c_{m+1}I_{m+1}(n)\Big)p_m(x)=\sum_{m=0}^{n+1}I_m(n+1)p_m(x).$$

Equating the coefficients of $p_m(x)$ yields a mixed recurrence relation with respect to m and n

$$a_{m-1}I_{m-1}(n)+b_mI_m(n)+c_{m+1}I_{m+1}(n)=I_m(n+1). \tag{5.4}$$

Similarly, we substitute (5.3) in $xv_n'(x)=\frac{n}{n+1}v_{n+1}'(x)$ and use (5.2) to get, after a shift of index,

$$\sum_{m=0}^{n+1}\Big(\alpha_{m-1}I_{m-1}(n)+\beta_mI_m(n)+\gamma_{m+1}I_{m+1}(n)\Big)p_m'(x)=\frac{n}{n+1}\sum_{m=0}^{n+1}I_m(n+1)p_m'(x).$$

By equating the coefficients of $p_m'(x)$, we get a mixed recurrence relation in the variables m and n

$$\alpha_{m-1}I_{m-1}(n)+\beta_mI_m(n)+\gamma_{m+1}I_{m+1}(n)=\frac{n}{n+1}I_m(n+1). \tag{5.5}$$

Combining (5.4) and (5.5), we get out with a recurrence equation with respect to m

$$\begin{aligned}\alpha_{m-1}I_{m-1}(n)+\beta_mI_m(n)+\gamma_{m+1}I_{m+1}(n)&=\\ &\frac{n}{n+1}\Big(a_{m-1}I_{m-1}(n)+b_mI_m(n)+c_{m+1}I_{m+1}(n)\Big),\end{aligned}$$

that is

$$\begin{aligned}&\Big(\alpha_{m-1}-\frac{n}{n+1}a_{m-1}\Big)I_{m-1}(n)+\Big(\beta_m-\frac{n}{n+1}b_m\Big)I_m(n)\\ &+\Big(\gamma_{m+1}-\frac{n}{n+1}c_{m+1}\Big)I_{m+1}(n)=0.\end{aligned} \tag{5.6}$$

Exercise We consider the Jacobi polynomials bases $v_n(x)=(x+1)^n$.

1. Show that

$$xv_n(x)=v_{n+1}(x)-v_n(x),\ \ xv_n'(x)=\frac{n}{n+1}v_{n+1}'(x)-v_n'(x).$$

2. We suppose that

$$(x+1)^n = v_n(x) = \sum_{m=0}^{n} I_m(n) p_m(x).$$

Show that $I_m(n)$ is solution of the recurrence relation

$$\Big(\alpha_{m-1} - \frac{n}{n+1} a_{m-1}\Big) I_{m-1}(n) + \Big(\beta_m - \frac{n}{n+1}(b_m+1)+1\Big) I_m(n) \\ + \Big(\gamma_{m+1} - \frac{n}{n+1} c_{m+1}\Big) I_{m+1}(n) = 0.$$

The next step is to substitute a, b, c, d, e for each family and solve the recurrence relation to get the inversion coefficients.

The coefficients $I_m(n)$ of the inversion formula

$$(1+x)^n = \sum_{m=0}^{n} I_m(n) P_m^{(\alpha,\beta)}(x) \tag{5.7}$$

of the Jacobi polynomials are solutions of the recurrence equation

$$\begin{aligned} &2\ (m+1+\alpha+\beta)(\alpha+\beta+m)(2m+3+\alpha+\beta)(2m+\alpha+\beta+2)(m-n-1) \\ &\times I_{m-1}(n) + 2\ (2m+3+\alpha+\beta)(m+1+\alpha+\beta)(\alpha+\beta+2m-1) \\ &\times \Big(\alpha\,\beta + 2m\alpha - n\alpha + \beta^2 + 2m\beta + n\beta + 2m^2 + 2\beta + 2m\Big) I_m(n) \\ &+ 2\ (\beta+m+1)(1+m+\alpha)(\alpha+\beta+2m-1)(2m+\alpha+\beta) \\ &\times (\alpha+\beta+m+n+2) I_{m+1}(n) = 0. \end{aligned}$$

The coefficients $I_m(n)$ of the inversion formula

$$x^n = \sum_{m=0}^{n} I_m(n) L_m^{(\alpha)}(x) \tag{5.8}$$

of the Laguerre polynomials are solutions of the recurrence equation

$$(-m+n+1)\, I_{m-1}(n) + (\alpha+2m-n)\, I_m(n) + (-1-m-\alpha)\, I_{m+1}(n) = 0.$$

The coefficients $I_m(n)$ of the inversion formula

$$x^n = \sum_{m=0}^{n} I_m(n) H_m(x) \tag{5.9}$$

of the Hermite polynomials are solutions of the recurrence equation

$$(m-n-1)\,I_{m-1}\,(n)+2\,m\,(m+1)\,I_{m+1}\,(n)=0.$$

The coefficients $I_m(n)$ of the inversion formula

$$x^n=\sum_{m=0}^{n} I_m(n)y_m(x;\alpha) \tag{5.10}$$

of the Bessel polynomials are solutions of the recurrence equation

$$\begin{aligned}
&2\,(m+\alpha)\,(2\,m+3+\alpha)\,(2+2\,m+\alpha)\,(1+m+\alpha)\,(m-n-1)\,I_{m-1}\,(n)\\
&-2\,m\,(2\,m+3+\alpha)\,(2\,m-1+\alpha)\,(1+m+\alpha)\,(\alpha+2+2\,n)\,I_m\,(n)\\
&-2\,m\,(m+1)\,(2\,m-1+\alpha)\,(\alpha+2\,m)\,(\alpha+m+n+2)\,I_{m+1}\,(n)=0.
\end{aligned}$$

To solve the above recurrence equations, we can use the Petkovšek-van-Hoeij algorithm implemented in Maple by the command `'LREtools/hypergeomsols'(rec, R(m), {}, output=basis)` [5, 10, 14, 21]. The solution is given up to a multiplicative constant:

- for the Jacobi polynomials by

$$I_m(n)=\frac{(-1)^m\,\Gamma\,(m+1+\alpha+\beta)\,\Gamma\,(m-n)\,(m+1/2+\beta/2+\alpha/2)}{\Gamma\,(\alpha+\beta+m+n+2)\,\Gamma\,(\beta+m+1)},$$

- for the Laguerre polynomials by

$$I_m(n)=\frac{\Gamma\,(m-n)}{\Gamma\,(1+m+\alpha)},$$

- for the Bessel polynomials by

$$I_m(n)=\frac{\Gamma\,(m-n)\,\Gamma\,(1+m+\alpha)\,(\alpha/2+m+1/2)}{\Gamma\,(m+1)\,\Gamma\,(\alpha+m+n+2)}.$$

This means for example that for the Bessel polynomials

$$x^n=\sum_{m=0}^{n}\frac{\Gamma\,(m-n)\,\Gamma\,(1+m+\alpha)\,(\alpha/2+m+1/2)}{\Gamma\,(m+1)\,\Gamma\,(\alpha+m+n+2)}\times constant\times y_m(x;\alpha).$$

To get the constant, we equate the coefficients of x^n in both sides of (5.10) (noting that $y_n(x;\alpha)=\frac{(n+\alpha+1)_n}{2^n}x^n+\ldots$ and $\Gamma(m-n)=(-n)_m\Gamma(-n)$) to get

$$constant=\frac{2(-2)^n}{\Gamma(-n)}.$$

This leads to the inversion formula

$$x^n = (-2)^n \sum_{m=0}^{n} (2m+\alpha+1) \frac{(-n)_m \Gamma(\alpha+m+1)}{m!\Gamma(n+m+\alpha+2)} y_m(x;\alpha).$$

Exercise Show that the following inversion formula is valid:
for the Jacobi polynomials

$$(1+x)^n = 2^n \Gamma(\beta+n+1) \sum_{m=0}^{n} (-1)^m (-n)_m \frac{(\alpha+\beta+2m+1)\Gamma(\alpha+\beta+m+1)}{\Gamma(\beta+m+1)\Gamma(\alpha+\beta+n+m+2)} P_m^{(\alpha,\beta)}(x),$$

for the Laguerre polynomials

$$x^n = (1+\alpha)_n \sum_{m=0}^{n} \frac{(-n)_m}{(1+\alpha)_m} L_m^{(\alpha)}(x),$$

for the Hermite polynomials

$$x^n = \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-\frac{n}{2})_m (-\frac{n}{2}+\frac{1}{2})_m}{m! 2^{n-2m}} H_{n-2m}(x) = \frac{n!}{2^n} \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{m!(n-2m)!} H_{n-2m}(x).$$

In general, to find the coefficients $C_m(n)$ in the connection formula [1–3, 6–9, 11, 12, 16–19]

$$p_n(x) = \sum_{m=0}^{n} C_m(n) q_m(x),$$

we combine

$$p_n(x) = \sum_{j=0}^{n} A_j(n) x^j \text{ and } x^j = \sum_{m=0}^{j} I_m(j) q_m(x),$$

which yields the representation

$$p_n(x) = \sum_{j=0}^{n} \sum_{m=0}^{j} A_j(n) I_m(j) q_m(x),$$

and then, interchanging the order of summation gives

$$C_m(n) = \sum_{j=0}^{n-m} A_{j+m}(n) I_m(j+m),$$

or

$$p_n(x) = \sum_{m=0}^{n} C_{n-m}(n) q_{n-m}(x).$$

For orthogonal polynomials with even weight such as the Hermite and Gegenbauer polynomials, we have the relations

$$p_n(x) = \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} A_j(n) x^{n-2j} \text{ and } x^j = \sum_{m=0}^{\lfloor \frac{j}{2} \rfloor} I_m(j) q_{j-2m}(x),$$

from which we deduce

$$x^{n-2j} = \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor - j} I_m(n-2j) q_{n-2j-2m}(x).$$

Finally, we combine the above two expressions and substitute m by $m - j$ to get

$$C_m(n) = \sum_{j=0}^{m} A_j(n) I_{m-j}(n-2j),$$

with

$$p_n(x) = \sum_{m=0}^{n} C_m(n) q_{n-2m}(x).$$

Since the summand $F(j, m, n) := A_j(n) I_m(j)$ of $C_m(n)$ turns out to be a hypergeometric term with respect to (j, m, n), i.e., the term ratios $F(j+1, m, n)/F(j, m, n)$, $F(j, m+1, n)/F(j, m, n)$, and $F(j, m, n+1)/F(j, m, n)$ are rational functions, Zeilberger's (combined with the Petkovšek-van-Hoeij) algorithm applies [5, 10, 14, 15, 21]. If a hypergeometric term solution exists, the representation of $C_m(n)$ follows then from the initial values $C_n(n) = k_n/\bar{k}_n$, $C_{n+s}(n) = 0$, $s = 1, 2, \ldots$, where k_n, $\bar{k}_n$ are, respectively, the leading coefficients of $p_n(x)$ and $q_n(x)$.

The following connection relations between classical orthogonal polynomials are valid:

$$P_n^{(\alpha,\beta)}(x) = \sum_{m=0}^{n} (2m+\gamma+\beta+1) \frac{\Gamma(n+\beta+1)\Gamma(n+m+\alpha+\beta+1)}{\Gamma(m+\beta+1)\Gamma(n+\alpha+\beta+1)} \times \frac{\Gamma(m+\gamma+\beta+1)(\alpha-\gamma)_{n-m}}{\Gamma(n+m+\gamma+\beta+2)(n-m)!} P_m^{(\gamma,\beta)}(x),$$

$$P_n^{(\alpha,\beta)} = \sum_{m=0}^{n}(-1)^{n-m}(2m+\alpha+\delta+1)\frac{\Gamma(n+\alpha+1)\Gamma(n+m+\alpha+\beta+1)}{\Gamma(m+\alpha+1)\Gamma(n+\alpha+\beta+1)}$$
$$\times\frac{\Gamma(m+\alpha+\delta+1)(\beta-\delta)_{n-m}}{\Gamma(n+m+\alpha+\delta+2)(n-m)!}P_m^{(\alpha,\delta)}(x),$$

$$P_n^{(\alpha,\beta)}(x) = \sum_{m=0}^{n}\frac{(m+\alpha+1)_{n-m}(n+\alpha+\beta+1)_m}{(n-m)!(m+\gamma+\delta+1)_m}$$
$$\times {}_3F_2\left(\left.\begin{matrix} m-n, n+m+\alpha+\beta+1, m+\gamma+1 \\ m+\alpha+1, 2m+\gamma+\delta+2 \end{matrix}\right| 1\right)P_m^{(\gamma,\delta)}(x),$$

$$L_n^{(\alpha)}(x) = \sum_{m=0}^{n}\frac{(\alpha-\beta)_{n-m}}{(n-m)!}L_m^{(\beta)}(x),$$

$$y_n(x;\alpha) = \sum_{m=0}^{n}(-1)^m(2m+\beta+1)$$
$$\times\frac{(-n)_m(n+\alpha+1)_m\Gamma(m+\beta+1)\Gamma(\beta-\alpha+1)}{m!\Gamma(n+m+\beta+2)\Gamma(m-n+\beta-\alpha+1)}y_m(x;\beta).$$

The following multiplication formulas of orthogonal polynomials of a continuous variable are valid:

$$P_n^{(\alpha,\beta)}(ax) = \sum_{m=0}^{n}\sum_{j=0}^{n-m}\frac{(-a)^m(1-a)^j(-n)_{m+j}(\alpha+1)_n(n+\alpha+\beta+1)_{m+j}}{2^j n! j!(\alpha+1)_{m+j}(\alpha+\beta+m+1)_m}$$
$$\times {}_2F_1\left(\left.\begin{matrix} \alpha+m+1, -j \\ \alpha+\beta+2m+2 \end{matrix}\right| \frac{2a}{a-1}\right)P_m^{(\alpha,\beta)}(x),$$

$$L_n^{(\alpha)}(ax) = \sum_{m=0}^{n}\frac{(\alpha+1)_n a^m(1-a)^{n-m}}{(n-m)!(\alpha+1)_m}L_m^{(\alpha)}(x),$$

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

$$y_n(ax;\alpha) = \sum_{m=0}^{n}\frac{(-a)^m(-n)_m(\alpha+n+1)_m}{m!(\alpha+m+1)_m}$$
$$\times {}_2F_1\left(\left.\begin{matrix} m-n, \alpha+m+n+1 \\ \alpha+2m+2 \end{matrix}\right| a\right)y_m(x;\alpha).$$

Acknowledgements Many thanks to the organizers of the AIMS–Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications, Douala, October 5–12, 2018.

References

1. I. Area, E. Godoy, A. Ronveaux, A. Zarzo, Solving connection and linearization problems within the Askey scheme and its q-analogue via inversion formulas. J. Comput. Appl. Math. **136**, 152–162 (2001)
2. R. Askey, G. Gasper, Jacobi polynomial expansions of Jacobi polynomials with nonnegative coefficients. Proc. Camb. Philos. Soc. **70**, 243–255 (1971)
3. R. Askey, G. Gasper, Convolution structures for Laguerre polynomials. J. Anal. Math. **31**, 48–68 (1977)
4. H. Chaggara, W. Koepf, Duplication coefficients via generating functions. Complex Var. Elliptic Equ. **52**, 537–549 (2007)
5. T. Cluzeau, M. van Hoeij, Computing hypergeometric solutions of linear recurrence equations. Appl. Algebra Eng. Commun. Comput. **17**, 83–115 (2006)
6. E.H. Doha, H.M. Ahmed, Recurrences and explicit formulae for the expansion and connection coefficients in series of Bessel polynomials. J. Phys. A **37**, 8045–8063 (2004)
7. J.L. Fields, J. Wimp, Expansions of hypergeometric functions in hypergeometric functions. Math. Comp. **15**, 390–395 (1961)
8. E. Godoy, A. Ronveaux, A. Zarzo, I. Area, Minimal recurrence relations for connection coefficients between classical orthogonal polynomials: continuous case. J. Comput. Appl. Math. **84**, 257–275 (1997)
9. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
10. W. Koepf, *Hypergeometric Summation—An Algorithmic Approach to Summation and Special Function Identities*, 2nd edn. (Springer Universitext, Springer, London, 2014)
11. W. Koepf, D. Schmersau, Representations of orthogonal polynomials. J. Comput. Appl. Math. **90**, 57–94 (1998)
12. S. Lewanowicz, The hypergeometric functions approach to the connection problem for the classical orthogonal polynomials. Technical Report, Institute of Computer Science, University of Wroclaw (2003)
13. P. Njionou Sadjang, Moments of classical orthogonal polynomials, Ph.D. Dissertation, Universität Kassel (2013)
14. M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients. J. Symb. Comput. **14**, 243–264 (1992)
15. M. Petkovšek, H. Wilf, D. Zeilberger, $A = B$ (A. K. Peters, Wellesley, 1996)
16. E.D. Rainville, *Special Functions* (The Macmillan Company, New York, 1960)
17. A. Ronveaux, A. Zarzo, E. Godoy, Recurrence relations for connection between two families of orthogonal polynomials. J. Comput. Appl. Math. **62**, 67–73 (1995)
18. J. Sánchez-Ruiz, J.S. Dehesa, Expansions in series of orthogonal hypergeometric polynomials. J. Comput. Appl. Math. **89**, 155–170 (1997)
19. D.D. Tcheutia, On connection, linearization and duplication coefficients of classical orthogonal polynomials, Ph.D. Dissertation, Universität Kassel (2014)
20. D.D. Tcheutia, M. Foupouagnigni, W. Koepf, P. Njionou Sadjang, Coefficients of multiplication formulas for classical orthogonal polynomials. Ramanujan J. **39**, 497–531 (2016)
21. M. van Hoeij, Finite singularities and hypergeometric solutions of linear recurrence equations. J. Pure Appl. Algebra **139**, 109–131 (1999)
22. A. Zarzo, I. Area, E. Godoy, A. Ronveaux, Results for some inversion problems for classical continuous and discrete orthogonal polynomials. J. Phys. A **30**, 35–40 (1997)

Classical Orthogonal Polynomials of a Discrete and a q-Discrete Variable

Patrick Njionou Sadjang

Abstract The classical orthogonal polynomials of discrete and q-discrete orthogonal polynomials are introduced from their difference and q-difference equations. Some structure formulas are proved for the Charlier and the Al-Salam Carlitz polynomials from their generating functions.

Keywords Orthogonal polynomials · Generating function · Inversion formula · Connection formula · Addition formula · Multiplication formula

Mathematics Subject Classification (2000) 33C45, 33D45, 33D15, 33F10, 68W30

1 Introduction

Let $\mathcal{P}$ be the linear space of polynomials with complex coefficients. A polynomial sequence $\{P_n\}_{n\geq 0}$ in $\mathcal{P}$ is called a polynomial set if and only if $\deg P_n = n$ for all nonnegative integers n.

Let α denote a nondecreasing function with a finite or an infinite number of points of increase in the interval $(a; b)$. The latter interval may be infinite. We assume that the numbers μ_n defined by

$$\mu_n = \int_a^b x^n d\alpha(x) \tag{1.1}$$

exist for $n = 0, 1, 2, \ldots$. These numbers are called *canonical moments of the measure* $d\alpha(x)$. The integral (1.1) can be considered as a Riemann-Stieltjes integral

P. Njionou Sadjang (✉)
University of Douala, Faculty of Industrial Engineering, Douala, Cameroon
e-mail: patrick.njionou@aims-cameroon.org

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_6

(with nondecreasing $\alpha(x)$) or equivalently as measure integral with measure $d\alpha(x)$. In the continuous case, $d\alpha(x) = \alpha'(x)\,dx$. In the discrete case, the measure $d\alpha(x)$ is a weighted sum of Dirac measures (point measures) ϵ_x at the points of increase x_k of $\alpha(x)$,

$$d\alpha(x) = \sum_{k=0}^{N} \alpha_k \epsilon_{x_k}$$

where α_k denotes the increment of $\alpha(x)$ at x_k, $N \in \mathbb{N}$ or $N = \infty$. In this case, the integral can be computed as the sum

$$\int_a^b x^n d\alpha(x) = \sum_{k=0}^{N} \alpha_k\, x_k^n\ .$$

Note that the Dirac measure ϵ_x at the point y is defined by

$$\epsilon_x(y) = \begin{cases} 1 \text{ if } & y = x \\ 0 \text{ if } & y \neq x. \end{cases}$$

In the q-discrete case, the measure $d\alpha(x)$ takes the form

$$d\alpha(x) = \sum_{k\in\mathbb{Z}} \left(\rho(q^k)\epsilon_{q^k} + \rho(-q^k)\epsilon_{-q^k}\right).$$

Definition 1 ([3, P. 244, Def. 5.2.1]) We say that a polynomial set $\{p_n(x)\}_0^\infty$ is orthogonal with respect to the measure $d\alpha(x)$ if $\forall n, m \in \mathbb{N}$

$$\int_a^b p_n(x)p_m(x)d\alpha(x) = h_n\delta_{mn}, \quad h_n \neq 0, \tag{1.2}$$

where δ_{mn} is the Kronecker delta notation defined by $\delta_{mn} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$.

2 Classical Orthogonal Polynomials of a Discrete Variable

2.1 Definitions and Preliminary Results

Definition 2 Let f be a function of the variable x. The forward and the backward operators Δ and ∇ are, respectively, defined by:

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

For $m \in \mathbb{N} = \{1, 2, 3, \dots\}$, one sets

$$\Delta^{m+1} f(x) = \Delta(\Delta^m f(x)).$$

It should be noted that Δ and ∇ transform a polynomial of degree n ($n \geq 1$) in x into a polynomial of degree $n-1$ in x and a polynomial of degree 0 into the zero polynomial.

The operator Δ fulfils the following properties.

Proposition 3 *Let f and g be two functions in the variable x, a and b be two complex numbers. The following properties are valid.*

1. $\Delta(af(x) + bg(x)) = a\Delta f(x) + b\Delta g(x)$ *(linearity);*
2. $\Delta[f(x)g(x)] = f(x+1)\Delta g(x) + g(x)\Delta f(x) = f(x)\Delta g(x) + g(x+1)\Delta f(x)$, *(product rule).*

Definition 4 ([12, P. 4]) The Pochhammer symbol or shifted factorial is defined by

$$(a)_0 := 1 \quad \text{and} \quad (a)_n = a(a+1)(a+2)\cdots(a+n-1), \quad a \neq 0 \quad n = 1, 2, 3, \dots .$$

The following notation (falling factorial) will also be used:

$$a^{\underline{0}} := 1 \quad \text{and} \quad a^{\underline{n}} = a(a-1)(a-2)\cdots(a-n+1), \quad n = 1, 2, 3, \dots .$$

It should be noted that the Pochhammer symbol and the falling factorial are linked as follows:

$$(-a)_n = (-1)^n a^{\underline{n}}.$$

Definition 5 ([12, P. 5]) The hypergeometric series ${}_rF_s$ is defined by

$${}_rF_s\left(\begin{matrix} a_1, \cdots, a_r \\ b_1, \cdots, b_s \end{matrix} \middle| z\right) := \sum_{n=0}^{\infty} \frac{(a_1, \cdots, a_r)_n}{(b_1, \cdots, b_s)_n} \frac{z^n}{n!},$$

where

$$(a_1, \dots, a_r)_n = (a_1)_n \cdots (a_r)_n.$$

An example of a summation formula for the hypergeometric series is given by the binomial theorem ([12, P. 7])

$${}_1F_0\left(\begin{matrix} -a \\ - \end{matrix} \middle| -z\right) = \sum_{n=0}^{\infty} \binom{a}{n} z^n = (1+z)^a, \quad |z| < 1, \tag{2.1}$$

where

$$\binom{a}{n} = \frac{(-1)^n}{n!}(-a)_n.$$

The following proposition is obtained by simple computations.

Proposition 6 *For all $i \in \mathbb{N}$ the following relations are valid.*

- $\Delta x^{\underline{i}} = i x^{\underline{i-1}}$;
- $\nabla x^{\underline{i}} = i(x-1)^{\underline{i-1}}$;
- $x x^{\underline{i}} = x^{\underline{i+1}} + i x^{\underline{i}}$;
- $\Delta\nabla x^{\underline{i}} = i(i-1)(x-1)^{\underline{i-2}}$;
- $x\Delta\nabla x^{\underline{i}} = i(i-1)x^{\underline{i-1}}$;
- $x^2\Delta\nabla x^{\underline{i}} = i(i-1)x^{\underline{i}} + i(i-1)^2 x^{\underline{i-1}}$;
- $x\Delta x^{\underline{i}} = i x^{\underline{i}} + i(i-1)x^{\underline{i-1}}$;
- $x(x-1)^{\underline{i}} = x^{\underline{i+1}}$.

Definition 7 A polynomial set

$$y(x) = p_n(x) = k_n x^n + \ldots \quad (n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}, \quad k_n \neq 0) \tag{2.2}$$

is a family of discrete classical orthogonal polynomials (also known as the Hahn class) if it is the solution of a difference equation of the type

$$\sigma(x)\Delta\nabla y(x) + \tau(x)\Delta y(x) + \lambda_n y(x) = 0. \tag{2.3}$$

They are known to satisfy the Pearson-type equation

$$\Delta[\sigma(x)\rho(x)] = \tau(x)\rho(x) \tag{2.4}$$

where $\rho(x)$ is the discrete weight function for which the p_n's are orthogonal.

Proposition 8 *Let $(p_n(x))_n$ be a family of classical discrete orthogonal polynomials satisfying the Pearson-type difference equation (2.4), then the family $(\Delta p_{n+1}(x))_n$ is orthogonal with respect to the weight function ρ_1 defined by*

$$\rho_1(x) = \sigma(x+1)\rho(x+1).$$

More generally, we have the following proposition.

Proposition 9 *Let $(p_n(x))_n$ be a family of classical discrete orthogonal polynomials satisfying the Pearson-type difference equation (2.4), then the family $(\Delta^k p_{n+k}(x))_n$ is orthogonal with respect to the weight function ρ_k defined by*

$$\rho_k(x) = \rho(x+k)\prod_{j=1}^{k}\sigma(x+j).$$

Theorem 10 (Rodrigues Formula) *Let $(P_n(x))_n$ be a family of classical discrete orthogonal polynomials. If we set $P_{m,n}(x) = \Delta^m P_n(x)$ $(m \leq n)$, then*

$$P_{m,n}(x) = \frac{A_{mn} B_n}{\rho_m(x)} \nabla^{n-m}[\rho_n(x)] \tag{2.5}$$

with

$$\begin{aligned} A_{mn} &= \frac{n!}{(n-m)!} \prod_{k=0}^{m-1} \left(\tau' + \frac{n+k-1}{2}\sigma''\right); \\ A_{0n} &= 1; \\ B_n &= \frac{\Delta^n P_n(x)}{A_{nn}} = \frac{1}{A_{nn}} P_n^{(n)}(x). \end{aligned}$$

Corollary 11 *For $m = 0$ in (2.5) it follows that*

$$P_n = P_{0n} = y_n(x) = \frac{B_n}{\rho(x)} \nabla^n[\rho(x)]. \tag{2.6}$$

2.2 General Polynomial Solutions of the Hypergeometric Discrete Difference Equation

In this section we provide the hypergeometric representation of the classical discrete orthogonal polynomials. More precisely, we find an explicit representation of the polynomial solutions of the difference equation (2.3). First, we look for the solutions of the difference equation

$$\sigma(x)\Delta\nabla y(x) + \tau(x)\Delta y(x) + \lambda y(x) = 0 \tag{2.7}$$

where σ, τ and λ have the form

- $\sigma(x) = ax^2 + bx + c \quad a, b, c \in \mathbb{R}$;
- $\tau(x) = dx + e \quad d \in \mathbb{R}^*, e \in \mathbb{R}$.
- λ is a constant.

Next we choose λ such that we obtain a polynomial solution. Since the suitable polynomial basis for the forward (or the backward) difference is the falling factorial, we write the solution has

$$y(x) = \sum_{k=0}^{\infty} \alpha_k x^{\underline{k}}. \tag{2.8}$$

Theorem 12 *The infinite series $y(x)$ defined by (2.8) is solution of (2.7) if and only if the α_i's are solutions of the recurrence equation*

$$R(a,b,c,d,e)\alpha_{i+2} + S(a,b,c,d,e)\alpha_{i+1} + T(a,b,c,d,e)\alpha_i = 0 \tag{2.9}$$

where

$$\begin{aligned}
R(a,b,c,d,e) &= ai^4 + (d+b+5a)i^3 + (4d+9a+4b+c+e)i^2 \\
&\quad +(3c+7a+3e+5d+5b)i + 2d + 2e + 2b + 2 \\
&\quad c + 2a; \\
S(a,b,c,d,e) &= 2ai^3 + (3a+b+2d)i^2 + (a+b+3d+e+\lambda)i + \\
&\quad d + e + \lambda; \\
R(a,b,c,d,e) &= ai^2 + (-a+d)i + \lambda.
\end{aligned}$$

Proof The following relations are easily verified

$$\begin{aligned}
\sigma(x)\Delta\nabla y(x) &= \sum_{i=2}^{+\infty}\alpha_i\left\{ax^2\Delta\nabla x^{\underline{i}} + bx\Delta\nabla x^{\underline{i}} + c\Delta\nabla x^{\underline{i}}\right\} \\
&= \sum_{i=2}^{+\infty} i(i-1)\alpha_i\left\{ax^{\underline{i}} + ((i-1)a+b)x^{\underline{i-1}} + c(x-1)^{\underline{i-2}}\right\}; \\
\tau(x)\Delta y(x) &= \sum_{i=1}^{+\infty}\alpha_i\left\{dx\Delta x^{\underline{i}} + e\Delta x^{\underline{i}}\right\} \\
&= \sum_{i=1}^{+\infty} i\alpha_i\left\{dx^{\underline{i}} + ((i-1)d+e)x^{\underline{i-1}}\right\} \\
\lambda y(x) &= \sum_{i=0}^{+\infty}\lambda\alpha_i x^{\underline{i}}.
\end{aligned}$$

In order the convert the term $(x-1)^{\underline{i-2}}$ into a term of the form $x^{\underline{i+j}}$, with $j \in \mathbb{Z}$, we multiply all the previous expressions by x and use the relation

$$x(x-1)^{\underline{i-2}} = x^{\underline{i-1}}$$

to deduce from Proposition 6 that

$$\begin{aligned}
x\sigma(x)\Delta\nabla y(x) &= \sum_{i=2}^{+\infty}\left\{A_i x^{\underline{i+1}} + B_i x^{\underline{i}} + C_i x^{\underline{i-1}}\right\}\\
&= \sum_{i=2}^{+\infty}\left\{A_i + B_{i+1} + C_{i+2}\right\}x^{\underline{i+1}} + B_2 x^{\underline{2}} + C_2 x^{\underline{1}} + C_3 x^{\underline{2}};\\
x\tau(x)\Delta y(x) &= \sum_{i=1}^{+\infty}\left\{A_i' x^{\underline{i+1}} + B_i' x^{\underline{i}} + C_i' x^{\underline{i-1}}\right\}\\
&= \sum_{i=1}^{+\infty}\left\{A_i' + B_{i+1}' + C_{i+2}'\right\}x^{\underline{i+1}} + B_1' x^{\underline{1}} + C_2' x^{\underline{1}} + C_1' x^{\underline{0}};\\
x\lambda y(x) &= \sum_{i=0}^{+\infty}\left\{A_i'' + B_{i+1}''\right\}x^{\underline{i+1}} + B_0'' x^{\underline{0}}.
\end{aligned}$$

with

$$\begin{aligned}
A_i &= ai(i-1)\alpha_i;\\
B_i &= i(i-1)[(2i-1)a+b]\alpha_i;\\
C_i &= i(i-1)[(i-1)^2a+(i-1)b+c]\alpha_i;\\
A_i' &= di\alpha_i;\\
B_i' &= i[(2i-1)d+e]\alpha_i;\\
C_i' &= i[(i-1)^2d+(i-1)e]\alpha_i;\\
A_i'' &= \lambda\alpha_i;\\
B_i'' &= \lambda i\alpha_i.
\end{aligned}$$

Finally, using Eqs. (2.7) and (2.9) follows. □

Remark 13 In the previous statement, the polynomial σ is of the form $\sigma(x) = ax^2 + bx + c$ where c is any real number. For the classical discrete orthogonal polynomials, σ is of the form $\sigma(x) = ax^2 + bx$ since their lattices are selected such that they start with $x = 0$. The previous theorem reduces to

Theorem 14 *If $\sigma(x) = ax^2 + bx$, then the series y defined by (2.8) is solution of (2.7) if and only if the α_i's verify the recurrence relation*

$$(i+1)[i(ai+b)+(id+e)]\alpha_{i+1} - [ai(i-1)+di+\lambda]\alpha_i = 0. \tag{2.10}$$

Equation (2.10) can be written as

$$\alpha_{i+1} = -\frac{ai(i-1)+di+\lambda}{(i+1)[i(ai+b)+(id+e)]}\alpha_i. \tag{2.11}$$

Note that for y to be a polynomial of degree n it necessary that $\alpha_n \neq 0$ and $\alpha_i = 0$ for $i > n$. These conditions imply

$$\lambda = \lambda_n = -n[(n-1)a+d].$$

Replacing λ by λ_n in (2.11), we get

$$\alpha_{i+1} = \frac{(n-i)[a(n+i-1)+d]}{(i+1)[i(ai+b)+(id+e)]}\alpha_i;$$

or otherwise stated

$$\alpha_{i+1} = \frac{(n-i)[a(n+i-1)+d]}{(i+1)[ai^2+(b+d)i+e)]}\alpha_i; \tag{2.12}$$

which gives

$$\alpha_i = (-1)^n C_i(a,b,d,e)\frac{(-n)_i}{i!} \tag{2.13}$$

with

$$C_i(a,b,d,e) = \prod_{k=0}^{i} \frac{a(n+k-2)+d}{a(k-1)^2+(b+d)(k-1)+e}\alpha_0.$$

2.3 *The Four Classical Discrete Orthogonal Polynomials*

2.3.1 Charlier Polynomials

The Charlier polynomials satisfy the difference equation

$$x\Delta\nabla y(x) + (\gamma - x)\Delta y(x) + \lambda y(x) = 0.$$

In this case we have

$$a = 0, \quad b = 1, \quad d = -1, \quad \text{and} \quad e = \gamma.$$

Equation (2.12) becomes in this case

$$\alpha_{i+1} = \frac{1}{\gamma}\frac{i-n}{i+1}\alpha_i.$$

Taking the normalization $\alpha_0 = 1$, we get:

$$\alpha_i = \left(\frac{1}{\gamma}\right)^i \frac{(-n)_i}{i!},$$

and finally we obtain the hypergeometric representation

$$\begin{aligned} C_n(x,\gamma) &= \sum_{k=0}^{n} \left(\frac{1}{\gamma}\right)^k \frac{(-n)_k}{k!} x^{\underline{k}} \\ &= \sum_{k=1}^{n} \left(-\frac{1}{\gamma}\right)^k \frac{(-n)_k(-x)_k}{i!} \\ &= {}_2F_0\left(\begin{matrix} -n, -x \\ - \end{matrix} \middle| -\frac{1}{\gamma} \right). \end{aligned}$$

The weight function of the Charlier polynomials are solution of the Pearson equation

$$\Delta(x\rho(x)) = (\gamma - x)\rho(x),$$

which is equivalent to

$$\frac{\rho(x+1)}{\rho(x)} = \frac{\gamma}{x+1},$$

and so

$$\rho(i) = \frac{\gamma^i}{i!}.$$

2.3.2 Meixner Polynomials

The Meixner polynomials satisfy the difference equation

$$x\Delta\nabla y(x) + ((\mu-1)x + \gamma\mu)\Delta y(x) + \lambda y(x) = 0.$$

Therefore we have

$$a = 0, \quad b = 1, \quad d = \mu - 1 \quad \text{et} \quad e = \gamma\mu.$$

Equation (2.12) becomes

$$\alpha_{i+1} = \frac{(n-i)(\mu-1)}{(i+1)\mu(i+\gamma)}\alpha_i = \frac{(n-i)}{(i+1)}\frac{\mu-1}{\mu(i+\gamma)}\alpha_i = \Big(1-\frac{1}{\mu}\Big)\frac{(n-i)}{(i+1)}\frac{1}{i+\gamma}\alpha_i.$$

Taking the normalization $\alpha_0 = 1$ we get

$$\alpha_i = (-1)^n\Big(1-\frac{1}{\mu}\Big)^i\frac{(-n)_i}{i!}\frac{1}{(-\gamma)_i}.$$

and finally we get the hypergeometric representation

$$\begin{aligned} M_n(x,\gamma,\mu) &= \sum_{k=0}^{n}(-1)^k\Big(1-\frac{1}{\mu}\Big)^k\frac{(-n)_k}{i!}\frac{1}{(-\gamma)_k}(-x)_k \\ &= {}_2F_1\left(\begin{matrix} -n, -x \\ \gamma \end{matrix}\,\middle|\, 1-\frac{1}{\mu}\right). \end{aligned}$$

The weight function of the Meixner polynomials satisfies the Pearson equation

$$\Delta(x\rho(x)) = ((\mu-1)x+\gamma\mu)\rho(x).$$

which is equivalent to

$$\frac{\rho(x+1)}{\rho(x)} = \frac{\mu(\gamma+x)}{x+1}$$

and so

$$\rho(x) = C\frac{\mu^x\Gamma(\gamma+x)}{\Gamma(x+1)}.$$

2.3.3 Kravchuk Polynomials

The Kravchuk polynomials satisfy the difference equation

$$x\Delta\nabla y(x) + \Big(-\frac{1}{q}x+\frac{Np}{q}\Big)\Delta y(x) + \lambda y(x) = 0,$$

with $p+q=1, 0<p<1$. Therefore we have

$$a=0, \quad b=1, \quad d=-\frac{1}{q} \quad \text{et} \quad e=\frac{Np}{q}.$$

Equation (2.12) becomes

$$\alpha_{i+1} = -\frac{(n-i)(\mu-1)}{p(i+1)(i+N)}\alpha_i = -\frac{(n-i)}{(i+1)}\frac{1}{p}\frac{1}{i+N}\alpha_i.$$

Taking the normalization $\alpha_0 = 1$, we get

$$\alpha_i = (-1)^i \left(\frac{1}{p}\right)^i \frac{(-n)_i}{i!}\frac{1}{(-N)_i},$$

and finally we get the hypergeometric representation

$$\begin{aligned} K_n(x, p, N) &= \sum_{k=1}^{n} \left(\frac{1}{p}\right)^i \frac{(-n)_k}{k!}\frac{1}{(-N)_k}(-x)_k \\ &= {}_2F_1\left(\begin{matrix} -n, -x \\ -N \end{matrix} \middle| -\frac{1}{p}\right). \end{aligned}$$

Similarly to the previous families we get the weight function

$$\rho(x) = C\frac{N!p^x q^{N-x}}{\Gamma(x+1)\Gamma(N+1-x)}.$$

2.3.4 Hahn Polynomials

The Hahn polynomials satisfy the difference equation

$$(x^2 + \mu x)\Delta\nabla y(x) + \tau(x)\Delta y(x) + \lambda y(x) = 0$$

with

$$\tau(x) = -(2N + \mu + \nu - 2)x + (N + \nu + 1)(N - 1).$$

Doing as for the previous families, we get the hypergeometric representation

$$\begin{aligned} Q_n(x, \mu, \nu, N) &= \sum_{k=0}^{n} \frac{(-n)_k(n+\nu+\mu+1)_k(-x)_k}{(\alpha+N)_k(-N)_k} \\ &= {}_3F_2\left(\begin{matrix} -n, n+\mu+\nu+1, -x \\ \alpha+N, -N \end{matrix} \middle| -\frac{1}{p}\right) \end{aligned}$$

and the weight function

$$\rho(x) = \frac{1}{\Gamma(x+1)\Gamma(x+\mu+1)\Gamma(N+\nu-x)\Gamma(N-x)} \quad (\nu > -1, \mu > -1).$$

2.4 Some Structure Formulas for the Charlier Polynomials

It is not difficult to prove that the Charlier polynomials have the exponential generating function [12, P. 248]

$$e^t\left(1-\frac{t}{\alpha}\right)^x = \sum_{n=0}^{\infty} C_n(x;\alpha)\frac{t^n}{n!}. \tag{2.14}$$

Indeed

$$\begin{aligned} e^t\left(1-\frac{t}{\alpha}\right)^x &= \left(\sum_{n=0}^{\infty}\frac{t^n}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{x^{\underline{n}}}{(-\alpha)^n}\frac{t^n}{n!},\right) \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{n}{k}\frac{x^{\underline{k}}}{(-\alpha)^k}\right)\frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{(-n)_k(-x)_k}{k!}\frac{1}{(-\alpha)^k}\right)\frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty} C_n(x;\alpha)\frac{t^n}{n!}. \end{aligned}$$

In what follows, we use only this generating function to prove some structure relations for these polynomials.

2.4.1 The Inversion Formula

Proposition 15 *The Charlier polynomials fulfil the following inversion formula.*

$$x^{\underline{n}} = \sum_{k=0}^{n}(-1)^k\binom{n}{k}\alpha^n C_k(x;\alpha). \tag{2.15}$$

Proof From the generating function (2.14), we get

$$\left(1-\frac{t}{\alpha}\right)^{x} = e^{-t}\sum_{n=0}^{\infty} C_n(x;\alpha)\frac{t^n}{n!} = \left(\sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!}\right)\left(\sum_{n=0}^{\infty} C_n(x;\alpha)\frac{t^n}{n!}\right)$$
$$= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}C_k(x;\alpha)\right)\frac{t^n}{n!}.$$

Using the fact that

$$\left(1-\frac{t}{\alpha}\right)^{x} = \sum_{n=0}^{\infty} x^{\underline{n}}\frac{(-t/\alpha)^n}{n!} = \sum_{n=0}^{\infty}\frac{x^{\underline{n}}}{(-\alpha)^n}\frac{t^n}{n!},$$

the result follows. □

2.4.2 A Connection Formula

Note that by doing the change of variable $t = \alpha t$, the generating function (2.14) can be written as

$$e^{\alpha t}\,(1-t)^{x} = \sum_{n=0}^{\infty}\alpha^n C_n(x;\alpha)\frac{t^n}{n!}. \tag{2.16}$$

Theorem 16 *The following connection formula holds true*

$$C_n(x;\beta) = \frac{1}{\beta^n}\sum_{k=0}^{n}\binom{n}{k}(\beta-\alpha)^{n-k}\alpha^k C_k(x;\alpha). \tag{2.17}$$

Proof From the relation (2.16), we get

$$\sum_{n=0}^{\infty}\beta^n C_n(x;\beta)\frac{t^n}{n!} = e^{\beta t}\,(1-t)^x = e^{(\beta-\alpha)t}e^{\alpha t}\,(1-t)^x$$
$$= \left(\sum_{n=0}^{\infty}(\beta-\alpha)^n\frac{t^n}{n!}\right)\left(\sum_{n=0}^{\infty}\alpha^n C_n(x,\alpha)\frac{t^n}{n!}\right)$$
$$= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{n}{k}(\beta-\alpha)^{n-k}\alpha^k C_k(x;\alpha)\right)\frac{t^n}{n!}.$$

The result follows by collecting the coefficients of t^n on both sides. □

2.4.3 An Addition Formula

Theorem 17 *The following addition formula holds true*

$$C_n(x+y;\alpha+\beta)=\sum_{k=0}^{n}\binom{n}{k}\frac{\alpha^k\beta^{n-k}}{(\alpha+\beta)^n}C_k(x;\alpha)C_{n-k}(y;\beta). \tag{2.18}$$

Proof From the generating function (2.16), we get

$$\begin{aligned}\sum_{n=0}^{\infty}(\alpha+\beta)^n C_n(x+y;\alpha+\beta)\frac{t^n}{n!}\\ &= e^{\alpha t}\,(1-t)^x\,e^{\beta y}\,(1-t)^y\\ &=\left(\sum_{n=0}^{\infty}\alpha^n C_n(x;\alpha)\frac{t^n}{n!}\right)\left(\sum_{n=0}^{\infty}\beta^n C_n(y;\beta)\frac{t^n}{n!}\right)\\ &=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{n}{k}\alpha^k\beta^{n-k}C_k(x;\alpha)C_{n-k}(y;\beta)\right)\frac{t^n}{n!}.\end{aligned}$$

□

Remark 18 In particular, when $\alpha=\beta$, (2.18) becomes

$$C_n(x+y;2\alpha)=\frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}C_k(x;\alpha)C_{n-k}(y;\alpha).$$

3 Classical Orthogonal Polynomials of a q-Discrete Variable

3.1 Definitions and Preliminary Results

This section contains some preliminary definitions and results that are useful for a better reading of the manuscript. The q-Pochhammer symbol, the q-binomial coefficients, the q-hypergeometric series, the q-derivative, the q-integral are defined. The reader will consult the references [12, 17] for more informations about these concepts.

Definition 19 The basic hypergeometric or q-hypergeometric series ${}_r\phi_s$ is defined by the series

$${}_r\phi_s\left(\left.\begin{matrix}a_1,\cdots,a_r\\ b_1,\cdots,b_s\end{matrix}\right|q;z\right):=\sum_{n=0}^{\infty}\frac{(a_1,\cdots,a_r;q)_n}{(b_1,\cdots,b_s;q)_n}\left((-1)^n q^{\binom{k}{2}}\right)^{1+s-r}\frac{z^n}{(q;q)_n},$$

where

$$(a_1, \cdots, a_r)_n := (a_1; q)_n \cdots (a_r; q)_n,$$

with

$$(a_i; q)_n = \begin{cases} \prod_{j=0}^{n-1} (1 - a_i q^j) & \text{if } n = 1, 2, 3, \cdots \\ 1 & \text{if } n = 0 \end{cases}.$$

For $n = \infty$ we set

$$(a; q)_\infty = \prod_{n=0}^{\infty} (1 - aq^n), \ |q| < 1.$$

The notation $(a; q)_n$ is the so-called q-Pochhammer symbol.

The following so-called q-binomial theorem if one of the most important identities for basic hypergeometric series.

Proposition 20 ([12, Page 16]) *The basic hypergeometric series fulfil the following identity*

$${}_1\phi_0\left(\begin{array}{c} a \\ - \end{array} \middle| q; z\right) = \sum_{n=0}^{\infty} \frac{(a; q)_n}{(q; q)_n} z^n = \frac{(az; q)_\infty}{(z; q)_\infty}, \quad 0 < |q| < 1, \ \ |z| < 1. \tag{3.1}$$

From the definition of $(a; q)_\infty$, it follows that for $0 < |q| < 1$, and for a nonnegative integer n, we have

$$(a; q)_n = \frac{(a; q)_\infty}{(aq^n; q)_\infty}.$$

Definition 21 For any complex number λ,

$$(a; q)_\lambda = \frac{(a; q)_\infty}{(aq^\lambda; q)_\infty}, \quad 0 < |q| < 1,$$

where the principal value of q^λ is taken.

We will also use the following common notations

$$[a]_q = \frac{1 - q^a}{1 - q}, \quad a \in \mathbb{C}, \quad q \neq 1,$$

$$\begin{bmatrix} n \\ m \end{bmatrix}_q = \frac{(q; q)_n}{(q; q)_m (q; q)_{n-m}}, \quad 0 \le m \le n,$$

and

$$(x \ominus y)_q^n = (x-y)(x-qy)\cdots(x-q^{n-1}y). \tag{3.2}$$

called the q-bracket, the q-binomial coefficients and the q-power, respectively.

Remark 22 Note that in (3.2), for $x = 1$ we obtain the q-Pochhammer symbol and for $y = 0$ we have the classical power. It is also straightforward to see that for $x \neq 0$, $(x \ominus y)_q^n = x^n \left(\frac{y}{x}; q\right)_n$.

Remark 23 Notice that the q-factorial and the q-Pochhammer symbol are linked in the following way

$$[n]_q! = \frac{(q;q)_n}{(1-q)^n}.$$

The natural extension of the q-power to the real numbers λ is

$$(a \ominus b)_q^\lambda = a^\lambda \frac{(b/a;q)_\infty}{(q^\lambda b/a;q)_\infty} = a^\lambda (b/a;q)_\lambda.$$

Definition 24 The q-derivative operator is defined by [10, 12]

$$D_q f(x) = \frac{f(x) - f(qx)}{(1-q)x}, \quad x \neq 0,$$

and

$$D_q f(0) = f'(0)$$

provided that f is differentiable at $x = 0$.

The q-derivative operator satisfies the important product rule

$$D_q(f(x)g(x)) = f(x)D_q g(x) + g(qx)D_q f(x). \tag{3.3}$$

Definition 25 (See [10]) Suppose $0 < a < b$. The definite q-integral is defined as

$$\int_0^b f(x)d_qx = (1-q)b\sum_{n=0}^{\infty} q^n f(q^n b), \tag{3.4}$$

and

$$\int_a^b f(x)d_qx = \int_0^b f(x)d_qx - \int_0^a f(x)d_qx. \tag{3.5}$$

Definition 26 The q-Gamma function is defined by

$$\Gamma_q(x) := \frac{(q;q)_\infty}{(q^x;q)_\infty}(1-q)^{1-x}, \quad 0 < q < 1. \tag{3.6}$$

Remark 27 From Definition 21, the q-Gamma function is also represented by

$$\Gamma_q(x) = (1-q)^{1-x}(q;q)_{x-1}.$$

Note also that the q-Gamma function satisfies the functional equation

$$\Gamma_q(x+1) = [x]_q \Gamma_q(x), \quad \text{with} \quad \Gamma_q(1) = 1.$$

Note that for arbitrary complex α,

$$\begin{bmatrix} \alpha \\ k \end{bmatrix}_q = \frac{(q^{-\alpha};q)_k}{(q;q)_k}(-1)^k q^{\alpha k - \binom{k}{2}} = \frac{\Gamma_q(\alpha+1)}{\Gamma_q(k+1)\Gamma_q(\alpha-1)}. \tag{3.7}$$

The exponential function has two different natural q-extensions, denoted by $e_q(z)$ and $E_q(z)$, which can be defined by

$$e_q(z) := \sum_{n=0}^{\infty} \frac{z^n}{[n]_q!}, \quad 0 < |q| < 1, \quad |z| < 1, \tag{3.8}$$

and

$$E_q(z) := \sum_{n=0}^{\infty} \frac{q^{\binom{n}{2}}}{[n]_q!} z^n, \quad 0 < |q| < 1. \tag{3.9}$$

Using the identity (3.1), it follows that

$$e_q(z) = \frac{1}{((1-q)z;q)_\infty},$$

and

$$E_q(z) = (-(1-q)z;q)_\infty.$$

These q-analogues of the exponential function are clearly related by

$$e_q(z)E_q(-z) = 1.$$

Definition 28 A polynomial set $p_n(x)$, given by (2.2), is a family of classical q-discrete orthogonal polynomials (also known as the polynomials of the q-Hahn tableau) if it is the solution of a q-difference equation of the type

$$\sigma(x)D_qD_{q^{-1}}y(x)+\tau(x)D_qy(x)+\lambda_ny(x)=0. \tag{3.10}$$

Here the polynomials $\sigma(x)$ and $\tau(x)$ are involved in the Pearson type equation

$$D_q(\sigma(x)\rho(x))=\tau(x)\rho(x),$$

where the function $\rho(x)$ is the q-discrete weight function associated to the family.

3.2 Polynomial Solutions of the q-Difference Equation

Theorem 29 (See [7]) *Let $P_n(x)$ be a polynomial system given by the q-differential equation (3.10) with $\sigma(x)=ax^2+bx+c$, and $\tau(x)=dx+e$. Then, the power series coefficients $C_m(n)$ given by*

$$P_n(x)=\sum_{m=0}^{n}C_m(n)x^m \tag{3.11}$$

satisfy the recurrence equation

$$\begin{aligned}(a[m]_{\frac{1}{q}}[m-1]_q+d[m]_q-\lambda_n)C_m(n)+(b[m+1]_{\frac{1}{q}}[m]_q+e[m+1]_q)C_{m+1}(n)\\ +c[m+2]_{\frac{1}{q}}[m+1]_qC_{m+2}(n)=0,\end{aligned} \tag{3.12}$$

with $C_n(n)=1$, $C_{n+1}(n)=0$. In particular, if $c=0$, then the recurrence equation

$$(a[m]_{1/q}[m-1]_q+d[m]_q-\lambda_n)C_m(n)+(b[m+1]_{1/q}[m]_q+e[m+1]_q)C_{m+1}(n), \tag{3.13}$$

is valid, and therefore $P_n(x)$ has the following q-hypergeometric representation up to a constant K_n:

$$P_n(x)=K_n\,{}_2\phi_1\left(\left.\begin{matrix}q^{-n},\frac{a-d+dq}{a}q^{n-1}\\ \frac{b-e+eq}{b}\end{matrix}\right|q;-\frac{aq}{b}x\right),\qquad ab\neq 0, \tag{3.14}$$

$$P_n(x)=K_n\,{}_1\phi_1\left(\left.\begin{matrix}q^{-n}\\ \frac{b-e+eq}{b}\end{matrix}\right|q;\frac{d(1-q)q^n}{b}x\right)\qquad a=0,\ b\neq 0, \tag{3.15}$$

$$P_n(x)=K_n\,{}_1\phi_0\left(\left.\begin{matrix}q^{-n}\\ -\end{matrix}\right|q;-\frac{dq^n}{e}x\right),\qquad a=b=0. \tag{3.16}$$

Proof Substituting the power series (3.11) into the q-differential equation (3.10), and equating the coefficients yields the recurrence equation (3.12).

For $c = 0$ this recurrence equation degenerates to a two-term recurrence equation, and hence establishes the q-hypergeometric representations (3.14)–(3.16), using the initial value $C_n(n) = 1$, $C_{n+1}(n) = 0$. □

We would like to mention that the recurrence equation (3.12) carries complete information about the q-hypergeometric representations given in the theorem.

Theorem 30 (See [7]) *Let $P_n(x)$ be a polynomial system given by the q-differential equation (3.10) with $\sigma(x) = ax^2 + bx + c$, and $\tau(x) = dx + e$. Then, the power series coefficients $C_m(n)$ given by*

$$P_n(x) = \sum_{m=0}^{n} C_m(n)(x;q)_m \tag{3.17}$$

satisfy the recurrence equation

$$\begin{aligned}
&q^n\left(q^{m+2}-1\right)\left(q^{m+1}-1\right)\left(a+q^{m+1}b+cq^{2m+2}\right)C_{m+2}(n)\\
&-\left(q^{m+1}-1\right)q\Big(-q^{n+1}a-aq^n+q^{n+2m+1}b-q^{m+1+n}b\\
&+q^{2m+2+n}e-q^{n+2m+1}e+q^{m+2n}a+q^{m+2n+1}d-q^{m+2n}d+q^{m+1}a\Big)C_{m+1}(n)\\
&-\left(-q^m+q^n\right)\left(q^{n+m}a+q^{m+1+n}d-q^{n+m}d-aq\right)q^2C_m(n)=0,
\end{aligned} \tag{3.18}$$

where $m = -2, -1, 0, \cdots, n$ and $C_m(n) = 0$ outside the set of (n, m) such that $0 \le m \le n$, with $C_n(n) = 1$, $C_{n+1}(n) = 0$.

Proof We first remark the following relations:

$$\begin{aligned}
&D_q(x;q)_m = -\frac{[m]_q}{1-x}(x;q)_m \text{ or } D_q(x;q)_m = -[m]_q(qx;q)_{m-1},\\
&D_{\frac{1}{q}}(x;q)_m = -[m]_q(x;q)_{m-1},\\
&x(qx;q)_n = q^{-n-1}(qx;q)_n - q^{-n-1}(qx;q)_{n+1}.
\end{aligned}$$

From these relations, we obtain

$$\begin{aligned}
(x;q)_m &= (1-q^{-m})(qx;q)_{m-1} + q^{-m}(qx;q)_m,\\
x(qx;q)_{m-1} &= q^{-m}(qx;q)_{m-1} - q^{-m}(qx;q)_m,\\
x(qx;q)_{m-2} &= q^{-m+1}(qx;q)_{m-2} - q^{-m+1}(qx;q)_{m-1},\\
x^2(qx;q)_{m-2} &= q^{-2m+2}(qx;q)_{m-2} - (q^{-2m+2}+q^{-2m+1})(qx;q)_{m-1}\\
&\quad +q^{-2m+1}(qx;q)_m,\\
(qx;q)_m &= \frac{(x;q)_m}{1-x}.
\end{aligned}$$

Next, we substitute $P_n(x)$ in the q-differential equation (3.10) and obtain (using the preceding relations and simplification)

$$\sum_{m=0}^{n} C_m(n)(x;q)_{m+1}\Big(a[m]_q[m-1]_q q^{-2m+1} + d[m]_q q^{-m} + \lambda_n q^{-m}\Big)$$
$$+\sum_{m=-1}^{n-1} C_{m+1}(n)(x;q)_{m+1}\Big(-a[m+1]_q[m]_q(q^{-2m}+q^{-2m-1}) - b[m+1]_q[m]_q q^{-m}$$
$$-d[m+1]_q q^{-m-1} - e[m+1]_q + \lambda_n(1-q^{-m-1})\Big) + \sum_{m=-2}^{n-2} C_{m+2}(n)(x;q)_{m+1}\times$$
$$\Big(a[m+2]_q[m+1]_q q^{-2m-2} + b[m+2]_q[m+1]_q q^{-m-1} + c[m+2]_q[m+1]_q\Big) = 0.$$

Since $(x;q)_m$ is a linearly independent family, equating the coefficients of $(x;q)_n$ yields the constant

$$\lambda_n = -a[n]_{\frac{1}{q}}[n-1]_q - d[n]_q.$$

Equating the coefficients of $(x;q)_{m+1}$, yields the desired recurrence equation satisfied by the coefficients $C_m(n)$. □

The above computations show that in the general case, we get a q-holonomic three-term recurrence equation for $C_m(n)$. In order to find solutions which are q-hypergeometric terms—hence satisfying a first-order q-holonomic recurrence—in some specific situations, we can use a q-version of Petkovšek's algorithm ([16], see e. g. [13]) which was given by Abramov, Paule and Petkovšek [1] and by Böing and Koepf [4]. This algorithm can be used utilizing the `qrecsolve` command of the `qsum` package in Maple [4].

The q-Petkovšek algorithm can be successfully used for several instances in this paper. However, this algorithm is rather inefficient and therefore not at all suitable for many of our complicated questions posed. Fortunately Horn [8, 9] published a refined version based on ideas by Mark Van Hoeij [5] which is much more efficient. Sprenger [19] presented a Maple implementation of this refined version `qHypergeomSolveRE` in his package `qFPS` which finds easily the q-hypergeometric term solutions of all q-recurrence equations of this paper. It is due to this algorithm that we can state all our results, especially in Corollary 31.

In particular, we can solve the recurrence equations of Theorems 29 and 30 for all particular systems and therefore obtain the q-hypergeometric representations up to a constant K_n. Here, we consider the monic cases such that by equating the highest coefficients of (3.11) and (3.17), we can recover the constant K_n. This method yields

Corollary 31 *The following representations of monic orthogonal polynomials of the q-Hahn class are valid:*

- *the Big q-Jacobi polynomials*

$$\tilde{P}_n(x,\alpha,\beta,\gamma;q) = \frac{(\alpha q;q)_n(\gamma q;q)_n}{(\alpha\beta q^{n+1};q)_n}{}_3\phi_2\left(\begin{matrix} q^{-n}, \alpha\beta q^{n+1}, x \\ \alpha q, \gamma q \end{matrix}\,\middle|\, q;q\right),$$

- *the q-Hahn polynomials*

$$\tilde{Q}_n(x,\alpha,\beta,N,q) = \frac{(\alpha q;q)_n(q^{-N};q)_n}{(\alpha\beta q^{n+1};q)_n}{}_3\phi_2\left(\begin{matrix} q^{-n}, \alpha\beta q^{n+1}, x \\ \alpha q, q^{-N} \end{matrix}\,\middle|\, q;q\right)$$

- *the Big q-Laguerre polynomials*

$$tildeP_n(x,\alpha,\beta,q) = (\alpha q;q)_n(\beta q;q)_n{}_3\phi_2\left(\begin{matrix} q^{-n}, 0, x \\ \alpha q, \beta q \end{matrix}\,\middle|\, q;q\right)$$

- *the Little q-Jacobi polynomials*

$$\tilde{p}_n(x;\alpha,\beta|q) = \frac{(-1)^n q^{\binom{n}{2}}(\alpha q;q)_n}{(\alpha\beta q^{n+1};q)_n}{}_2\phi_1\left(\begin{matrix} q^{-n}, \alpha\beta q^{n+1} \\ \alpha q \end{matrix}\,\middle|\, q;qx\right)$$

- *the Alternative q-Charlier polynomials*

$$\tilde{K}_n(x,\alpha,q) = \frac{(-1)^n q^{\frac{n(n-1)}{2}}}{(-\alpha q^n;q)_n}{}_2\phi_1\left(\begin{matrix} q^{-n}, -\alpha q^n \\ 0 \end{matrix}\,\middle|\, q;qx\right)$$

- *the Little q-Laguerre/Wall polynomials*

$$\tilde{p}_n(x,\alpha|q) = (-1)^n q^{\binom{n}{2}}(\alpha q;q)_n{}_2\phi_1\left(\begin{matrix} q^{-n}, 0 \\ \alpha q \end{matrix}\,\middle|\, q,qx\right)$$

- *the q-Meixner polynomials*

$$\tilde{M}_n(x,\beta,\gamma;q) = (-\gamma)^n q^{n^2}(\beta q;q)_n{}_2\phi_1\left(\begin{matrix} q^{-n}, x \\ \beta q \end{matrix}\,\middle|\, q;-\frac{q^{n+1}}{\gamma}\right)$$

- *the q-Charlier polynomials*

$$\tilde{C}_n(x,\alpha,q) = (-\alpha)^n q^{-n^2}{}_2\phi_1\left(\begin{matrix} q^{-n}, x \\ 0 \end{matrix}\,\middle|\, q;-\frac{q^{n+1}}{\alpha}\right)$$

- *the q-Laguerre polynomials*

$$\tilde{L}_n^{(\alpha)}(x;q) = \frac{(-1)^n (q^{\alpha+1} : q)_n}{q^{n(n+\alpha)}} {}_1\phi_1 \left(\begin{array}{c} q^{-n} \\ q^{\alpha+1} \end{array} \middle| q; -xq^{\alpha+n+1} \right)$$

- *the Stieltjes-Wigert polynomials*

$$\tilde{S}_n(x;q) = (-1)^n q^{-n^2} {}_1\phi_1 \left(\begin{array}{c} q^{-n} \\ 0 \end{array} \middle| q; -xq^{n+1} \right)$$

- *the Al Salam-Carlitz II polynomials*

$$\tilde{V}_n^{(\alpha)}(x,q) = (-\alpha)^n q^{-\binom{n}{2}} {}_2\phi_0 \left(\begin{array}{c} q^{-n}, x \\ - \end{array} \middle| q; \frac{q^n}{\alpha} \right)$$

- *the Discrete q-Hermite II polynomials*

$$\tilde{H}_n(x,q) = q^{-\frac{n(n-1)}{2}} {}_2\phi_0 \left(\begin{array}{c} q^{-n}, x \\ - \end{array} \middle| q, -q^n \right)$$

Theorem 32 *Let $P_n(x)$ be a polynomial system given by the q-differential equation (3.10) with $\sigma(x) = (x-1)(x-a)$, and $\tau(x) = dx + e$. Then, the power series coefficients $C_m(n)$ given by*

$$P_n(x) = \sum_{m=0}^{n} C_m(n)(x \ominus 1)_q^m, \tag{3.19}$$

satisfy the recurrence relation

$$\begin{aligned} &\left(q^{m+1}-1\right)\left(q^{2m}-q^m-aq^m+a+dq^{2m+1}-dq^{2m}+eq^{m+1}-eq^m\right)C_{m+1}(n) \\ &\quad+\left(-q^{m+1}+q^{2m}+q-q^m+dq^{2m+1}-dq^{2m}-dq^{m+1}+dq^m+\lambda_n q^{m+2}\right. \\ &\qquad\left.-2\lambda_n q^{m+1}+\lambda_n q^m\right)C_m(n)=0. \end{aligned} \tag{3.20}$$

Proof First remark that $D_q D_{\frac{1}{q}} = \frac{1}{q} D_{\frac{1}{q}} D_q$ and rewrite (3.10) as

$$(x-a)(x-1)D_{\frac{1}{q}} D_q P_n(x) + q(dx+e)D_q P_n(x) + q\lambda_n P_n(x) = 0, \tag{3.21}$$

Next, substituting (3.19) in (3.10), and collecting the coefficients of $C_{m+1}(n)$ and $C_m(n)$, the desired recurrence relation in obtained. $\square$

Corollary 33 *The Al-Salam-Carlitz and the Discrete q-Hermite polynomials have the following monic q-hypergeometric representation*

$$U_n^{(a)}(x;q) = (-a)^n q^{\binom{n}{2}} {}_2\phi_0 \left(\begin{array}{c} q^{-n}, x^{-1} \\ 0 \end{array} \middle| \frac{qx}{a} \right),$$

$$h_n(x;q) = U_n^{(-1)}(x;q) = q^{\binom{n}{2}} {}_2\phi_0 \left(\begin{array}{c} q^{-n}, x^{-1} \\ 0 \end{array} \middle| -qx \right).$$

Proof For the Al-Salam-Carlitz I polynomials we have $d = \frac{1}{1-q}$ and $e = \frac{a+1}{q-1}$ and $\lambda_n = \frac{q^n-1}{q^{n-1}(q-1)^2}$. Substituting these data in (3.20), we obtain the following recurrence relation

$$aq^n \left(q^{m+1} - 1\right) C(m+1) - q\left(-q^n + q^m\right) C(m) = 0.$$

This yields the following

$$C_m(n) = K_n \left(\frac{q}{a}\right)^m \frac{(q^{-n};q)_m}{(q;q)_m}.$$

In order to obtain the monic representation, we determine the leading coefficients and get

$$K_n = (-a)^n q^{\binom{n}{2}}.$$

Finally, using the relation $(x \ominus 1)_q^n = x^n \left(\frac{1}{x};q\right)_n$, the desired q-hypergeometric representation is obtained. □

3.3 Some Structure Formulas for the Al-Salam Carlitz I Polynomials

Note that the Al-Salam Carlitz I polynomials have the generating function [12]

$$\frac{e_q(xt)}{e_q(t)e_q(at)} = \sum_{n=0}^{\infty} U_n^{(a)}(x;q) \frac{t^n}{[n]_q!}, \tag{3.22}$$

Using only this generating function, we prove

3.3.1 The Inversion Formula

Proposition 34 (See [2, 15]) *The following inversion formulas hold for the Al-Salam Carlitz I polynomials* $U_n^{(a)}(x;q)$*:*

$$(x\ominus 1)_q^n=\sum_{k=0}^{n}a^{n-k}\begin{bmatrix}n\\k\end{bmatrix}_q U_k^{(a)}(x;q). \tag{3.23}$$

Proof We first remark that [14, (5.19)]

$$(x\ominus y)_q^n=\sum_{k=0}^{n}(-y)^{n-k}q^{\binom{n-k}{2}}\begin{bmatrix}n\\k\end{bmatrix}_q x^k.$$

Next, taking into account that $e_q(x)E_q(-x)=1$ and multiplying the generating function (3.22) by $e_q(at)$, the left-hand side gives

$$\begin{aligned}\frac{e_q(xt)}{e_q(t)}&=e_q(xt)E_q(-t)\\&=\left(\sum_{k=0}^{\infty}\frac{x^nt^n}{[n]_q!}\right)\left(\sum_{k=0}^{\infty}\frac{q^{\binom{n}{2}}}{[n]_q!}(-t)^n\right)\\&=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(-1)^{n-k}q^{\binom{n-k}{2}}\begin{bmatrix}n\\k\end{bmatrix}_q x^k\right)\frac{t^n}{[n]_q!}\\&=\sum_{n=0}^{\infty}(x\ominus 1)_q^n\frac{t^n}{[n]_q!},\end{aligned}$$

and the right-hand side gives

$$\begin{aligned}e_q(at)\sum_{n=0}^{\infty}U_n^{(a)}(x;q)\frac{t^n}{[n]_q!}&=\left(\sum_{n=0}^{\infty}a^n\frac{t^n}{[n]_q!}\right)\left(\sum_{n=0}^{\infty}U_n^{(a)}(x;q)\frac{t^n}{[n]_q!}\right)\\&=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a^{n-k}\begin{bmatrix}n\\k\end{bmatrix}_q U_k^{(a)}(x;q)\right)\frac{t^n}{[n]_q!}.\end{aligned}$$

Hence we have

$$\sum_{n=0}^{\infty}(x\ominus 1)_q^n\frac{t^n}{[n]_q!}=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a^{n-k}\begin{bmatrix}n\\k\end{bmatrix}_q U_k^{(a)}(x;q)\right)\frac{t^n}{[n]_q!}.$$

So (3.23) is proved. Note that this result is proved in [15] using Verma's q-extension [20] of the Fields and Wimp inversion formula [6]. □

3.3.2 A Connection Formula

Before we state the connection formula, we prove the following proposition.

Proposition 35 *The following expansion applies*

$$\frac{e_q(\alpha z)}{e_q(\beta z)} = \sum_{n=0}^{\infty} \frac{(\alpha \ominus \beta)_q^n}{[n]_q!} z^n.$$

Proof Using the equation $\dfrac{e_q(\alpha z)}{e_q(\beta z)} = e_q(\alpha z) E_q(-\beta z)$ we get

$$\begin{aligned}
\frac{e_q(\alpha z)}{e_q(\beta z)} &= \sum_{n=0}^{\infty} \frac{(\alpha z)^n}{[n]_q!} \times \sum_{n=0}^{\infty} \frac{q^{\binom{n}{2}}(-\beta z)^n}{[n]_q!} \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q q^{\binom{n-k}{2}} \alpha^k (-\beta)^{n-k} \right) \frac{z^n}{[n]_q!} \\
&= \sum_{n=0}^{\infty} \frac{(\alpha \ominus \beta)_q^n}{[n]_q!} z^n.
\end{aligned}$$

□

Theorem 36 *The following connection formula holds for the Al-Salam Carlitz I polynomials*

$$U_n^{(b)}(x;q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q (a \ominus b)_q^k U_k^{(a)}(x;q).$$

Proof From the generating function (3.22), we get

$$\begin{aligned}
\sum_{n=0}^{\infty} U_n^{(b)}(x;q) \frac{t^n}{[n]_q!} &= \frac{e_q(xt)}{e_q(t) e_q(bt)} \\
&= \frac{e_q(at)}{e_q(bt)} \frac{e_q(xt)}{e_q(t) e_q(at)} \\
&= \frac{e_q(at)}{e_q(bt)} \sum_{n=0}^{\infty} U_n^{(a)}(x;q) \frac{t^n}{[n]_q!}
\end{aligned}$$

$$= \left(\sum_{n=0}^{\infty} \frac{(a \ominus b)_q^n}{[n]_q!} t^n\right) \left(\sum_{n=0}^{\infty} U_n^{(a)}(x;q) \frac{t^n}{[n]_q!}\right)$$

$$= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q (a \ominus b)_q^k U_k^{(a)}(x;q)\right) \frac{t^n}{[n]_q!}$$

Hence the theorem is proved. □

3.3.3 A q-Addition Formula

We first recall the following definition and result.

Definition 37 (See [18]) Let a and b two real or complex numbers. Then, the Ward q-addition of a and b is given by

$$(a \oplus_q b)^n := \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q a^k b^{n-k}, \quad n = 0, 1, 3, \ldots . \tag{3.24}$$

Proposition 38 (See [11]) *For $x, y \in \mathbb{R}$, the following formula applies*

$$e_q(x)e_q(y) = e_q(x \oplus_q y). \tag{3.25}$$

Theorem 39 *The following q-addition formula holds for the Al-Salam Carlitz I polynomials*

$$U_n^{(a)}(x \oplus_q y) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q y^{n-k} U_k^{(a)}(x;q)$$

$$= \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q x^{n-k} U_k^{(a)}(y;q).$$

Proof From the generating function (3.22), we get

$$\sum_{n=0}^{\infty} U_n^{(a)}(x \oplus_q y) \frac{t^n}{[n]_q!} = \frac{e_q((x \oplus_q y)t)}{e_q(t)e_q(at)} = \frac{e_q(xt)e_q(yt)}{e_q(t)e_q(at)}$$

$$= \left(\sum_{n=0}^{\infty} y^n \frac{t^n}{[n]_q!}\right) \left(\sum_{n=0}^{\infty} U_n^{(a)}(x) \frac{t^n}{[n]_q!}\right)$$

$$= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q y^{n-k} U_k^{(a)}(x;q)\right) \frac{t^n}{[n]_q!}$$

which proves the result. □

3.3.4 A Multiplication Formula

Theorem 40 *The following multiplication formula holds true for any non-zero real number* α

$$U_n^{(a)}(\alpha x; q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q (\alpha \ominus 1)_q^{n-k} x^{n-k} U_k^{(a)}(x; q). \tag{3.26}$$

Proof From the generating function (3.22), we get

$$\begin{aligned} \sum_{n=0}^{\infty} U_n^{(a)}(\alpha x; q) \frac{t^n}{[n]_q!} &= \frac{e_q(\alpha x t)}{e_q(t) e_q(at)} = \frac{e_q(\alpha x t)}{e_q(xt)} \frac{e_q(xt)}{e_q(t) e_q(at)} \\ &= \left(\sum_{n=0}^{\infty} (\alpha x \ominus x)_q^n \frac{t^n}{[n]_q!} \right) \left(\sum_{n=0}^{\infty} U_n^{(a)}(x; q) \frac{t^n}{[n]_q!} \right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q (\alpha x \ominus x)_q^{n-k} U_k^{(a)}(x; q) \right) \frac{t^n}{[n]_q!}. \end{aligned}$$

The result follows from the relation $(\alpha x \ominus x)_q^{n-k} = (\alpha \ominus 1)_q^{n-k} x^{n-k}$. $\square$

References

1. S. Abramov, P. Paule, M. Petkovšek, q-Hypergeometric solutions of q-difference equations. Discrete Math. **180**, 3–22 (1998)
2. W.A. Al-Salam, q-Appell polynomials. Ann. Mat. Pura Appl. **77**, 31–45 (1967)
3. G.E. Andrews, R. Askey, R. Roy, Special functions, in *Encyclopedia of Mathematics and its Applications*, vol. 71 (Cambridge University Press, Cambridge, 2000)
4. H. Böing, W. Koepf, Algorithms for q-hypergeometric summation in computer algebra. J. Symb. Comput. **11**, 1–23 (1999)
5. T. Cluzeau, M. van Hoeij, Computing hypergeometric solutions of linear recurrence equations. Appl. Algebra Eng. Commun. Comput. **17**, 83–115 (2005)
6. J.L. Fields, J. Wimp, Expansions of the hypergeometric functions in hypergeometric functions. Math. Comp. **15**, 390–395 (1961)
7. M. Foupouagnigni, W. Koepf, D.D. Tcheutia, P. Njionou Sadjang, Representations of q-orthogonal polynomials. J. Symb. Comput. **47**, 347–1371 (2012)
8. P. Horn, Faktorisierung in Schief-Polynomringen, Ph.D. Thesis, Universität Kassel (2008). Available at: http://www.urn.fi/urn:nbn:de:hebis:34-2009030226513
9. P. Horn, W. Koepf, T. Sprenger, m-Fold hypergeometric solutions of linear recurrence equations revisited. Math. Comput. Sci. **6**, 61–77 (2012)
10. V. Kac, P. Cheung, *Quantum Calculus* (Springer, Berlin, 2001)
11. T. Kim, q-Extension of the Euler Formula and Trigometric functions. Russ. J. Math. Phys. **13**, 275–278 (2007)
12. R. Koekoek, P.A. Lesky, R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues* (Springer, Berlin, 2010)

13. W. Koepf, *Hypergeometric Summation* (Vieweg, Braunschweig, 1998)
14. P. Njionou Sadjang, Moments of classical orthogonal polynomials, Ph.D. Thesis, Universität Kassel (2013). Available at http://nbn-resolving.de/urn:nbn:de:hebis:34-2013102244291
15. P. Njionou Sadjang, W. Koepf, M. Fouppuagnigni, On moments of classical orthogonal polynomials. J. Math. Anal. Appl. **424**, 122–151 (2015)
16. M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients. J. Symb. Comput. **14**, 243–264 (1992)
17. E.D. Rainville, *Special Functions* (The Macmillan Company, New York, 1960)
18. M. Schork, q-Dedekind type sums related to q-zeta function and basic L-series. J. Math. Anal. Appl. **318**, 333–351 (2006)
19. T. Sprenger, Algorithmen für q-holonome Funktionen und q-hypergeometrische Reihen, Ph.D. Thesis. Universität Kassel (2009). Available at http://urn:nbn:de:hebis:34-2009072129130
20. A. Verma, Certain expansions of the basic hypergeometric functions. Math. Comp. **20**, 151–157 (1966)

Computer Algebra, Power Series and Summation

Wolfram Koepf

Abstract Computer algebra systems can do many computations that are relevant for orthogonal polynomials and their representations. In this preliminary training we will introduce some of those important algorithms: the automatic computation of differential equations and formal power series, hypergeometric representations, and the algorithms by Fasenmyer, Gosper, Zeilberger and Petkovšek/van Hoeij.

Keywords Computer algebra · Algorithms for power series · Algorithms for summation

Mathematics Subject Classification (2000) Primary 68W30, 33F10; Secondary 30B10, 33C20

1 Introduction

I will use the computer algebra system *Maple* to program and demonstrate the methods considered. Of course one could also easily use any other general-purpose system like *Mathematica*, Maxima, Reduce or Sage.

Such general-purpose computer algebra systems

- contain a high-level programming language,
- are dialog-oriented and not compiling,
- are able to plot functions,
- and can compute with symbols.

W. Koepf (✉)
Institute of Mathematics, University of Kassel, Kassel, Germany
e-mail: koepf@mathematik.uni-kassel.de

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_7

The mostly used algorithms are:

- algorithms of linear algebra with many variables and coefficients that are rational functions,
- multivariate polynomial factorization (over $\mathbb{Q}$),

which can be treated algorithmically and are efficiently available in all general-purpose computer algebra systems. High-end algorithms contain algorithms for modular arithmetic, algebraic numbers, solving polynomial systems, differentiation and integration, solving differential equations, as well as Taylor polynomials and power series.

2 Taylor Polynomials

The coefficients $B_k(x)$ of the power series

$$\sqrt{\frac{\left(\frac{1+z}{1-z}\right)^x - 1}{2xz}} = \sum_{k=0}^{\infty} B_k(x)\, z^k$$

are polynomials of degree k. In a 1978 publication Malcolm S. Robertson [16] conjectured that the coefficients of all these polynomials $B_k(x)$ are non-negative. He had computed $B_k(x)$ for $1 \leqq k \leqq 6$. This is reproduced by the *Maple* computation

```
> g:=sqrt((((1+z)/(1-z))^x-1)/(2*x*z));
```

$$g := 1/2\sqrt{2}\sqrt{\frac{1}{xz}\left(\left(\frac{1+z}{1-z}\right)^x - 1\right)}$$

```
> map(expand,series(g,z,7));
```

$$1 + 1/2\,xz + \left(\frac{5\,x^2}{24} + 1/6\right)z^2 + \left(1/16\,x^3 + x/4\right)z^3 + \left(\frac{31}{360} + \frac{25\,x^2}{144} + \frac{79\,x^4}{5760}\right)z^4$$
$$+ \left(\frac{41\,x}{240} + \frac{7\,x^3}{96} + \frac{3\,x^5}{1280}\right)z^5 + \left(\frac{863}{15120} + \frac{85\,x^2}{576} + \frac{79\,x^4}{3840} + \frac{71\,x^6}{193536}\right)z^6 + O\left(z^7\right)$$

Robertson repeated his conjecture in the article [17] published in 1989. In this publication Robertson further conjectured that the coefficients A_k of the univariate power series

$$\sqrt{\frac{e^x - 1}{x}} = \sum_{k=0}^{\infty} A_k\, x^k$$

are all positive. It turns out that both conjectures are false, as the following computations show![1]

```
> f:=sqrt((exp(x)-1)/x);
```

$$f := \sqrt{\frac{e^x - 1}{x}}$$

```
> series(f,x=0,14);
```

$$1 + x/4 + \frac{5\,x^2}{96} + \frac{x^3}{128} + \frac{79\,x^4}{92160} + \frac{3\,x^5}{40960} + \frac{71\,x^6}{12386304} + \frac{113\,x^7}{247726080} + \frac{3053\,x^8}{118908518400} + \frac{x^9}{22649241600} + \frac{17\,x^{10}}{930128855040} + \frac{19\,x^{11}}{744103084032} + \frac{935917\,x^{12}}{1218840851644416000} - \frac{20287103\,x^{13}}{43878270659198976000} + O\left(x^{14}\right)$$

```
> expand(coeff(series(g,z,14),z,13));
```

$$\frac{102672775873\,x}{1307674368000} + \frac{437\,x^{11}}{2179989504} - \frac{20287103\,x^{13}}{5356234211328000} + \frac{2299\,x^9}{15925248000} + \frac{3735911\,x^3}{54743040} + \frac{2031271\,x^5}{258048000} + \frac{2042249\,x^7}{5852528640}$$

By these computations both conjectures are disproved since the thirteenth coefficient of $f(x)$ is negative, and $B_{13}(x)$ also has a negative coefficient. The falsification of these Robertson's conjectures was published in 1991 [10]. Please notice that the reviewing process of my article almost 30 years ago was still very hard! The reviewers were complaining why we should believe in the output of a computer? Nowadays computer algebra computations are so common, and we know that they are safer than hand computations, that no reviewer will complain about such computations any more.

3 Holonomic Power Series

Sometimes one is not only interested in a Taylor polynomial approximation, but in the full Taylor series. Assume, given an expression $f(x)$ depending on the variable x, we would therefore like to compute a formula for the coefficient a_k of the power series

$$f(x) = \sum_{k=0}^{\infty} a_k\, x^k \tag{3.1}$$

representing $f(x)$. A well-known example of that type is given by

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} .$$

[1]Please note that *Maple* regularly does *not* sort its output to save computation time so that the Taylor coefficients might appear in wrong order.

Here is an algorithm for this purpose [11].

Algorithm 3.1 (FPS (Koepf [11]))

- **Input:** *expression* $f(x)$.
- **HolonomicDE:** *Determine a holonomic differential equation DE (i.e. homogeneous and linear with polynomial coefficients) by computing the derivatives of* $f(x)$ *iteratively.*
- **DEtoRE:** *Convert DE to a holonomic recurrence equation RE for* a_k.
- **RSolve:** *Solve RE for* a_k.
- **Output:** a_k *resp.* $\sum\limits_{k=0}^{\infty} a_k x^k$.

Functions satisfying a holonomic DE and sequences satisfying a holonomic RE are also called holonomic. A function is holonomic if and only if it is the generating function of a holonomic sequence.

The above algorithm is embedded into the *Maple* system via the `convert(..., FormalPowerSeries)` command, and we present some examples:

```
> convert(sin(x),FPS);
```

$$\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$$

```
> convert(arcsin(x)^2,FPS);
```

$$\sum_{k=0}^{\infty} \frac{(k!)^2 4^k x^{2k+2}}{(k+1)(2k+1)!}$$

How does this algorithm work? Here are the details:

Algorithm 3.2 (HolonomicDE (Koepf [11]))

- **Input:** *expression* $f(x)$.
- *Iterate for* $J = 1, 2, \ldots$*:*
- *Compute* $c_0 f(x) + c_1 f'(x) + \cdots + f^{(J)}(x)$ *with still undetermined coefficients* c_j.
- *Sort this linear combination w.r.t. linearly independent functions over* $\mathbb{Q}(x)$ *and determine their coefficients* $\in \mathbb{Q}(x)$.[2]
- *Set these coefficients zero, and solve the corresponding linear system for the unknowns* $c_0, c_1, \ldots, c_{J-1}$.
- **Output:** *DE* $:= c_0 f(x) + c_1 f'(x) + \cdots + f^{(J)}(x) = 0$ *(or else multiply by the common denominator of the* c_j*s).*

Holonomic functions have interesting algebraic properties. The existence of a holonomic differential equation shows that the dimension of the vector space

$$V_f = \langle f(x), f'(x), f''(x), \ldots \rangle$$

[2]For details how to select the linearly independent functions, see [8].

over the field of rational functions $\mathbb{Q}(x)$ is finite. This argument yields

Algorithm 3.3 ((Stanley [19]), Maple (Salvy and Zimmermann [18]))

- *Let a function $f(x)$ be given by a holonomic differential equation DE1 of order n, and let a function $g(x)$ be given by a holonomic differential equation DE2 of order m.*
- *Then there are linear algebra algorithms showing that $f+g$ is holonomic of degree $\leqq n+m$, and $f \cdot g$ is holonomic of degree $\leqq n \cdot m$.*
- *Similar statements are valid for holonomic sequences.*

By these algebraic properties the ring of holonomic functions has a so-called *normal form* which consists of the differential equation together with enough initial values. The same is true for the ring of holonomic sequences whose normal form consists of their recurrence equation together with enough initial values. Unfortunately, holonomic functions are not closed under division, an example of which is given by $\tan x = \frac{\sin x}{\cos x}$ (for details of the proof, see [12]).

The following computation computes the differential equation for $f(x) := \arcsin^2(x)$ and converts it towards the corresponding recurrence equation for its coefficients a_k for $f(x)$ given by (3.1):

```
> bind(FormalPowerSeries):
> DE:=HolonomicDE(arcsin(x)^2,F(x));
```

$$DE := \left\{\frac{d}{dx}F(x) + 3x\frac{d^2}{dx^2}F(x) + \left(x^2-1\right)\frac{d^3}{dx^3}F(x),\, F(0)=0, D(F)(0)=0, \left(D^{(2)}\right)(F)(0)=2\right\}$$

```
> RE:=SimpleRE(arcsin(x)^2,x,a(k));
```

$$RE := (k+1)^3\, a(k+1) - (k+1)(k+2)(k+3)\, a(k+3) = 0$$

which can be easily solved for a_k using the hypergeometric coefficient formula (3.3) that we will consider soon.

By computing their normal forms two functions can be identified as identical! The following computation shows, for example, the addition theorem of the sine function, completely automatically.

```
> HolonomicDE(sin(x+y),F(x));
```

$$\left\{\frac{d^2}{dx^2}F(x) + F(x),\, F(0) = \sin(y),\, D(F)(0) = \cos(y)\right\}$$

```
> HolonomicDE(sin(x)*cos(y)+cos(x)*sin(y),F(x));
```

$$\left\{\frac{d^2}{dx^2}F(x) + F(x),\, F(0) = \sin(y),\, D(F)(0) = \cos(y)\right\}$$

A very rich class of holonomic functions are the hypergeometric functions/series. The formal power series

$${}_pF_q\left(\begin{array}{c} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{array}\middle|\, z\right) = \sum_{k=0}^{\infty} A_k\, z^k = \sum_{k=0}^{\infty} \alpha_k\ ,$$

whose summands $\alpha_k = A_k z^k$ have a rational term ratio (factored over $\mathbb{Q}$ or, if necessary, over $\mathbb{C}$)

$$\frac{\alpha_{k+1}}{\alpha_k} = \frac{A_{k+1}\, z^{k+1}}{A_k\, z^k} = \frac{(k+a_1)\cdots(k+a_p)}{(k+b_1)\cdots(k+b_q)}\frac{z}{(k+1)} ,$$

is called the *generalized hypergeometric series*. For this reason, the summand $\alpha_k = A_k z^k$ of a hypergeometric series is called a *hypergeometric term*.

For the coefficients of the generalized hypergeometric function one gets the following formula using the shifted factorial (*Pochhammer symbol*) $(a)_k = a(a+1) \cdots (a+k-1)$

$${}_pF_q\left(\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix} \middle| z\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!} .$$

Therefore every holonomic recurrence equation of first order defines a hypergeometric term which can be determined by the hypergeometric coefficient formula.

If the recurrence is brought into the form

$$\frac{\alpha_{k+1}}{\alpha_k} = \frac{(k+a_1)\cdots(k+a_p)}{(k+b_1)\cdots(k+b_q)}\frac{z}{(k+1)} ,$$

or, equivalently

$$(k+b_1)\cdots(k+b_q)(k+1)\,\alpha_{k+1} - (k+a_1)\cdots(k+a_p)\,z\,\alpha_k = 0 , \tag{3.2}$$

then α_k is given by

$$\alpha_k = \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!} \cdot \alpha_0 . \tag{3.3}$$

The above method provides an algorithm to detect the hypergeometric representation for

$$S = \sum_{k=0}^{\infty} \alpha_k .$$

Algorithm 3.4 (Sumtohyper (Koepf [13]))

- **Input:** *α_k, a hypergeometric term w.r.t. k.*
- *Compute*

$$r_k := \frac{\alpha_{k+1}}{\alpha_k} \in \mathbb{Q}(k) .$$

- *Factorize r_k.*
- **Output:** *Reading off upper and lower parameters and the argument z from representation (3.2), and computing an initial value yields the corresponding hypergeometric series.*

As an example, for

$$\sin x = \sum_{k=0}^{\infty} \alpha_k = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} ,$$

we get

$$r_k = \frac{\alpha_{k+1}}{\alpha_k} = -\frac{x^2}{2(k+1)(2k+3)} = \frac{1}{(k+3/2)(k+1)} \left(-\frac{x^2}{4}\right) \in \mathbb{Q}(x,k)$$

and

$$\alpha_0 = x \ .$$

Therefore

$$\sin x = x \cdot {}_0F_1\left(\begin{array}{c} - \\ 3/2 \end{array} \middle| -\frac{x^2}{4}\right) .$$

After loading the `hsum` package [13]

```
> read "hsum17.mpl";
   `Package "Hypergeometric Summation", Maple V - Maple 17`
  `Copyright 1998-2014, Wolfram Koepf, University of Kassel`
```

this computation is done automatically by the *Maple* command

```
> Sumtohyper((-1)^k*x^(2*k+1)/(2*k+1)!,k);
```

$$x\mathit{Hypergeom}\left([], [3/2], -1/4\,x^2\right)$$

4 Fasenmyer's Algorithm

The sum evaluation

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

is well-known. However, the similar, but more advanced identities

$$\sum_{k=0}^{n} k \binom{n}{k} = n\, 2^{n-1} \quad \text{and} \quad \sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

might be less known. Our question is: How can we compute the right hand sides, given the sums on the left?

In general, consider the sum

$$S_n := \sum_{k=0}^{n} F(n,k) \quad \text{with} \quad \frac{F(n+1,k)}{F(n,k)}, \frac{F(n,k+1)}{F(n,k)} \in \mathbb{Q}(n,k),$$

where $F(n,k)$ therefore is a *hypergeometric term* w r. t. k and n. Fasenmyer's idea [5, 6] is to find first a *k-free recurrence for the summand $F(n,k)$*:

$$\sum_{j=0}^{J} \sum_{i=0}^{I} a_{ij}\, F(n+j, k+i) = 0 \quad \text{for some } a_{ij} \in \mathbb{Q}(n) \ .$$

This can be done by linear algebra!

If successful, since the recurrence is k-free, it can be summed from $k = -\infty, \ldots, \infty$ to get a *pure recurrence for the sum S_n*.

For

$$F(n,k) = \binom{n}{k} \qquad \text{and} \qquad S_n = \sum_{k=0}^{n} \binom{n}{k} = \sum_{k=-\infty}^{\infty} \binom{n}{k}$$

the first step yields the *Pascal triangle identity*

$$F(n+1, k+1) = F(n, k+1) + F(n,k) \ .$$

If we sum this relation from $k = -\infty, \ldots, \infty$, we clearly get

$$S_{n+1} = S_n + S_n = 2\, S_n \ .$$

Since $S_0 = 1$, we finally have $S_n = 2^n$. The corresponding *Maple* command is given by

```
> fasenmyer(binomial(n,k),k,S(n),1);
```

$$S(n+1) - 2\, S(n) = 0$$

or—knowing the result—by

```
> fasenmyer(binomial(n,k)/2^n,k,S(n),1);
```

$$S(n+1) - S(n) = 0$$

and other examples are

```
> fasenmyer(k*binomial(n,k),k,S(n),1);
```

$$nS(n+1) - 2\,(n+1)\,S(n) = 0$$

```
> fasenmyer(k*binomial(n,k)/(n*2^(n-1)),k,S(n),1);
```

$$S(n+1) - S(n) = 0$$

```
> fasenmyer(binomial(n,k)^2,k,S(n),2);
```

$$(n+2)\,S(n+2) - 2\,S(n+1)\,(2\,n+3) = 0$$

```
> fasenmyer(binomial(n,k)^2/binomial(2*n,n),k,S(n),2);
```

$$S(n+2) - S(n+1) = 0$$

Fore more details, see [13, Chapter 4].

5 Gosper's Algorithm

Given a_k, a sequence s_k is called an *antidifference* of a_k if

$$a_k = s_{k+1} - s_k\ .$$

Given such an antidifference, then by telescoping the sum can be computed with arbitrary lower and upper bounds:

$$\sum_{k=m}^{n} a_k = (s_{n+1} - s_n) + (s_n - s_{n-1}) + \cdots + (s_{m+1} - s_m)$$
$$= s_{n+1} - s_m\ .$$

Gosper's algorithm computes a *hypergeometric term* antidifference s_k for a hypergeometric term a_k, i.e.

$$\frac{a_{k+1}}{a_k} \in \mathbb{Q}(k)\ .$$

Here are the computational steps: Given a_k with

$$\frac{a_{k+1}}{a_k} = \frac{u_k}{v_k}$$

with polynomials $u_k, v_k \in \mathbb{Q}[k]$, $\gcd(u_k, v_k) = 1$, it is easy to show that

$$s_k = \frac{g_k}{h_k}\, a_k$$

for certain polynomials $g_k, h_k \in \mathbb{Q}[k]$.

Whereas the denominator h_k can be written down explicitly, for g_k one gets the inhomogeneous recurrence equation

$$h_k\, u_k\, g_{k+1} - h_{k+1}\, v_k\, g_k = h_k\, h_{k+1}\, v_k \ .$$

To compute g_k one first finds a *degree bound* for g_k. The final step uses *linear algebra* to compute the coefficients of g_k by equating coefficients.

Gosper's algorithm [13, Chapter 5] was maybe the first algorithm which would not have been found without computer algebra. In his paper [7], Gosper stated:

> Without the support of MACSYMA and its developers, I could not have collected the experiences necessary to provoke the conjectures that led to this algorithm.

The next two examples for an application of Gosper's algorithm

```
> gosper((-1)^k*binomial(n,k),k);
```

$$-\frac{k\,(-1)^k \binom{n}{k}}{n}$$

```
> gosper(1/(k+1),k);
```

Error, (in gosper) No hypergeometric term antidifference exists

show that

- An antidifference of $a_k = (-1)^k \binom{n}{k}$ is $s_k = -\frac{k}{n}\, a_k$.
- The harmonic numbers $H_n = \sum_{k=1}^n \frac{1}{k}$ cannot be written as hypergeometric term.

Fore more details, see [12, Chapter 11] and [13, Chapter 5].

6 Zeilberger's Algorithm

Zeilberger's algorithm deals again—like Fasenmyer's—with definite sums of the form

$$S_n := \sum_{k=0}^{n} F(n,k) \quad \text{with} \quad \frac{F(n+1,k)}{F(n,k)},\ \frac{F(n,k+1)}{F(n,k)} \in \mathbb{Q}(n,k).$$

Applying Gosper to $F(n,k)$ w. r. t. k (by telescoping) always generates the identity $S_n = 0$, if it is successful. Therefore this method generally cannot be applied.

Doron Zeilberger's idea [13, Chapter 7]: Apply Gosper instead to

$$a_k := F(n,k) + \sum_{j=1}^{J} \sigma_j\, F(n+j,k) \quad \text{for some} \quad J \geqq 1 \ .$$

Doron Zeilberger's idea generates the following algorithm.

Algorithm 6.1 (Zeilberger [21])

- *Apply Gosper's algorithm to*

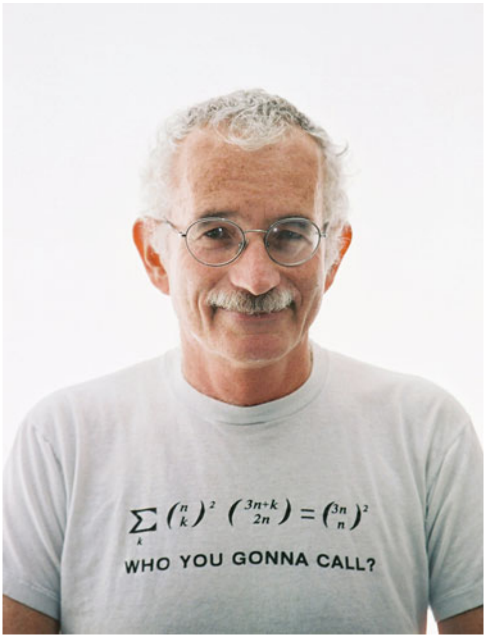
$$a_k := F(n,k) + \sum_{j=1}^{J} \sigma_j \, F(n+j,k) \quad \textit{for some } J \geqq 1 \; .$$

- *In the final linear algebra step, solve not only for the coefficients of g_k, but also for σ_j $(j = 1, \ldots, J)$.*
- *If successful, then this algorithm generates the recurrence equation*

$$S_n + \sum_{j=1}^{J} \sigma_j \, S_{n+j} = 0$$

  with rational functions $\sigma_j \in \mathbb{Q}(n)$ $(j = 1, \ldots, J)$ for S_n.
- *Multiplication with the common denominator yields a holonomic recurrence equation for S_n.*

If the resulting recurrence equation is of first order, then an application of Algorithm 3.4 yields the hypergeometric term solution.



Reprinted with kind permission of Tamar Legrange Zeilberger

We give an example:

```
> sumrecursion(binomial(n,k)^2,k,S(n));
```

$$-(n+1)\,S(n+1)+2\,(2\,n+1)\,S(n)=0$$

Since the resulting recurrence equation is of first order, we can solve it explicitly by the command

```
> closedform(binomial(n,k)^2,k,n);
```

$$\frac{(2\,n)!}{(n!)^2}$$

The next example solves the problem given on Doron Zeilberger's T-shirt.

```
> closedform(binomial(n,k)^2*binomial(3*n+k,2*n),k,S(n));
```

$$\frac{(\mathit{pochhammer}(2/3,n))^2\,(\mathit{pochhammer}(1/3,n))^2\,(4^n)^2}{((2\,n)!)^2}\left(\frac{729}{16}\right)^n$$

Of course this result looks rather different from the right hand side on Zeilberger's T-shirt although the results are equivalent. But since we know the desired result, we can prove it easily by the computation

```
> closedform(binomial(n,k)^2*binomial(3*n+k,2*n)/
> binomial(3*n,n)^2,k,n);
```

$$1$$

Fore more details, see [13, Chapter 7].

7 CAOP

CAOP [4] is a web tool for calculating formulas for orthogonal polynomials belonging to the Askey–Wilson scheme using Maple. The implementation of CAOP was originally done by René Swarttouw as part of the Askey–Wilson Scheme Project performed at RIACA in Eindhoven in 2004. The present site http://www.caop.org/ is a completely revised version of this project which has been done by Torsten Sprenger under my supervision in 2012 and is maintained at the University of Kassel.

If we select, for example, the *Laguerre polynomials*, the CAOP system will let us know that they are defined as

$$L_n^{(\alpha)}(x)=\sum_{k=0}^{n}\binom{n+\alpha}{n-k}\frac{(-x)^k}{k!}=\frac{(\alpha+1)_n}{n!}\,{}_1F_1\left(\left.\begin{matrix}-n\\ \alpha+1\end{matrix}\right|x\right),$$

both in binomial series and in hypergeometric notation. The system will immediately start to compute both a recurrence equation for $L_n^{(\alpha)}(x)$ w.r.t. n and a differential equation w.r.t. x using Zeilberger type algorithms. All computations are done on a server at the University of Kassel. The complete *Maple* codes which are used can be displayed. In principle, one can even multiply $P_n(x) = L_n^{(\alpha)}(x)$ by an arbitrary factor and redo the computations. For $L_n^{(\alpha)}(x)$, one gets

$$x\,P_n''(x) + (-x+\alpha+1)\,P_n'(x) + n\,P_n(x) = 0$$

and

$$(n+2)\,P_{n+2}(x) - (-x+2\,n+\alpha+3)\,P_{n+1}(x) + (n+\alpha+1)\,P_n(x) = 0\,.$$

8 Petkovšek's and van Hoeij's Algorithm

Assume, S_n is a hypergeometric term, however Zeilberger's algorithm generates not the recurrence of first order, but a holonomic recurrence of higher order

$$RE : \sum_{j=0}^{J} P_j(n)\,S_{n+j} = 0$$

with polynomial coefficients $P_j \in \mathbb{Q}[n]$.

Then *Petkovšek's algorithm* ([15], see [13, Chapter 9])—which is quite similar to Gosper's—finds every solution which is a linear combination of hypergeometric terms. Combining Zeilberger's with Petkovšek's algorithm leads to a decision procedure for hypergeometric term summation.

We want to compute

$$S_n = \sum_{k=0}^{n} \frac{\binom{3k+1}{k}\binom{3n-3k}{n-k}}{3k+1}\,.$$

First we realize that the summation bounds $k = 0, \ldots, n$ are the natural ones since $-n$ is an upper parameter of the corresponding hypergeometric representation, shown by the computation

```
> summand:=binomial(3*k+1,k)*binomial(3*n-3*k,n-k)/(3*k+1);
```

$$summand := \frac{\binom{3\,k+1}{k}\binom{3\,n-3\,k}{n-k}}{3\,k+1}$$

```
> Sumtohyper(summand,k);
```

$$\frac{\Gamma(3\,n+1)\,Hypergeom([-n, 1/3, 2/3, -n+1/2], [-n+2/3, -n+1/3, 3/2], 1)}{\Gamma(2\,n+1)\,\Gamma(n+1)}$$

By

```
> RE:=sumrecursion(summand,k,S(n));
```

$$RE := 4\,(2\,n+5)(2\,n+3)(n+2)\,S(n+2) - 12\,(2\,n+3)\left(9\,n^2+27\,n+22\right)S(n+1)$$
$$+\,81\,(3\,n+2)(3\,n+4)(n+1)\,S(n) = 0$$

we therefore have computed a recurrence equation for S_n. Applying Petkovšek's algorithm yields two linearly independent hypergeometric term solutions with term ratios S_{n+1}/S_n given by

```
> rechyper(RE,S(n));
```

$$\left\{27\,\frac{n+1}{4\,n+6},\,3/2\,\frac{(3\,n+2)\,(3\,n+4)}{(2\,n+3)\,(n+1)}\right\}$$

Therefore S_n must be a linear combination of the corresponding hypergeometric terms

```
> result:=a*pochhammer(1,n)/pochhammer(3/2,n)*(27/4)^n
> +b*pochhammer(2/3,n)*pochhammer(4/3,n)/
> (pochhammer(1,n)*pochhammer(3/2,n))*(27/4)^n;
```

$$result := \frac{a\,pochhammer\,(1,n)}{pochhammer\,(3/2,n)}\left(\frac{27}{4}\right)^n + \frac{b\,pochhammer\,(2/3,n)\,pochhammer\,(4/3,n)}{pochhammer\,(1,n)\,pochhammer\,(3/2,n)}\left(\frac{27}{4}\right)^n$$

Using two initial values, we can compute a and b

```
> init0:=eval(add(subs(n=0,summand),k=0..0));
```

$$init0 := 1$$

```
> init1:=eval(add(subs(n=1,summand),k=0..1));
```

$$init1 := 4$$

```
> sol:=solve({eval(result,n=0)=init0,eval(result,n=1)
>    =init1},{a,b});
```

$$sol := \{a = 0, b = 1\}$$

so that we finally get for S_n

```
> res1:=subs(sol,result);
```

$$res1 := \frac{pochhammer\,(2/3,n)\,pochhammer\,(4/3,n)}{pochhammer\,(1,n)\,pochhammer\,(3/2,n)}\left(\frac{27}{4}\right)^n$$

We claim that

$$S_n = \binom{3n+1}{n},$$

which is proved by

```
> res2:=binomial(3*n+1,n);
```

$$res2 := \binom{3\,n+1}{n}$$

```
> simplify(res1/res2);
```

$$1$$

However, Petkovšek's algorithm is rather inefficient. Mark van Hoeij ([20], see [13, Chapter 9]), gave a much faster algorithm for the same purpose, based on *the singularity behavior* of the constituting Γ functions that occur in the solutions. Details can be found in [13, Chapter 9].

Mark van Hoeij implemented his algorithm in *Maple*. Applying this implementation to the same example as above, we get

```
> res:=LREtools[hypergeomsols](RE,S(n),{S(0)=1,S(1)=4});
```

$$res := 3/4\,\frac{\sqrt{3}\Gamma\,(4/3+n)\,\Gamma\,(2/3+n)}{\sqrt{\pi}\Gamma\,(n+3/2)\,\Gamma\,(n+1)}\left(\frac{27}{4}\right)^{n}$$

which gives—as announced—the result in terms of the Gamma function. We can again show that the result constitutes the same term as above by

```
> simplify(res/res2);
```

$$1$$

9 Hypergeometric Identities

All examples from this section can be found in [13, Chapter 7].

Clausen's formula gives the cases when a Clausen ${}_3F_2$ function is the square of a Gauss ${}_2F_1$ function:

$$\left({}_2F_1\left(\begin{array}{c} a, b \\ a+b+\frac{1}{2} \end{array}\middle|\, x\right)\right)^2 = {}_3F_2\left(\begin{array}{c} 2a, 2b, a+b \\ 2a+2b, a+b+\frac{1}{2} \end{array}\middle|\, x\right).$$

The right hand side can be detected from the left hand side by Zeilberger's algorithm. Using the Cauchy product, we get for the coefficient of the left hand square:

```
> sumrecursion(hyperterm([a,b],[a+b+1/2],x,k)*
> hyperterm([a,b],[a+b+1/2],x,n-k),k,S(n));
```

$$-(n+1)\,(n+2\,b+2\,a)\,(2\,a+2\,b+1+2\,n)\,S\,(n+1)$$
$$+\,2\,x\,(n+2\,b)\,(2\,a+n)\,(a+b+n)\,S\,(n)=0$$

which is of first order so that we can easily convert the summand into the corresponding hypergeometric term by the command

```
> Closedform(hyperterm([a,b],[a+b+1/2],x,k)*
> hyperterm([a,b],[a+b+1/2],x,n-k),k,n);
```

$$Hyperterm\,([2\,a, 2\,b, a+b], [2\,b+2\,a, a+b+1/2], x, n)$$

This proves Clausen's formula from left to right. The question remains, how we can get the full identity without prior knowledge about the parameter choice? This can be done using differential equations, by solving a non-linear system. The generic function ${}_2F_1(a, b; c; x)$ satisfies the differential equation

```
> DE2F1:=sumdiffeq(hyperterm([a,b],[c],x,k),k,C(x));
```

$$DE2F1 := \left(\frac{d^2}{dx^2}C\,(x)\right)x\,(x-1)+(ax+bx-c+x)\,\frac{d}{dx}C\,(x)+abC\,(x)=0$$

and therefore its square satisfies[3]

```
> DE2F1square:=map(factor,gfun['diffeq*diffeq'](DE2F1,
>   DE2F1,C(x))):
```

On the other hand, the generic function ${}_3F_2(A, B, C; D, E; x)$ satisfies the differential equation

```
> DE3F2:=op(1,sumdiffeq(hyperterm([A,B,C],[D,E],x,k),k,C(x))):
```

Since the coefficient of the highest derivative in the left hand differential equation is $(x - 1)$ times the coefficient of the highest derivative in the right hand differential equation, we multiply the right hand side by $(x - 1)$ and subtract to get zero.

```
> pol:=expand(subs(diff(C(x),x$3)=y^3,diff(C(x),x$2)=y^2,
> diff(C(x),x)=y,C(x)=1,DE3F2*(x-1)-DE2F1square)):
```

Next, by equating the coefficients we can set all coefficients zero, assuming that a and b are arbitrary. Therefore we solve for the remaining variables c, A, B, C, D, and E and get

```
> solve({coeffs(expand(pol),[x,y])},{c,A,B,C,D,E});
```

$$\{A = 2b, B = a + b, C = 2a, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = 2a, B = a + b, C = 2b, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = 2b, B = a + b, C = 2a, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\},$$
$$\{A = 2a, B = a + b, C = 2b, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\},$$
$$\{A = 2b, B = 2a, C = a + b, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = a + b, B = 2a, C = 2b, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = 2b, B = 2a, C = a + b, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\},$$
$$\{A = a + b, B = 2a, C = 2b, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\},$$
$$\{A = a + b, B = 2b, C = 2a, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = 2a, B = 2b, C = a + b, D = 2b + 2a, E = a + b + 1/2, c = a + b + 1/2\},$$
$$\{A = a + b, B = 2b, C = 2a, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\},$$
$$\{A = 2a, B = 2b, C = a + b, D = a + b + 1/2, E = 2b + 2a, c = a + b + 1/2\}$$

Please notice that all twelve solutions are equivalent since the upper parameters A, B, C as well as the lower parameters D, E can be permuted arbitrarily.

Dougall (1907) found the following astonishing identity (see [12], Table 6.1, p. 108)

$$
{}_7F_6\left(\left.\begin{array}{c} a, 1+\frac{a}{2}, b, c, d, 1+2a-b-c-d+n, -n \\ \frac{a}{2}, 1+a-b, 1+a-c, 1+a-d, b+c+d-a-n, 1+a+n \end{array}\right| 1\right)
$$
$$
= \frac{(1+a)_n(a+1-b-c)_n(a+1-b-d)_n(a+1-c-d)_n}{(1+a-b)_n(1+a-c)_n(1+a-d)_n(1+a-b-c-d)_n}\,.
$$

[3]We don't print the lengthy outputs.

Again, if one knows the left-hand side, Zeilberger's algorithm generates the right-hand side immediately.

```
> Closedform(hyperterm([a,1+a/2,b,c,d,1+2*a-b-c-d+n,-n],
> [a/2,1+a-b,1+a-c,1+a-d,b+c+d-a-n,1+a+n],1,k),k,n);
```

$$\mathit{Hyperterm}([a+1, -c+1+a-b, 1+a-d-b, 1+a-d-c, 1],\\ [1+a-d, -b+1+a, -c+1+a, -b-c-d+a+1], 1, n)$$

Apéry [2] proved that

$$\zeta(3) = \sum_{j=1}^{\infty} \frac{1}{j^3}$$

is irrational. In his proof, he used a holonomic recurrence equation for the so-called *Apéry numbers*

$$A_n := \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 .$$

Again, Zeilberger's algorithm detects their recurrence equation immediately.

```
> sumrecursion(binomial(n,k)^2*binomial(n+k,k)^2,k,A(n));
```

$$(n+2)^3 A(n+2) - (2n+3)\left(17n^2+51n+39\right) A(n+1) + (n+1)^3 A(n) = 0$$

Shortly before Zeilberger's algorithm came up, Perlstadt [14] was able to publish a paper deriving recurrence equations for

$$S_n^{(r)} := \sum_{k=0}^{n} \binom{n}{k}^r$$

for $r = 5$ and 6. Using Zeilberger's algorithm, this is now "trivial".

```
> TIME:=time():
> sumrecursion(binomial(n,k)^5,k,S(n));
> time()-TIME;
```

$$\left(55n^2+143n+94\right)(n+3)^4 S(n+3)\\ -\left(1155n^6+14553n^5+75498n^4+205949n^3+310827n^2+245586n+79320\right) S(n+2)\\ -\left(19415n^6+205799n^5+900543n^4+2082073n^3+2682770n^2+1827064+514048\right) S(n+1)\\ +32\left(55n^2+253n+292\right)(n+1)^4 S(n) = 0$$

0.140

For the next examples, we omit again the lengthy outputs.

```
> TIME:=time():
> sumrecursion(binomial(n,k)^6,k,S(n)):
> time()-TIME;
```

0.125

```
>   TIME:=time():
>   sumrecursion(binomial(n,k)^10,k,S(n)):
>   time()-TIME;
```

$$3.016$$

Even for power 10, our computation needs only few seconds although the resulting recurrence equation of order 5 has *very complicated* coefficients.

10 Generating Functions

Assume, we want to compute the *generating function* of the Legendre polynomials

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} \binom{-n-1}{k} \left(\frac{1-x}{2}\right)^k ,$$

namely

$$F(z) := \sum_{n=0}^{\infty} P_n(x)\, z^n \ .$$

How can we approach this question?

The recurrence equation for the coefficients (the Legendre polynomials) can be converted into a differential equation for the generating function. In certain cases, this differential equation can be solved.

```
>   RE:=sumrecursion(binomial(n,k)*binomial(-n-1,k)*((1-x)/2)^k,
>      k,P(n));
```

$$RE := (n+2)\, P\,(n+2) - x\,(2\,n+3)\, P\,(n+1) + (n+1)\, P\,(n) = 0$$

```
>   DE:=gfun[rectodiffeq]({RE,P(0)=1,P(1)=x},P(n),F(z));
```

$$DE := \left\{(-x+z)\, F\,(z) + \left(-2\,xz + z^2 + 1\right) \frac{d}{dz} F\,(z)\, ,\, F\,(0) = 1\right\}$$

```
>   dsolve(DE,F(z));
```

$$F\,(z) = \left(\sqrt{-2\,xz + z^2 + 1}\right)^{-1}$$

Assume, we want to compute the *exponential generating function* of the Legendre polynomials

$$G(z) := \sum_{n=0}^{\infty} \frac{P_n(x)}{n!}\, z^n \ .$$

How can we approach this question?

Although, in principle, one can use the same method here (since *Maple*'s `dsolve` is really very powerful!), in this case, we apply a different approach. For this purpose, we use one particular hypergeometric representation

$$P_n(x) = x^n \, {}_2F_1\left(\begin{matrix} -\frac{n}{2}, -\frac{n-1}{2} \\ 1 \end{matrix} \,\middle|\, 1 - \frac{1}{x^2} \right)$$

of the Legendre polynomials. Then after interchanging the order of summation, Zeilberger's algorithm yields the generating function:

$$G(z) = e^{xz} \, {}_0F_1\left(\begin{matrix} - \\ 1 \end{matrix} \,\middle|\, \frac{z^2\,(x-1)^2}{4} \right) = e^{xz} \, I_0(z\sqrt{1-x^2}) \, ,$$

$I_n(z)$ denoting the modified Bessel function, by the computations

```
> summand:=x^n*hyperterm([-n/2,(1-n)/2],[1],1-1/x^2,k)/n!*z^n;
```

$$summand := \frac{x^n pochhammer\,(-n/2, k)\, pochhammer\,(1/2 - n/2, k)\, z^n}{(k!)^2\, n!} \left(1 - x^{-2}\right)^k$$

After interchanging the order of summation, we get for the summand of the outer sum the first order recurrence

```
> RE:=sumrecursion(summand,n,S(k));
```

$$RE := -4\, (k+1)^2\, S\,(k+1) + (x-1)\,(x+1)\, z^2 S\,(k) = 0$$

and the initial value

```
> init:=sum(subs(k=0,summand),n=0..infinity);
```

$$init := \mathrm{e}^{xz}$$

```
> outersummand:=rsolve({RE,S(0)=init},S(k));
```

$$outersummand := \frac{\left(1/4\, \left(x^2-1\right) z^2\right)^k \mathrm{e}^{xz}}{(\Gamma\,(k+1))^2}$$

```
> hyper:=Sumtohyper(outersummand,k);
```

$$hyper := \mathrm{e}^{xz} Hypergeom\left([], [1], 1/4\, (x-1)\,(x+1)\, z^2\right)$$

```
> convert(subs(Hypergeom=hypergeom,hyper),StandardFunctions);
```

$$\mathrm{e}^{xz} I_0\left(\sqrt{x-1}\sqrt{x+1} z\right)$$

11 Almkvist–Zeilberger Algorithm

Similarly to definite summation, Almkvist and Zeilberger [1] generated algorithms to compute holonomic recurrence and differential equations for definite integrals instead of series. In this case

$$S_n = \int_a^b F(n,t)\, dt$$

or

$$S(x) = \int_a^b F(x,t)\,dt$$

with F, hyperexponential w. r. t. the continuous variables x and t and hypergeometric w. r. t. the discrete variable n, for suitably chosen bounds a and b.

As an example, we would like to compute the definite integral

$$J(x) := \int_0^\infty \exp\left(-\frac{x^2}{t^2} - t^2\right) dt\ .$$

Almkvist–Zeilberger's algorithm shows that $J(x)$ satisfies the holonomic differential equation

$$J''(x) - 4\,J(x) = 0\ .$$

From this differential equation, using two initial values, we can deduce

$$J(x) = \frac{\sqrt{\pi}}{2}\,e^{-2x}\ .$$

By the Cauchy integral formula

$$f^{(n)}(x) = \frac{n!}{2\pi i}\int_\gamma \frac{f(t)}{(t-x)^{n+1}}\,dt\ , \tag{11.1}$$

a *Rodrigues formula* concerning the nth derivative can be written in terms of a definite integral.

Therefore, using the Almkvist–Zeilberger algorithm, one can deduce from the following Rodrigues representation

$$P_n(x) = \frac{1}{2^n\,n!}\frac{d^n}{dx^n}\left((x^2-1)^n\right) \tag{11.2}$$

again the recurrence and differential equations of the Legendre polynomials by the computations[4]

```
> rodriguesrecursion(1/(2^n*n!),(x^2-1)^n,x,P(n));
```

$$(n+2)\,P\,(n+2) - x\,(2\,n+3)\,P\,(n+1) + (n+1)\,P\,(n) = 0$$

```
> rodriguesdiffeq(1/(2^n*n!),(x^2-1)^n,n,P(x));
```

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}P\,(x)\right)(x-1)\,(x+1) + 2\,x\frac{\mathrm{d}}{\mathrm{d}x}P\,(x) - n\,(n+1)\,P\,(x) = 0$$

[4] The commands `rodriguesrecursion` and `rodriguesdiffequ` take representation (11.1) into consideration and invoke the Almkvist–Zeilberger algorithms. For details, see [13, Chapter 13].

therefore identifying those polynomials given by (11.2) as the Legendre polynomials.

If $F(z)$ is the generating function of the sequence $a_n\, f_n(x)$, i.e.

$$F(z) = \sum_{n=0}^{\infty} a_n\, f_n(x)\, z^n \ ,$$

then by Cauchy's integral formula and Taylor's theorem, we have

$$f_n(x) = \frac{1}{a_n} \cdot \frac{F^{(n)}(0)}{n!} = \frac{1}{a_n} \cdot \frac{1}{2\pi i} \int_\gamma \frac{F(t)}{t^{n+1}}\, dt \ .$$

Therefore, if the generating function is hyperexponential, we can again apply the Almkvist–Zeilberger algorithm to deduce recurrence and differential equations. Starting from the generating function

$$F(z) = \frac{1}{\sqrt{1 - 2xz + z^2}}$$

we can therefore compute the recurrence and differential equations for the Legendre polynomials

```
> F:=1/sqrt(1-2*x*z+z^2);
```

$$F := \left(\sqrt{-2\,xz + z^2 + 1}\right)^{-1}$$

```
> RE:=GFrecursion(F,1,z,P(n));
```

$$RE := (n+2)\, P\,(n+2) - x\,(2\,n+3)\, P\,(n+1) + (n+1)\, P\,(n) = 0$$

```
> DE:=GFdiffeq(F,1,z,n,P(x));
```

$$DE := \left(\frac{d^2}{dx^2} P\,(x)\right)(x-1)\,(x+1) + 2\,x \frac{d}{dx} P\,(x) - n\,(n+1)\, P\,(x) = 0$$

from their generating function.

These and many more examples are given in detail in [13, Chapter 13].

12 Basic Hypergeometric Series

As can be seen on the CAOP web site, the Askey–Wilson scheme consists of

1. the classical continuous orthogonal polynomial systems (OPS),
2. the classical discrete OPS on the lattice $\mathbb{Z}$ (or $\mathbb{N}_{\geq 0}$),
3. the Askey–Wilson polynomials on a quadratic lattice,
4. the OPS of the Hahn class on the basic lattice $q^{\mathbb{Z}}$,
5. and the Askey–Wilson polynomials on a q-quadratic lattice.

All details can be found in [9]. The first three types are all represented by hypergeometric series, whereas the last two have *basic hypergeometric* representations. For this purpose, one defines the basic hypergeometric series by

$$
{}_r\phi_s\left(\begin{array}{c} a_1, a_2, \ldots, a_r \\ b_1, b_2, \ldots, b_s \end{array} \middle| \, q\,;\, x \right) := \sum_{k=0}^{\infty} \frac{(a_1, a_2, \ldots, a_r;\, q)_k}{(b_1, b_2, \ldots, b_s;\, q)_k} \frac{x^k}{(q;\, q)_k} \left((-1)^k q^{\binom{k}{2}}\right)^{1+s-r}
$$

where $(a_1, a_2, \ldots, a_r;\, q)_k$ is a short form for the product $\prod_{j=1}^{r} \left(a_j;\, q\right)_k$, and

$$
(a;\, q)_k := \begin{cases} \prod_{j=0}^{k-1} \left(1 - a\, q^j\right) & \text{if } k > 0 \\ 1 & \text{if } k = 0 \\ \prod_{j=1}^{|k|} \left(1 - a\, q^{-j}\right)^{-1} & \text{if } k < 0 \\ \prod_{j=0}^{\infty} \left(1 - a\, q^j\right) & \text{if } k = \infty \text{ (if } |q| < 1) \end{cases}
$$

denotes the *q-Pochhammer symbol*. The relation

$$
{}_2\phi_1\left(\begin{array}{c} q^{-n}, b \\ c \end{array} \middle| \, q\,;\, \frac{c\, q^n}{b} \right) = \frac{(c/b;\, q)_n}{(c;\, q)_n}
$$

is a q-variant of the *Chu–Vandermonde Identity*

$$
\sum_{k=0}^{N} \binom{m}{k} \binom{n}{N-k} = \binom{n}{N} {}_2F_1\left(\begin{array}{c} -m, -N \\ n+1-N \end{array} \middle| \, 1 \right) = \binom{m+n}{N} .
$$

Applying Zeilberger's algorithm to the latter and its q-variant to the former yields these results from left to right.

```
> result1:=closedform(binomial(m,k)*binomial(n,N-k),k,N);
```

$$
\mathit{result1} := \frac{\mathit{pochhammer}\,(-n-m, N)\,(-1)^N}{N!}
$$

```
> result2:=closedform(binomial(m,k)*binomial(n,N-k)/
> binomial(m+n,N),k,N);
```

$$
\mathit{result2} := 1
$$

```
> convert(Sumtohyper(binomial(m,k)*binomial(n,N-k),k),binomial);
```

$$
\mathit{Hypergeom}\,([-N, -m], [n+1-N], 1) \binom{n}{N}
$$

We load the qsum package (see [3] and [13])

```
> read "qsum17.mpl";
```

    'Package "q-Hypergeometric Summation", Maple V-17'
    'Copyright 1998-2013, Harald Boeing,
    Wolfram Koepf & Torsten Sprenger, University of Kassel'

and get using the q-variant of Zeilberger's algorithm ([13, Chapter 7])

```
> RE:=qsumrecursion(qphihyperterm([q^(-n),b],[c],q,c*q^n/b,k),
> q,k,CV(n));
```

$$RE := -b\left(cq^n - q\right) CV(n) + \left(cq^n - bq\right) CV(n-1) = 0$$

```
> qsumrecursion(qphihyperterm([q^(-n),b],[c],q,c*q^n/b,k),q,k,
> CV(n),rec2qhyper=true);
```

$$\left[CV(n) = \frac{1}{qpochhammer\,(c,q,n)} qpochhammer\left(\frac{c}{b},q,n\right), 0 \le n\right]$$

```
> qsumrecursion(qphihyperterm([q^(-n),b],[c],q,c*q^n/b,k)/
> (qpochhammer(c/b,q,n)/qpochhammer(c,q,n)),q,k,CV(n),
> rec2qhyper=true);
```

$$[CV(n) = 1, 0 \le n]$$

As examples of OPS of the q-Hahn class, we consider the *Little* and *Big q-Legendre polynomials*:

$$p_n(x|q) = {}_2\phi_1\left(\begin{matrix} q^{-n}, q^{n+1} \\ q \end{matrix} \middle| q\,;\, q\,x\right)$$

and

$$P_n(x;c;q) = {}_3\phi_2\left(\begin{matrix} q^{-n}, q^{n+1}, x \\ q, c\,q \end{matrix} \middle| q\,;\, q\right).$$

Using again the q-variant of Zeilberger's algorithm, we get for these OPS

```
> LqL:=qphihyperterm([q^(-n),q^(n+1)],[q],q,q*x,k);
```

$$LqL := \frac{qpochhammer\left(q^{-n},q,k\right) qpochhammer\left(q^{n+1},q,k\right)(qx)^k}{(qpochhammer\,(q,q,k))^2}$$

```
> qsumrecursion(LqL,q,k,p(n));
```

$$q^n\left(q^n-1\right)\left(q^n+q\right)p(n) + \left(q^{2n}-q\right)\left(q^{2n}x + q^{n+1}x + q^n x + qx - 2q^n\right)p(n-1)$$
$$+\left(q^n+1\right)\left(q^n-q\right)q^n p(n-2) = 0$$

```
> BqL:=qphihyperterm([q^(-n),q^(n+1),x],[q,c*q],q,q,k);
```

$$BqL := \frac{qpochhammer\left(q^{-n},q,k\right) qpochhammer\left(q^{n+1},q,k\right) qpochhammer\,(x,q,k)\, q^k}{(qpochhammer\,(q,q,k))^2\, qpochhammer\,(cq,q,k)}$$

```
> qsumrecursion(BqL,q,k,P(n));
```

$$\left(q^n+q\right)q\left(q^n-1\right)\left(cq^n-1\right)P(n)$$
$$+\left(q^{2n}-q\right)\left(q^{2n}x - 2cq^{n+1} + q^{n+1}x - 2q^{n+1} + q^n x + qx\right)P(n-1)$$
$$-\left(q^n+1\right)\left(q^n-q\right)\left(-cq+q^n\right)q^n P(n-2) = 0$$

References

1. G. Almkvist, D. Zeilberger, The method of differentiating under the integral sign. J. Symb. Comput. **10**, 571–591 (1990)
2. Apéry, R.: Irrationalité de $\zeta(2)$ et $\zeta(3)$. Astérisque **61**, 11–13 (1979)
3. H. Böing, W. Koepf, Algorithms for q-hypergeometric summation in computer algebra. J. Symb. Comput. **28**, 777–799 (1999)
4. CAOP: Computer algebra and orthogonal polynomials. www.caop.org
5. M.C. Fasenmyer, Some generalized hypergeometric polynomials. Ph.D. Thesis, University of Michigan (1945)
6. M.C. Fasenmyer, A note on pure recurrence relations. Am. Math. Mon. **56**, 14–17 (1949)
7. R.W. Gosper Jr., Decision procedure for indefinite hypergeometric summation. Proc. Natl. Acad. Sci. U. S. A. **75**, 40–42 (1978)
8. D. Gruntz, W. Koepf, Maple package on formal power series. Maple Tech. Newsl. **2**(2), 22–28 (1995). Appeared also as `SC-93-31.pdf`, Konrad-Zuse-Zentrum Berlin, see http://www.mathematik.uni-kassel.de/~koepf/Publikationen/
9. R. Koekoek, P.A. Lesky, R. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010)
10. W. Koepf, On two conjectures of M. S. Robertson. Complex Variables **16**, 127–130 (1991)
11. W. Koepf, Power series in computer algebra. J. Symb. Comput. **13**, 581–603 (1992)
12. W. Koepf, *Computeralgebra. Eine algorithmisch orientierte Einführung* (Springer, Berlin, 2006)
13. W. Koepf, *Hypergeometric Summation. An Algorithmic Approach to Summation and Special Function Identities,* 1st edn. (Vieweg, Wiesbaden, 1998); 2nd edn. Springer Universitext (Springer, London, 2014)
14. M.A. Perlstadt, Some recurrences for sums of powers of binomial coefficients. J. Number Theory **27**, 304–309 (1987)
15. M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients. J. Symb. Comput. **14**, 243–264 (1992)
16. M.S. Robertson, Complex powers of p-valent functions and subordination, in *Complex Analysis, Proceedings of SUNY Conference, Brockport 1976*. Lecture Notes in Mathematics, vol. 36 (1978), pp. 1–33
17. M.S. Robertson, A subordination problem and a related class of polynomials, in *Univalent Functions, Fractional Calculus, and Their Applications*, ed. by H.M. Srivastava, Sh. Owa (eds.) (Ellis Horwood Limited, Chichester, 1989), pp. 245–266
18. B. Salvy, P. Zimmermann, GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. ACM Trans. Math. Softw. **20**, 163–177 (1994)
19. R.P. Stanley, Differentiably finite power series. Eur. J. Comb. **1**, 175–188 (1980)
20. M. van Hoeij, Finite singularities and hypergeometric solutions of linear recurrence equations. J. Pure Appl. Algebra **139**, 109–131 (1998)
21. D. Zeilberger, A fast algorithm for proving terminating hypergeometric identities. Discrete Math. **80**, 207–211 (1990)

On the Solutions of Holonomic Third-Order Linear Irreducible Differential Equations in Terms of Hypergeometric Functions

Merlin Mouafo Wouodjié

Abstract We present here an algorithm that combines change of variables, exp-product and gauge transformation to represent solutions of a given irreducible third-order linear differential operator L, with rational function coefficients and without Liouvillian solutions, in terms of functions $S \in \left\{ {}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2 \right\}$ where ${}_pF_q$ with $p \in \{0, 1, 2\}$, $q \in \{1, 2\}$, is the generalized hypergeometric function. That means we find rational functions r, r_0, r_1, r_2, f such that the solution of L will be of the form

$$y = \exp\left(\int r\,dx\right)\left(r_0 S(f(x)) + r_1 (S(f(x)))' + r_2 (S(f(x)))''\right).$$

An implementation of this algorithm in Maple is available.

Keywords Hypergeometric functions · Operators · Transformations · Change of variables · Exp-product · Gauge transformation · Singularities · Generalized exponents · Exponent differences · Rational functions · Zeroes · Poles

Mathematics Subject Classification (2000) 34-XX, 33C10, 33C2, 34B30, 34Lxx

1 Introduction

We consider a differential equation of type

$$a_n(x)\ y^{(n)}(x) + a_{n-1}(x)\ y^{(n-1)}(x) + \ldots + a_0(x)\ y(x) = 0,\ n \in \mathbb{N}_{>0}.$$

M. Mouafo Wouodjié (✉)
Institute of Mathematics, University of Kassel, Kassel, Germany
e-mail: merlin@aims-cameroon.org

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_8

When the coefficients $a_i(x), i = 0, \dots, n$ are polynomials of the variable x, the differential equation is said to be holonomic.

Remark 1.1 For a given homogeneous linear differential equation with rational function coefficients, one can multiply by their common denominator and get a holonomic differential equation with the same space of solutions.

Every holonomic differential equation

$$a_n(x)\ y^{(n)}(x) + a_{n-1}(x)\ y^{(n-1)}(x) + \ldots + a_0(x)\ y(x) = 0, \tag{1.1}$$

with $a_i(x) \in \mathbb{Q}[x]$, corresponds to a differential operator L given by

$$\mathrm{L} = \sum_{i=0}^{n} a_i(x) D_x^i,$$

and vice versa. Hence, by the solutions of a differential operator L we mean the solutions y of the holonomic differential equation $\mathrm{L}y = 0$.

Definition 1.2 Let

$$\mathrm{L} = \sum_{i=0}^{n} a_i(x) D_x^i,$$

and $n_0 = \max\{i = 0, 1, \dots, n \mid a_i(x) \neq 0\}$. n_0 is called the order of L, and $a_{n_0}(x)$ its leading coefficient.

The ring of differential operators with coefficients in $\mathrm{K} = \mathbb{Q}[x]$ is denoted by $\mathrm{K}[D_x]$. In our context, $D_x := \frac{d}{dx}$.

2 Previous Works

2.1 First-Order Holonomic Differential Equations

First-order holonomic differential equations are of type

$$a_1(x)\ y' + a_0(x)\ y = 0 \quad \text{with } a_0(x), a_1(x) \in \mathbb{Q}[x], \quad a_1(x) \neq 0. \tag{2.1}$$

That means, non-zero solutions satisfy

$$\frac{y'}{y} = -\frac{a_0(x)}{a_1(x)},$$

and they can be easily computed in the form

$$y(x) = c \cdot \exp\left(\int -\frac{a_0(x)}{a_1(x)}\,dx\right) \quad \text{with } c \in \mathbb{R}.$$

Those solutions are called hyperexponential functions.

Example Let us consider the following differential equation

$$(x^2+1)y' - xy = 0.$$

Non-zero solutions are

$$\begin{aligned} y(x) &= c \cdot \exp\left(\int -\frac{-x}{x^2+1}\,dx\right) = c \cdot \exp\left(\frac{1}{2}\ln\left(x^2+1\right)\right) \\ &= c \cdot \sqrt{x^2+1} \quad \text{with } c \in \mathbb{R}. \end{aligned}$$

2.2 *Second-Order Holonomic Differential Equations*

Second-order holonomic differential equations are of the form

$$a_2(x)\, y'' + a_1(x)\, y' + a_0(x)\, y = 0 \tag{2.2}$$

with $a_0(x), a_1(x), a_2(x) \in \mathbb{Q}[x]$, $a_2(x) \neq 0$. Let L be the associated differential operator of (2.2). In comparison with the first-order holonomic differential equations, the situation here is quite different. Either L **is reducible or irreducible**.

2.2.1 L Is Reducible

Here L has a nontrivial (non-commutative) factorization. In this case the right factor is of first order leading to a hyperexponential solution of $\mathrm{L}(y) = 0$, again. To check the reducibility of an operator, we can use some known algorithms like *Beke's algorithm* or the algorithm in [9]. Beke's algorithm was extended by Mark van Hoeij (see [12]) in his PhD thesis on factorization of linear differential operators.

Example Let us consider the following differential equation

$$(x+7)y'' - x(x+7)y' + xy = 0. \tag{2.3}$$

Its associated operator can be factorized as follows:

$$(x+7)D_x^2 - x(x+7)D_x + x = (D_x - x)\cdot((x+7)D_x - 1)$$

Hence, solving $(x+7)y' - y = 0$, we get a solution of (2.3) given by

$$y(x) \ = \ \exp\left(\int \frac{1}{x+7}\,dx\right) = \exp\left(\ln\left(x+7\right)\right) = x+7.$$

2.2.2 L Is Irreducible

In this case it is more difficult to find solutions of $\mathrm{L}(y) = 0$. There are some algorithms which try to find them in some particular forms. *Kovacic's algorithm* [5] (which finds Liouvillian solutions) is an example. Some complete algorithms which solve $\mathrm{L}(y) = 0$ in terms of special functions are given by Mark van Hoeij, Wolfram Koepf, Ruben Debeerst and Quan Yuan [1, 2, 13, 14]. This complete algorithm tries to find all solutions of the type

$$\exp\left(\int r\,dx\right)\left(r_0 S(f(x)) + r_1\left(S(f(x))\right)'\right) \tag{2.4}$$

where S is the special function that we want to solve in terms of it, and $r, r_0, r_1, f \in \mathbb{Q}(x)$ are parameters of the three following transformations which preserve the order of the operator:

1. change of variables $\xrightarrow{f}_C\colon y(x) \to y(f(x))$,
2. exp-product $\xrightarrow{r}_E\colon y \to \exp\left(\int r\,dx\right) y$, and
3. gauge transformation $\xrightarrow{r_0,r_1}_G\colon y \to r_0 y + r_1 y'$.

Example Let us consider the following differential equation

$$4(x-2)^2 y'' + (4x-8)y' + (-144x^4 + 1152x^3 - 3456x^2 + 4608x - 2305)y = 0. \tag{2.5}$$

Its associated operator cannot be factorized. By using Ruben Debeerst's code, we see that a solution of (2.5) is given by

$$y(x) = \mathrm{B}_{\frac{1}{4}}\left(3(x-2)^2\right),$$

where $\mathrm{B}_\nu(x)$ is the Bessel function of parameter ν.

Remark 2.1 Bessel functions belong to the class of special functions since they are expressed in terms of the most prominent special function solutions of holonomic differential equations called generalized hypergeometric functions.

2.2.3 Generalized Hypergeometric Functions

The generalized hypergeometric series ${}_pF_q$ are defined by

$$ {}_pF_q\left(\left.\begin{matrix} a_1, a_2, \ldots, a_p \\ b_1, b_2, \ldots, b_q \end{matrix}\right| x\right) = \sum_{k=0}^{+\infty} \frac{(a_1)_k \cdot (a_2)_k \cdots (a_P)_k}{(b_1)_k \cdot (b_2)_k \cdots (b_q)_k \cdot k!} x^k, \tag{2.6} $$

where $(a)_k$ denotes the *Pochhammer symbol*

$$ (a)_k := \begin{cases} 1 & \text{if } k = 0, \\ a \cdot (a+1) \cdots (a+k-1) & \text{if } k > 0\,. \end{cases} $$

It satisfies the following holonomic differential equation

$$ \delta(\delta + b_1 - 1) \cdots (\delta + b_q - 1) y(x) = x(\delta + a_1) \cdots (\delta + a_p) y(x) \tag{2.7} $$

where δ denotes the differential operator $\delta = x\frac{d}{dx}$. This equation has order $\max(p, q+1)$. For $p \leq q$ the series ${}_pF_q$ is convergent for all x. For $p > q + 1$ the radius of convergence is zero, and for $p = q + 1$ the series converges for $|x| < 1$. For $p \leq q + 1$ the series and its analytic continuation is called generalized hypergeometric function. We are interested here in the case $p < q + 1$ for which the radius of convergence is infinity.

2.3 Third-Order Holonomic Differential Equations

Third-order holonomic differential equations are of type

$$ a_3(x)\, y''' + a_2(x)\, y'' + a_1(x)\, y' + a_0(x)\, y = 0 \tag{2.8} $$

with $a_0(x), a_1(x), a_2(x), a_3(x) \in \mathbb{Q}[x]$, $a_3(x) \neq 0$. Let L be the associated differential operator of (2.8).

2.3.1 L Is Reducible

Here, solutions of $\mathrm{L}(y) = 0$ can be in some case easily computed, since we know how to solve first-order holonomic differential equations and also, in some particular cases, second-order holonomic differential equations. Michael Singer described (see [8]) in which situation L has so-called *Eulerian solutions* (solutions which can be expressed as products of second-order operators using sums, products, field operations, algebraic extensions, integrals, differentiations, exponentials, logarithms and change of variables). He showed that solving such an operator L can be reduced to solving second-order operators through factoring operators (see [3, 10–12]), or reducing operators to tensor products of lower order operators.

Example Let us consider the following differential equation

$$\begin{aligned}
&9x^6(x+7)^2y''' + 9x^4(x+7)(8x^2+47x+84)y'' \\
&+ x^2(2954x^2 - 882x + 90x^4 + 972x^3 - 21609)y' \\
&+ (-74088x - 259308 - 1274x^3 - 13818x^2)y = 0. \qquad (2.9)
\end{aligned}$$

The operator associated to (2.9) can be factorized in the following form

$$\begin{aligned}
&\left(x^2D_x + 12\right)\cdot\Big(9x^4(x+7)^2D_x^2 + 9(7+2x)x^3(x+7)D_x \\
&-21609 - 637x^2 - 6174x\Big). \qquad (2.10)
\end{aligned}$$

Hence, solving

$$9x^4(x+7)^2y'' + 9(7+2x)x^3(x+7)y - (21609 + 637x^2 + 6174x)y = 0 \qquad (2.11)$$

gives us a solution of (2.9). By using Ruben Debeerst's code, we see that a solution of (2.11) is given in terms of Bessel functions:

$$y(x) = \mathrm{B}_{\frac{2}{3}}\left((x+7)/x\right)$$

which is therefore a solution of (2.9).

2.3.2 L Is Irreducible

If **the differential operator** L associated to (2.8) **is irreducible**, it is difficult to solve the Eq. (2.8). In addition to being irreducible, if Liouvillian or Eulerian solutions of L are not allowed, then the first algorithm was published in my PhD thesis [6] and also in [7]. For L of order three for example, that is the case where this operator comes from certain special and useful functions such as hypergeometric functions. That is why we focus on those operators here and specially those of order three.

Example Let us consider the following third-order irreducible holonomic differential equation satisfied by the square of the Hermite polynomial:

$$y''' - 6xy'' + (8x^2 + 8n - 2)y' - 16xny = 0. \qquad (2.12)$$

If we assume that we don't know where (2.12) is coming from, it will be difficult to solve it. But, by using one of my codes which I have implemented in a Maple package called `Solver3` that can be downloaded from http://www.mathematik.uni-kassel.de/~merlin/, we get some solutions. The Maple function is called `Hyp1F1sqSolutions` and takes as input any irreducible third-order linear differential operator L and returns, if they exist, all the parameters of transformations $(r, r_2, r_1, r_0, f \in k(x))$ and the parameters of the function ${}_1F_1{}^2$ in terms of which we are solving (2.12).

```
> read "Solver3.mpl":
> L:= Dx^3-6*x*Dx^2+(8*x^2+8*n-2)*Dx-16*n*x;
```

$$L := Dx^3 - 6\,xDx^2 + \left(8\,x^2 + 8\,n - 2\right) Dx - 16\,nx$$

Using my code called `Hyp1F1sqSolutions` we get:

```
> Hyp1F1sqSolutions(L);
```

$$\left\{\left[[[[-n/2]\,, 1/2, [0]\,, [1]]]\,, x^2\right],\right.$$
$$\left.\left[\left[\left[[n/2]\,, 1/2, [4\,x]\,, \left[Dx^2 - 6\,xDx + 8\,x^2 + 4\,n - 4\right]\right]\right], -x^2\right]\right\}$$

Hence, (2.12) has ${}_1F_1{}^2$ type solutions. One of them has $a = -n/2$ and $b = 1/2$ as upper and lower parameters of ${}_1F_1$, respectively, and the transformation parameters are: $r = 0$, $r_0 = 1$, $r_1 = 0$, $r_2 = 0$ and $f = x^2$. This solution is

```
> y:=expand(exp(int(r,x))*(r0*(hypergeom([a], [b], x^2))^2
+r1*normal(diff((hypergeom([a], [b], x^2))^2,x))
+r2*normal(diff((hypergeom([a], [b], x^2))^2,x$2))));
```

$$y := \left({}_1\mathrm{F}_1(-n/2;\ 1/2;\ x^2)\right)^2$$

3 My Work

3.1 Our Main Objective

We develop a complete algorithm to detect the solutions of any third-order irreducible holonomic differential equation which are related to the following special functions: ${}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2$.

Remark 3.1 If y a is solution of a second-order holonomic differential equation, then y^2 is a solution of a holonomic differential equation of order three. That is the case for the function $y = {}_1F_1$. This is a rich source for third-order holonomic differential equations whose solutions are sought.

All our special functions ${}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2$ satisfy third-order holonomic irreducible differential equations. For example, the differential operator associated to the ${}_1F_1^2$ function is

$$\mathrm{L}_{11}^2 = x^2 D_x^3 + 3x\,(-x+b)\,D_x^2 - \left(-2x^2 + 4x(a+b) - b(2b-1)\right) D_x - 2a\,(-2x+2b-1)$$

where a and b are the upper and lower parameters of ${}_1F_1$, respectively.

3.2 The Method

To solve our differential equations, we use the three transformations

1. change of variables $\xrightarrow{f}_C$: $y(x) \to y(f(x))$,
2. exp-product $\xrightarrow{r}_E$: $y \to \exp\left(\int r\,dx\right) y$, and
3. gauge transformation $\xrightarrow{r_0,r_1,r_2}_G$: $y \to r_0 y + r_1 y' + r_2 y''$.

where $r, r_0, r_1, f \in \mathbb{Q}(x)$. Our goal is to find all solutions which can be written in the form:

$$\exp\left(\int r\,dx\right)\left(r_0 S(f(x)) + r_1 (S(f(x)))' + r_2 (S(f(x)))''\right)$$

where $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$ and $r, r_0, r_1, r_2, f \in \mathbb{Q}(x)$. That means for a given third-order irreducible holonomic differential equation, to find (if they exist) the transformation parameters r, r_0, r_1, r_2 and f, and also the parameter(s) associated to our chosen special function $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$.

Remark 3.2 The only case which is not covered since it requires different methods, is where the special function has finite radius of convergence: ${}_2F_1{}^2$ and ${}_3F_2$ are the only examples.

3.3 Our Inputs

As input, we consider a third-order irreducible holonomic differential equation with coefficients in $\mathbb{Q}[x]$

$$(eq):\quad a_3(x)\ y''' + a_2(x)\ y'' + a_1(x)\ y' + a_0(x)\ y = 0$$

that we want to solve in term of $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$. Let us call L the differential operator associated to (eq).

L_S is always the operator associated to the differential equation satisfied by $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$.

3.4 Steps of the Resolution

1. Find the singularities of the differential operator L
2. Find the generalized exponents of L
3. Find the transformation parameter(s) which are:

(a) the change of variable parameter f
(b) the parameter(s) of our chosen special function $S \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$
(c) the exp-product parameter r
(d) the gauge parameters r_0, r_1 and r_2

3.5 Singularities

Definition 3.3 A point $p \in \mathbb{C} \cup \{\infty\}$ is called a *singularity* of a holonomic differential operator L, if p is a zero of the leading coefficient of L. All the other points are called *regular* points.

Remark 3.4

- To understand the singularity at $x = \infty$, one can always use the change of variables $x \to \frac{1}{x}$ and deal with 0.
- At all regular points of L we can find a *fundamental system of power series solutions.*

If $p \in \mathbb{C} \cup \{\infty\}$, we define the *local parameter* t_p as

$$t_p = \begin{cases} x - p & if\ p \neq \infty, \\ \frac{1}{x} & if\ p = \infty\ . \end{cases}$$

Let $L_{\frac{1}{x}}$ denote the operator coming from L by the change of variables $x \to \frac{1}{x}$.

Definition 3.5 A singularity p of L is called

(i) *apparent singularity* if all solutions of L are regular at p,
(ii) *regular singular* ($p \neq \infty$) if $t_p^i \dfrac{a_{3-i}(x)}{a_3(x)}$ is regular at p for $1 \leq i \leq 3$,
(iii) regular singular ($p = \infty$) if $L_{\frac{1}{x}}$ has a regular singularity at $x = 0$, and
(iv) *irregular singular* otherwise.

The operators coming from our functions ${}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2$ have two singularities: one regular at $x = 0$ and the other irregular at $x = \infty$.

3.6 Generalized Exponents

Let us consider the following differential equation

$$(E7): \quad x^2(x+1)y'' - (6x + 7x^2)y' + (12 + 16x)y = 0.$$

By searching a solution of $(E7)$ in a neighborhood of $x = 0$ of the form $x^c \cdot G$ with $c \in \mathbb{Q}$, $G \in \mathbb{Q}[x][\ln(x)]$ such that G has a non-zero constant coefficient, we get

$$y(x) = x^3 \left(1 + x - x\ln(x)\right),$$

and therefore, $c = 3$. What happens if we want solutions of the same form but with c which is not a constant? The answer to this question leads us to the definition of the generalized exponent.

Definition 3.6 Let p be a point with local parameter t_p. An element $e \in \mathbb{Q}[{t_p}^{-1/n}]$, $n \in \mathbb{N}_{>0}$ is called a *generalized exponent* of L at the point p if there exists a *formal solution* of L of the form

$$y(x) = \exp\left(\int \frac{e}{t_p}\,dt_p\right) G, \quad G \in \mathbb{Q}(({t_p}^{1/n}))[\ln(t_p)], \tag{3.1}$$

where the constant term of the *Puiseux series* G is non-zero. For a given solution this representation is unique and $n \in \mathbb{N}$ is called the *ramification index* of e.

The set of generalized exponents at a point p is denoted by gexp(L, p).

Similarly, we call e a generalized exponent of the solution y at the point p if $y = y(x)$ has a representation (3.1) for some $G \in \mathbb{Q}(({t_p}^{1/n}))[\ln(t_p)]$.

There is an algorithm called `gen_exp` given by Mark van Hoeij which computes the generalized exponents of L at a given point.

3.6.1 Generalized Exponents and Singularities

- p is an irregular singularity of L if L has at p at least one non-constant generalized exponent.
- p is a regular point of L if the three generalized exponents of L at p are $0, 1, 2$.
- If p is an apparent singularity of L, then all the generalized exponents of L at p are non-negative integers.
- If p is a regular singularity of L, then all the generalized exponents of L at p are constants.

The operator L_{11}^2 coming from the function ${}_1F_1{}^2$ with parameters a and b has as generalized exponents:

1. at its regular singularity $x = 0$: $[0,\ 1 - b_1,\ 2(1 - b_2)]$,
2. at its irregular singularity $x = \infty$: $\left[2a_1,\ -2t^{-1} + 2(b_1 - a_1),\ -t^{-1} + b_1\right]$ with $t = \frac{1}{x}$.

3.6.2 Generalized Exponents and Exp-Product Transformation

Lemma 3.7 *Let* $\mathrm{L}_1, \mathrm{L}_2 \in \mathbb{Q}[x][\partial]$ *be two irreducible third-order holonomic differential operators such that* $\mathrm{L}_1 \xrightarrow{r}_E \mathrm{L}_2$ *and let e be a generalized exponent* *of* L_1 *at the point* $p \in \mathbb{C} \cup \{\infty\}$ *with the ramification index* $n_e \in \mathbb{N}^*$. *Furthermore, let* r *has at* p *the series representation*

$$r = \sum_{i=m_p}^{+\infty} r_i t_p^i, \quad m_p \in \mathbb{Z} \ \text{ with } \ r_i \in \mathbb{Q} \ \text{ and } \ r_{m_p} \neq 0.$$

where t_p *is the local parameter of* p.

1. *If* p *is not a pole of* r *then* $m_p \geq 0$ *and the generalized exponent of* L_2 *at* p *is*

$$\overline{e} = \begin{cases} e & \text{if } \ p \neq \infty, \\ e - r_0 t_\infty^{-1} - r_1 & \text{otherwise.} \end{cases}$$

2. *If* p *is a pole of* r *then* $m_p \leq -1$, *where* $-m_p$ *is the multiplicity order of* r *at* p, *and the generalized exponent of* L_2 *at* p *is given by*

$$\overline{e} = \begin{cases} e + \sum\limits_{i=m_p}^{-1} r_i t_p^{i+1} & \text{if } \ p \neq \infty, \\ e - \sum\limits_{i=m_\infty}^{1} r_i t_\infty^{i-1} & \text{otherwise.} \end{cases}$$

Proof The proof can be found in my PhD thesis [6, Lemma 3.32] and also in [7].

□

3.6.3 Generalized Exponents and Gauge Transformation

Lemma 3.8 *Let* $\mathrm{L}_1, \mathrm{L}_2 \in \mathbb{Q}[x][\partial]$ *be two irreducible third-order holonomic differential operators such that* $\mathrm{L}_1 \longrightarrow_G \mathrm{L}_2$ *and let* e *be a generalized exponent of* L_1 *at some point* p. *The operator* L_2 *has at* p *a generalized exponent* $\overline{e}$ *such that* $\overline{e} = e \bmod \frac{1}{n_e}\mathbb{Z}$, *where* $n_e \in \mathbb{N}^*$ *is the ramification index of* e.

Proof The proof can be found in my PhD thesis [6, Lemma 3.32] and also in [7].

□

3.6.4 Generalized Exponents and Change of Variable Transformation

Let us consider the case $\mathrm{L}_S \xrightarrow{f}_C \mathrm{M}$ with $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$ and $f \in \mathbb{Q}(x)\backslash\mathbb{Q}$. Since L_S has singularities at 0 and ∞, we will see how the generalized exponents of M look like at the points p such that $f(p) = 0$ and $f(p) = \infty$ (i.e. at the *zeroes* and *poles* of f).

Theorem 3.9 *Let* $\mathrm{M} \in \mathbb{Q}[x][\partial]$ *be an irreducible third-order linear differential operator such that* $\mathrm{L}_S \xrightarrow{f}_C \mathrm{M}$, *with* $S(x) \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$ *and* $f \in \mathbb{Q}(x) \setminus \mathbb{Q}$.

1. *Let* p *be a zero of* f *with multiplicity* $m_p \in \mathbb{N}^*$ *and* e *a generalized exponent of* L_S *at* $x = 0$ *with ramification index* $n_e \in \mathbb{N}^*$. *Then* p *is a regular singularity of* M *and the generalized exponent of* M *at* p *related to* e *is*

$$m_p \cdot e_0 - \sum_{i=1}^{n} \sum_{j=-i\cdot m_p}^{-1} \frac{j \cdot e_i \cdot \sigma^{-i}}{i} \overline{f}_{i,j+i\cdot m_p} t_p^{j/n_e} \tag{3.2}$$

where

$$\begin{cases} f = t_p^{m_p} \sum\limits_{j=0}^{+\infty} f_j t_p^j, \quad \textit{with}\ \ f_i \in k \ \ \textit{and} \ \ f_0 \neq 0, \\ e = \sum\limits_{i=0}^{n} e_i t^{-i} \quad \textit{with}\ \ bt^{n_e} = x,\ b \in k \setminus \{0\},\ n \in \mathbb{N} \ \textit{and}\ \ e_i \in k, \\ \left(\sum\limits_{j=0}^{+\infty} f_j t_p^j\right)^{-i/n_e} = \sum\limits_{j=0}^{+\infty} \overline{f}_{i,j} t_p^{j/n_e} \quad \textit{with}\ \ \overline{f}_{i,j} \in k,\ i = 1, \ldots, n \\ \textit{and}\ \sigma\ \ \textit{is solution of}\ \ X^{n_e} - b^{-1} = 0\ \ \textit{with unknown}\ \ X. \end{cases}$$

2. *Let* p *be a pole of* f *with multiplicity* $m_p \in \mathbb{N}^*$ *and* e *a generalized exponent of* L_S *at* $x = 0$ *with ramification index* $n_e \in \mathbb{N}^*$. *Then* p *is an irregular singularity of* M *and the generalized exponent of* M *at* p *related to* e *is*

$$m_p \cdot e_0 - \sum_{i=1}^{n} \sum_{j=-i\cdot m_p}^{-1} \frac{j \cdot e_i \cdot \sigma^{-i}}{i} \overline{f}_{i,j+i\cdot m_p} t_p^{j/n_e} \tag{3.3}$$

where

$$\begin{cases} f = t_p^{-m_p} \sum\limits_{j=0}^{+\infty} f_{j-m_p} t_p^j, \quad \textit{with}\ \ f_{j-m_p} \in k \ \ \textit{and} \ \ f_{-m_p} \neq 0, \\ e = \sum\limits_{i=0}^{n} e_i t^{-i} \ \ \textit{with}\ bt^{n_e} = t_\infty,\ b \in k \setminus \{0\},\ n \in \mathbb{N}\ \textit{and}\ \ e_i \in k, \\ \left(\sum\limits_{j=0}^{+\infty} f_{j-m_p} t_p^j\right)^{i/n_e} = \sum\limits_{j=0}^{+\infty} \overline{f}_{i,j} t_p^{j/n_e} \quad \textit{with}\ \ \overline{f}_{i,j} \in k,\ i = 1, \ldots, n \\ \textit{and}\ \sigma\ \ \textit{is solution of}\ \ X^{n_e} - b^{-1} = 0\ \ \textit{with unknown}\ \ X. \end{cases}$$

Proof The proof can be found in my PhD thesis [6, Theorem 2.11] and also in [7]. □

Let us apply this theorem where L_S is the operator L_{11}^2 coming from the function ${}_1F_1{}^2$ with parameters a and b.

1. If p is a zero of f with *multiplicity* $m_p \in \mathbb{N}^*$, then the generalized exponents of M at p are
$$\left[0,\ \ m_p\,(1-b_1)\,,\ \ 2m_p\,(1-b_1)\right].$$
2. If p is a pole of f with *multiplicity* $m_p \in \mathbb{N}^*$, then the generalized exponents of M at p are
$$\left[2m_p a_1,\quad -2m_p(a_1-b_1)+2\sum_{j=-m_p}^{-1} j f_j t_p^j,\ \ m_p b_1+\sum_{j=-m_p}^{-1} j f_j t_p^j\right],$$
where $f = t_p^{-m_p}\sum_{j=0}^{+\infty} f_{j-m_p} t_p^j$ with $f_{j-m_p} \in \mathbb{Q}$ and $f_{-m_p} \neq 0$.

3.7 How to Find the Transformation Parameter(s)

The task now is to find our three transformations such that
$$\mathrm{L}_S \xrightarrow{f}_C \mathrm{M} \xrightarrow{r}_E \mathrm{L}_1 \xrightarrow{r_0,r_1,r_2}_G \mathrm{L}$$
with $r, r_0, r_1, r_2, f \in \mathbb{Q}(x)$, $\ f \notin \mathbb{Q}$ and $\mathrm{M}, \mathrm{L}_1 \in \mathbb{Q}[x][\partial]$. A solution y of L in terms of S will be
$$y = \exp\left(\int r\,dx\right)\left(r_0 S(f(x)) + r_1 (S(f(x)))' + r_2 (S(f(x)))''\right).$$

Finding these transformations is equivalent to find their parameter(s). We proceed as follows:

1. in the first step, we find the change of variable parameter f and the parameter(s) associated to the function S,
2. in the second step, we find the parameters r, r_0, r_1 and r_2 for the exp-product and gauge transformations.

I will now show how to get these parameters in each of these steps, and refer for more details to my PhD thesis (see [6]) which can be downloaded from http://www.mathematik.uni-kassel.de/~merlin/ and also in [7].

3.7.1 How to Find the Change of Variable Parameter f

Let $S \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$.

1. If the ramification index of L_S at ∞ is 1 we compute the polar part of f from the generalized exponents of L at its irregular singularities. Then we get f by using the regular singularities of L or some information related to the degree of the numerator that f can have.
2. If the ramification index of L_S at ∞ is greater than 1, we put f in the form $f = \dfrac{\mathrm{A}}{\mathrm{B}}$ with $\mathrm{A}, \mathrm{B} \in k[x]$, B monic and $\gcd(\mathrm{A}, \mathrm{B}) = 1$. Using the generalized exponents at the irregular singularities of L, we can compute B and a bound for the degree of A. Hence, using also the fact that the ramification index of L_S is greater than 1, we can get the truncated series for f and some linear equations for the coefficients of A. By comparing the number of linear equations for the coefficients of A and the degree of A, we will deal with some cases which will help us to find A.

3.7.2 How to Find the Parameter(s) of Our Chosen Special Function $S \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$

Let us assume that we know f. By observing the forms of the generalized exponents of L at the zeroes and poles of f with their corresponding multiplicity orders, we find the parameter(s) of our special function S. Therefore, we obtain the differential operator L_S associated to S. That means we get the operator M coming from L_S by the change of variable transformation with parameter f

$$\mathrm{L}_S \xrightarrow{f}_C \mathrm{M}.$$

3.7.3 How to Find the Exp-Product Parameter r

Let us assume that we know the operator M such that

$$\mathrm{L}_S \xrightarrow{f}_C \mathrm{M} \longrightarrow_{EG} \mathrm{L}.$$

This theorem finds the exp-product parameter in the projective equivalence $(\longrightarrow_{EG})$.

Theorem 3.10 *Let* $\mathrm{M}, \mathrm{L} \in k(x)[\partial]$ *be two irreducible third-order linear differential operators such that* $\mathrm{M} \longrightarrow_{EG} \mathrm{L}$ *and* r *the parameter of the exp-product transformation. Let* NS *be the set of all non-apparent singularities of* L *and* $\mathbb{P}_0$ *the set of all the poles of* r. *For* $p \in \mathbb{P}_0 \cup \mathrm{NS}$, *let us set*

$$e_p^i = e_p^i(\mathrm{L}) - e_p^i(\mathrm{M}), \quad i = 1, 2, 3$$

where $e_p^i(\mathrm{M})$ *and* $e_p^i(\mathrm{L})$ *are the* i*-th generalized exponent of* M *and* L *at* p*, respectively, and* r *has series representation*

$$r = \sum_{i=-m_p}^{+\infty} r_{p,i} t_p^i, \quad m_p \in \mathbb{N} \ \ with \ \ r_{p,i} \in k \ \ and \ \ r_{p,-m_p} \neq 0.$$

If we assume that

1. $\mathbb{P}_0^{11} = \{p \in \mathbb{P}_0 | \ \{e_p^1(\mathrm{L}), e_p^2(\mathrm{L}), e_p^3(\mathrm{L})\} \subseteq \mathbb{Z} \ and \ r_{p,-1} \notin \mathbb{Z}\Big\} = \emptyset,$
2. M *is not the image of an exp-product transformation with rational function* $-r + a_p t_p^{-1}$ *with* $a_p \in k$ *and* $p \in \mathbb{P}_0$ *such that* $m_p \geq 2$ *if* $p \neq \infty$*,*

then

$$\sum_{p \in \mathrm{NS}\setminus\{\infty\}} \frac{e_p^i}{t_p} - t_\infty \cdot \overline{e_\infty^i} = r + \sum_{p \in \mathbb{S}\setminus\{\infty\}} \frac{b_p}{n_p} t_p^{-1} - \sum_{p \in \mathbb{P}_0^{12}\setminus(\mathrm{NS}\cup\{\infty\})} r_{p,-1} t_p^{-1} \tag{3.4}$$

where $\overline{e_\infty^i} = e_\infty^i - const(e_\infty^i)$ *with* $const(e_\infty^i)$ *the constant term of* e_∞^i, $b_p \in \mathbb{Z}$, $n_p = \max\left\{n_{e_p^i(\mathrm{L}_2)}, i = 1, 2, 3\right\}$ *with* $n_{e_p^i(\mathrm{L}_2)}$ *the ramification index of* $e_p^i(\mathrm{L}_2)$*, and* $\mathbb{P}_0^{12} = \{p \in \mathbb{P}_0 | \ \{e_p^1(\mathrm{L}_2), e_p^2(\mathrm{L}_2), e_p^3(\mathrm{L}_2), r_{p,-1}\} \subseteq \mathbb{Z}\Big\}.$

Proof The proof can be found in my PhD thesis [6, Lemma 3.32] and also in [7]. □

Lemma 3.11 *Let us consider the hypothesis and notations of Theorem 3.10, and assume that all the conditions of Theorem 3.10 are satisfied. Then the parameter* r *of the exp-product transformation is given by*

$$r = \sum_{p \in \mathbb{S}\setminus\{\infty\}} \frac{e_p^i}{t_p} - t_\infty \cdot \overline{e_\infty^i} + \sum_{p \in \mathbb{S}\setminus\{\infty\}} \frac{c_p}{n_p} t_p^{-1} \tag{3.5}$$

with $c_p \in \mathbb{Z}$ *and* $|c_p| < n_p$*.*

Proof The proof can be found in my PhD thesis [6, Lemma 3.32] and also in [7]. □

3.7.4 How to Find the Gauge Parameters r_0, r_1 and r_2

Let us assume that we know our exp-product parameter r. Therefore, we have the operator L_1 such that

$$\mathrm{L}_S \xrightarrow{f}_C \mathrm{M} \xrightarrow{r}_E \mathrm{L}_1 \xrightarrow{r_0, r_1, r_2}_G \mathrm{L}.$$

To find the gauge parameters $r_0, \ldots, r_{n-1}$ (gauge equivalence problem) between two operators L_1 and L we use Mark van Hoeij's gauge equivalence test (see [11]) which gives us those parameters as the coefficients of a second-order linear differential operator. There already exists in Maple an algorithm for that called `Homomorphisms`, implemented by Mark van Hoeij (see [11]), which takes as input two linear differential operators L_i and L_j of order n and returns a basis of all the operators of order $n-1$ satisfying the equation $\mathrm{L}_i = \mathrm{L}_j\mathrm{X}$ if the first input is L_i, and $\mathrm{L}_j = \mathrm{L}_i\mathrm{X}$ if the first input is L_j (X is the unknown of this equation). Otherwise it returns an empty list. Let $n = 3$ and $\mathrm{X} = a_2(x)\partial^2 + a_1(x)\partial + a_0(x)$, we deduce the gauge parameters r_0, r_1 and r_2 by taking

$$\begin{cases} r_0 = a_0(x), \\ r_1 = a_1(x), \\ r_2 = a_2(x). \end{cases} \tag{3.6}$$

3.7.5 Example

We have implemented the methods of this work in a Maple package called `Solver3` which can be downloaded from http://www.mathematik.uni-kassel.de/~merlin/. My PhD thesis [6] and also [7] explain the algorithms in more detail. We will take here just one example and show how our package can be used. Let us consider the following third-order irreducible holonomic differential equation satisfied by the square of the Laguerre polynomial[1]:

$$x^2y''' + (-3x^2 + 3x)y'' + (4nx + 2x^2 - 4x + 1)y' + (-4nx + 2n)y = 0. \tag{3.7}$$

Let our input operator L be the operator associated to this differential equation and see if we can solve it using our codes. That means if we can find the function $S \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$ and the transformation parameters such that

$$\mathrm{L}_S \xrightarrow{f}_C \mathrm{M} \longrightarrow_{EG} \mathrm{L}.$$

We upload first our Maple package called `Solver3` which can be downloaded from http://www.mathematik.uni-kassel.de/~merlin/.

```
> read "Solver3.mpl":
> L:= x^2*Dx^3+(-3*x^2+3*x)*Dx^2+(4*n*x+2*x^2-4*x+1)*Dx
-4*n*x+2*n;
```

$$L := x^2Dx^3 - \left(-3\,x^2 + 3\,x\right)Dx^2 + \left(4\,n\,x + 2\,x^2 - 4\,x + 1\right)Dx - 4\,n\,x + 2\,n$$

[1]To generate this differential equation, we use the `hsum` package from Wolfram Koepf (see [4]).

Using one of our codes called `Hyp1F1sqSolutions` we get:

```
> Hyp1F1sqSolutions(L);
```

$$\Big\{\Big[\Big[\Big[[n], 2, [2], \Big[\Big(x^2n^2 - 2nx^3 + x^4\Big) Dx^2 + \Big(-5x^2n^2 + 8nx^3 - 3x^4 + 4n^2x$$
$$-6nx^2 + 2x^3\Big) Dx + 2n^3x + 4x^2n^2 - 6nx^3 + 2x^4 - 12n^2x + 12nx^2 - 4x^3$$
$$+2n^2\Big]\Big]\Big], -x\Big], \Big[\Big[\Big[[-n], 2, \Big[x^{-1}\Big], \Big[xDx^2 + (-x+2)Dx + 2n + 1\Big]\Big]\Big], x\Big]\Big\}$$

Hence, (3.7) has ${}_1F_1{}^2$ type solutions. One of them has $a = -n$ and $b = 2$ as upper and lower parameters of ${}_1F_1$, respectively, and the transformation parameters are: $r = 1/x$, $r_0 = 2n + 1$, $r_1 = -x + 2$, $r_2 = x$ and $f = x$. This solution is

```
> y:=expand(exp(int(r,x))*(r0*(hypergeom([a], [b],
x^2))^2+ r1*normal(diff((hypergeom([a], [b], x^2))^2,
x))+r2*normal(diff((hypergeom([-n/2], [1/2], x^2))^2,
x$2))));
```

$$y := 2x\left({}_1\mathrm{F}_1(-n;\ 2;\ x^2)\right)^2 n + x\left({}_1\mathrm{F}_1(-n;\ 2;\ x^2)\right)^2 + 2\,{}_1\mathrm{F}_1(-n;\ 2;\ x^2)n$$
$$\times {}_1\mathrm{F}_1(-n+1;\ 3;\ x^2)x^3 - 4\,{}_1\mathrm{F}_1(-n;\ 2;\ x^2)n\,{}_1\mathrm{F}_1(-n+1;\ 3;\ x^2)x^2$$
$$+8n^2\left({}_1\mathrm{F}_1(-n/2+1;\ 3/2;\ x^2)\right)^2 x^4 + 8/3\,{}_1\mathrm{F}_1(-n/2;\ 1/2;\ x^2)n^2$$
$$\times {}_1\mathrm{F}_1(-n/2+2;\ 5/2;\ x^2)x^4 - 16/3\,{}_1\mathrm{F}_1(-n/2;\ 1/2;\ x^2)n$$
$$\times {}_1\mathrm{F}_1(-n/2+2;\ 5/2;\ x^2)x^4 - 4\,{}_1\mathrm{F}_1(-n/2;\ 1/2;\ x^2)$$
$$\times {}_1\mathrm{F}_1(-n/2+1;\ 3/2;\ x^2)nx^2$$

All explanations of the functioning of this algorithm are available in my PhD thesis [6] and also in [7].

4 Conclusion

We gave an algorithm to find $S \in \{{}_1F_1{}^2,\ {}_0F_2,\ {}_1F_2,\ {}_2F_2\}$ type solutions of an irreducible third-order linear differential operator without Liouvillian solutions and with rational function coefficients. We have also implemented this algorithm in Maple (available from http://www.mathematik.uni-kassel.de/~merlin/).

Acknowledgements This work has been supported by a DAAD scholarship (German Academic Exchange Service) and the University of Kassel by a "Promotions-Abschlussstipendium". All these institutions receive my sincere thanks.

Many thanks to the organizers of the AIMS–Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications, Douala, October 5–12, 2018.

References

1. R. Debeerst, Solving differential equations in terms of Bessel functions. Master's Thesis, Universität Kassel (2007)
2. R. Debeerst, M. van Hoeij, W. Koepf, Solving differential equations in terms of Bessel functions, in *Proceedings of the 2008 International Symposium on Symbolic and Algebraic Computation (ISSAC'08)* (2008), pp. 39–46
3. P. Horn, Faktorisierung in Schief-Polynomringen. Ph.D. Thesis, Universität Kassel (2008). https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2009030226513
4. W. Koepf, *Hypergeometric Summation—An Algorithmic Approach to Summation and Special Function Identities* (Springer, Berlin, 2014)
5. J. Kovacic, An algorithm for solving second-order linear homogeneous equations. J. Symb. Comput. **2**, 3–43 (1986)
6. M. Mouafo Wouodjié, On the solutions of holonomic third-order linear irreducible differential equations in terms of hypergeometric functions. Ph.D. Thesis, Universität Kassel (2018). https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2018060655613
7. M. Mouafo Wouodjié, W. Koepf, On the solutions of holonomic third-order linear irreducible differential equations in terms of hypergeometric functions. J. Symb. Comput. (2019). https://doi.org/10.1016/j.jsc.2019.08.002
8. M.F. Singer, Solving homogeneous linear differential equations in terms of second order linear differential equations. Am. J. Math. **107**, 663–696 (1985)
9. M. van der Put, M.F. Singer, *Galois Theory of Linear Differential Equations*. Comprehensive Studies in Mathematics, vol. 328 (Springer, Berlin, 2003)
10. M. van Hoeij, Factorization of linear differential operators. Ph.D. Thesis, Universitijt Nijmegen (1996)
11. M. van Hoeij, Rational solutions of the mixed differential equation and its application to factorization of differential operators, in *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation (ISSAC'96)* (1996), pp. 219–225
12. M. van Hoeij, Factorization of linear differential operators with rational functions coefficients. J. Symb. Comput. **24**, 237–561 (1997)
13. M. van Hoeij, Q. Yuan, Finding all Bessel type solutions for linear differential equations with rational function coefficients, in *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation (ISSAC'10)* (2010), pp. 37–44
14. Q. Yuan, Finding all Bessel type solutions for linear differential equations with rational function coefficients. Ph.D. Thesis, Florida State University (2012)

The Gamma Function

Daniel Duviol Tcheutia

Abstract After the so-called elementary functions as the exponential and the trigonometric functions and their inverses, the Gamma function is the most important special function of classical analysis. In this note, we present the definition and properties of the Gamma and the Beta functions.

Keyword Gamma function

Mathematics Subject Classification (2000) 33E50

1 Definition

The Gamma function $\Gamma(z)$ developed by Euler (1707–1783) is defined by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \ \mathrm{Re}(z) > 0. \tag{1.1}$$

If we consider the integral (1.1), it is known that at infinity the behaviour of the exponential function dominates the behaviour of any power function, so that $t^{z-1}e^{-t} \to 0$ as $t \to \infty$ for any value of z, and hence no problem is expected from the upper limit of the integral. When $t \to 0$, we have $e^{-t} \simeq 1$ and then for $c > 0$ very small and $z = x \in \mathbb{R}$, we may write (1.1) as

$$\begin{aligned}\Gamma(x) &\simeq \int_0^c t^{x-1} dt + \int_c^\infty e^{-t} t^{x-1} dt \\ &= \left[\frac{1}{x} t^x\right]_0^c + \int_c^\infty e^{-t} t^{x-1} dt.\end{aligned}$$

D. D. Tcheutia (✉)
Institute of Mathematics, University of Kassel, Kassel, Germany
e-mail: duvtcheutia@yahoo.fr

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_9

For the first term to remain finite as $t \to 0$, we must have $x > 0$. The main references are [1–5].

2 Properties of the Gamma and Beta Functions

$\Gamma(x) > 0$ for all $x \in (0, \infty)$ and for $x = 1$, we have

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

Using integration by parts, it follows that for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 0$,

$$\Gamma(z+1) = \int_0^\infty t^z e^{-t} dt = \left[-e^{-t} t^z\right]_0^\infty + z \int_0^\infty t^{z-1} e^{-t} dt = z\Gamma(z).$$

Using the latter recurrence relation $\Gamma(z + 1) = z\Gamma(z)$ and the initial condition $\Gamma(1) = 1$, it follows that for $z = n \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$, one gets

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n(n-1)\cdots 2 \cdot 1 \cdot \Gamma(1) = n!.$$

The Gamma function therefore can be seen as an extension of the factorial function to real and complex arguments.

From the recurrence relation $\Gamma(z + 1) = z\Gamma(z)$ we have

$$\Gamma(z) = \frac{1}{z}\Gamma(z+1). \tag{2.1}$$

Since $\Gamma(1) = 1$, we deduce from (2.1) that $\Gamma(0) = \infty$. With this result, we get

$$\Gamma(-1) = \frac{1}{-1}\Gamma(0) = \infty,\ \Gamma(-2) = \frac{1}{-2}\Gamma(-1) = \infty,\ \text{etc.}$$

That means $\Gamma(n) = \infty$ if n is zero or a negative integer.

The right-hand side of (2.1) is defined for $\mathrm{Re}(z + 1) > 0$, i.e., for $\mathrm{Re}(z) > -1$. By iteration, we get

$$\Gamma(z) = \frac{1}{z(z+1)\cdots(z+n-1)}\Gamma(z+n)\ (n \in \mathbb{N}). \tag{2.2}$$

Equation (2.2) enables us to define $\Gamma(z)$ for $\mathrm{Re}(z) > -n$ as an analytic function except for $z = 0, -1, -2, \ldots, -n+1$. Thus, $\Gamma(z)$ can be continued analytically to the whole complex z-plane except for simple poles at $z = 0, -1, -2, \ldots$. It is now possible to draw a graph of $\Gamma(x)$ $(x \in \mathbb{R})$ as shown in Fig. 1.

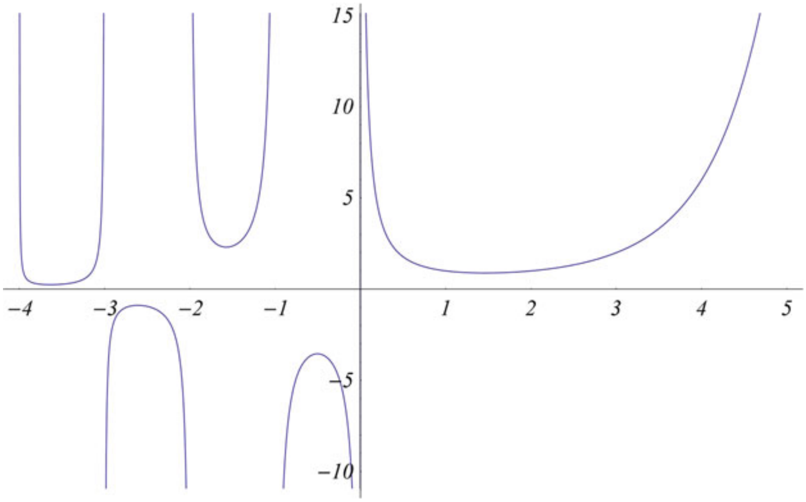


Fig. 1 The Gamma function on the real axis

The shifted factorial

$$(z)_n := z(z+1)\cdots(z+n-1) = \frac{\Gamma(z+n)}{\Gamma(z)} \ (n \in \mathbb{N}_0), \tag{2.3}$$

which occurs in (2.2), is called the Pochhammer symbol.

At the poles $-n$ $(n \in \mathbb{N}_0)$ of the Gamma function, we get

$$\begin{aligned}\lim_{z\to -n}(z+n)\Gamma(z) &= \lim_{z\to -n}\frac{(z+n)\Gamma(z+n)}{z(z+1)\cdots(z+n-1)}\\ &= \lim_{z\to -n}\frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n-1)} = \frac{(-1)^n}{n!}.\end{aligned}$$

This result may be interpreted as the residue of $\Gamma(z)$ at the simple poles $z = -n$.

We have the identity

$$\begin{aligned}(z)_n &= z(z+1)\cdots(z+n-2)(z+n-1)\\ &= (-1)^n(-z)(-z-1)\cdots(1-z-n+1)(1-z-n)\\ &= (-1)^n(1-z-n)(1-z-n+1)\cdots(1-z-n+n-2)(1-z-n+n-1)\\ &= (-1)^n(1-z-n)_n.\end{aligned}$$

Using the definition (2.3), we can rewrite this as

$$\frac{\Gamma(z+n)}{\Gamma(z)} = (-1)^n\frac{\Gamma(1-z)}{\Gamma(1-z-n)}$$

or equivalently

$$\Gamma(z)\Gamma(1-z) = (-1)^n \Gamma(z+n)\Gamma(1-(z+n)).$$

Since $z = 0, -1, -2, \ldots$ are the poles of $\Gamma(z)$, we deduce that $1/\Gamma(z)$ is analytic in the entire complex plane with zeros $0, -1, -2, \ldots$. It follows that the zeros of $1/\Gamma(1-z)$ are $1, 2, \ldots$. This means that

$$\frac{1}{\Gamma(z)\Gamma(1-z)}$$

is analytic in the entire complex plane with zeros $\ldots, -2, -1, 0, 1, 2, \ldots$ similar as the function $\sin(\pi z)$. It can be shown that the following relationship between the Gamma and circular functions is valid, where the last statement is the Euler product for the sine function:

$$\frac{1}{\Gamma(z)\Gamma(1-z)} = \frac{\sin(\pi z)}{\pi} = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right). \tag{2.4}$$

One similarly has

$$\Gamma(z) = \lim_{n\to\infty} \frac{n! n^z}{(z)_{n+1}} = \frac{e^{-\gamma z}}{z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^{-1} e^{\frac{z}{n}},$$

where

$$\gamma := \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln n\right) \approx 0.57721\,56649\,01532\,86060\,65120\,90082$$

denotes the Euler–Mascheroni constant. Equation (2.4) is called the reflection formula of the Gamma function.

Equation (2.4) with $z = \frac{1}{2}$ yields immediately

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \tag{2.5}$$

which, in view of (1.1), implies

$$\int_0^{\infty} \frac{e^{-t}}{\sqrt{t}} dt = \sqrt{\pi}.$$

Equation (2.2) combined with (2.5) yields

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!\sqrt{\pi}}{2^{2n} n!}, \quad \Gamma\left(-n + \frac{1}{2}\right) = (-1)^n \sqrt{\pi} \frac{2^{2n} n!}{(2n)!}, \quad n \in \mathbb{N}_0.$$

If we set $t = u^2$ in the definition (1.1) so that $dt = 2udu$, we get

$$\Gamma(z) = 2\int_0^\infty e^{-u^2}(u^2)^{z-1}udu = 2\int_0^\infty e^{-u^2}u^{2z-1}du.$$

This means that

$$\Gamma(z) = 2\int_0^\infty e^{-t^2}t^{2z-1}dt, \tag{2.6}$$

and for $z = \frac{1}{2}$, we derive the following result

$$\int_0^\infty e^{-t^2}dt = \frac{1}{2}\sqrt{\pi}.$$

The binomial coefficients can be expressed as

$$\binom{z}{n} = \frac{z(z-1)\cdots(z-n+1)}{n!} = (-1)^n\frac{(-z)_n}{n!},$$

or, equivalently, as

$$\binom{z}{n} = \frac{\Gamma(z+1)}{n!\Gamma(z-n+1)},$$

for arbitrary $z \in \mathbb{C}$, $z+1 \neq 0, -1, \ldots,$ and $z-n+1 \neq 0, -1, \ldots$. Since $\Gamma(-k) = \infty$ for $k \in \mathbb{N}_0$, we may set $1/\Gamma(-k) = 0$ which reads again as $1/k! = 0$ for $k = -1, -2, \ldots$. We deduce that for $k, n \in \mathbb{N}_0$, we have

$$\binom{n}{k} = 0 \text{ for } k < 0 \text{ and } k > n.$$

3 The Beta Function

The Beta function is defined by the integral

$$B(z, w) = \int_0^1 t^{z-1}(1-t)^{w-1}dt, \ \mathrm{Re}(z) > 0, \ \mathrm{Re}(w) > 0. \tag{3.1}$$

The substitution $t = 1 - u$ shows that

$$B(z, w) = \int_0^1 u^{w-1}(1-u)^{z-1}du = B(w, z). \tag{3.2}$$

By setting $t = \cos^2\theta$ so that $dt = -2\cos\theta\sin\theta d\theta$, we get

$$B(z, w) = \int_{\pi/2}^{0} (\cos^2\theta)^{z-1}(\sin^2\theta)^{w-1}(-2\cos\theta\sin\theta)d\theta$$
$$= 2\int_0^{\pi/2} \cos^{2z-1}\theta\,\sin^{2w-1}\theta d\theta. \tag{3.3}$$

Now we want to show that

$$B(z, w) = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)}. \tag{3.4}$$

We first consider the product

$$\Gamma(z)\,\Gamma(w) = \int_0^\infty t^{z-1}\,e^{-t}\,dt \cdot \int_0^\infty u^{w-1}\,e^{-u}\,du$$

and use the substitutions $t = x^2$ and $u = y^2$ to obtain

$$\Gamma(z)\,\Gamma(w) = 4\int_0^\infty e^{-x^2}\,x^{2z-1}\,dx \int_0^\infty e^{-y^2}\,y^{2w-1}\,dy$$
$$= 4\int_0^\infty\int_0^\infty e^{-x^2-y^2}\,x^{2z-1}\,y^{2w-1}\,dx\,dy.$$

Applying polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ to this double integral, we get

$$\Gamma(z)\,\Gamma(w) = 4\int_0^{\pi/2}\int_0^\infty e^{-r^2}\,r^{2z+2w-2}\,\cos^{2z-1}\theta\cdot\sin^{2w-1}\theta\cdot r\,dr\,d\theta$$
$$= 2\int_0^\infty e^{-r^2}\,r^{2z+2w-1}\,dr\cdot 2\int_0^{\pi/2}\cos^{2z-1}\theta\cdot\sin^{2w-1}\theta\,d\theta$$
$$= \Gamma(z+w)\,B(w, z) = \Gamma(z+w)\,B(z, w)$$

where Eqs. (2.6) and (3.3) are utilized. This proves (3.4).

Relation (3.4) not only confirms the symmetry property in (3.2), but also continues the Beta function analytically for all complex values of z and w, except when $z, w \in \{0, -1, -2, \ldots\}$. Thus we may write

$$B(z, w) = \begin{cases} \int_0^1 t^{z-1}(1-t)^{w-1}dt & (\mathrm{Re}(z) > 0,\ \mathrm{Re}(w) > 0) \\ \\ \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)} & (\mathrm{Re}(z) < 0,\ \mathrm{Re}(w) < 0,\ z, w \notin \{0, -1, -2, \ldots\}). \end{cases}$$

The following relations are valid:

$$B(z+1, w) = \frac{z}{z+w} B(z, w),$$
$$B(z, w+1) = \frac{w}{z+w} B(z, w).$$

Indeed, we have

$$\begin{aligned} B(z+1, w) &= \frac{\Gamma(z+1)\Gamma(w)}{\Gamma(z+w+1)} \\ &= \frac{z\Gamma(z)\Gamma(w)}{(z+w)\Gamma(z+w)} \\ &= \frac{z}{z+w} B(z, w). \end{aligned}$$

For further reading on the Gamma and Beta functions, one might go through the following books. This article presents the most important part of [3, Chap. 1].

References

1. G. Andrews, R.A. Askey, R. Roy, *Special Functions, Encyclopedia of Mathematics and its Applications* (Cambridge University Press, Cambridge, 1999)
2. W.W. Bell, *Special Functions for Scientists and Engineers* (D. van Nostrand Company, Toronto, 1968)
3. W. Koepf, *Hypergeometric Summation – An Algorithmic Approach to Summation and Special Function Identities*. Springer Universitext, 2nd edn. (Springer, London, 2014)
4. F.W.J. Olver, D.W. Lozier, R.F. Boisvert, C.W. Clark, *NIST Handbook of Mathematical Functions* (National Institute of Standards and Technology U.S. Department of Commerce and Cambridge University Press, Cambridge, 2010)
5. H.M. Srivastava, J. Choi, *Zeta and q-Zeta Functions and Associated Series and Integrals* (Elsevier, Amsterdam, 2012)

Part II
Recent Research Topics in Orthogonal Polynomials and Applications

Hypergeometric Multivariate Orthogonal Polynomials

Iván Area

To Prof. Eduardo Godoy, with appreciation and respect

Abstract In this lecture a comparison between univariate and multivariate orthogonal polynomials is presented. The first step is to review classical univariate orthogonal polynomials, including classical continuous, classical discrete, their q-analogues and also classical orthogonal polynomials on nonuniform lattices. In all these cases, the orthogonal polynomials are solution of a second-order differential, difference, q-difference, or divided-difference equation of hypergeometric type. Next, a review multivariate orthogonal polynomials is presented. In the approach we have considered, the main tool is the partial differential, difference, q-difference or divided-difference equation of hypergeometric type the polynomial sequences satisfy. From these equations satisfied, the equation satisfied by any derivative (difference, q-difference or divided-difference) of the polynomials is obtained. A big difference appears for nonuniform lattices, where bivariate Racah and for bivariate q-Racah polynomials satisfy a fourth-order divided-difference equation of hypergeometric type. From this analysis, we propose a definition of multivariate classical orthogonal polynomials. Finally, some open problems are stated.

Keywords Orthogonal polynomials · Hypergeometric equation

Mathematics Subject Classification (2000) Primary 33C45, 33C50; Secondary 39A13, 39A14

I. Area (✉)
Departamento de Matemática Aplicada II, E.E. Aeronáutica e do Espazo, Universidade de Vigo, Ourense, Spain
e-mail: area@uvigo.gal

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_10

1 Classical Univariate Case: From Hermite to q-Racah Polynomials

The class of classical univariate orthogonal polynomials with respect to a positive weight function contains the well-known families of Jacobi, Laguerre and Hermite polynomials. They are orthogonal with respect to the beta distribution, the normal distribution and the gaussian distribution, respectively. As for the Hermite polynomials, they have been studied since the works of P.S. Laplace in 1810, and later in detail by P. Chebyshev in 1859. As in many other branches of science, the works of Chebyshev were overlooked and the family was named after a work of Hermite in 1864; now, we can ensure that they were not new. Nevertheless, Hermite published some other works in 1865 introducing for the first time multidimensional orthogonal polynomials [22].

For each family of classical univariate orthogonal polynomials with respect to a positive weight function, a number of properties were proved. For instance, for monic Hermite polynomials we have

1. Orthogonality relation

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}dx = d_n^2\delta_{n,m} = \begin{cases} \sqrt{\pi}2^{-n}n!, & n=m, \\ 0, & n \neq m. \end{cases}$$

2. Rodrigues' representation

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

3. Recurrence relation

$$H_{n+1}(x) = xH_n(x) - \frac{n}{2}H_{n-1}(x).$$

4. Explicit representation

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

5. Generating function

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

6. Second-order linear differential equation

$$y''(x) - 2xy'(x) + 2ny(x) = 0.$$

There exist many other algebraic properties such as structure relation or derivative representation, just to cite some of them.

Following a historical approach, the next step was to prove that many of the above properties are indeed characterizations of the class of classical univariate orthogonal polynomials with respect to a positive weight function. Among many important contributions, the following ones might be cited [1, 9, 38]. In particular, classical univariate orthogonal polynomials with respect to a positive weight function are solution of a second order linear differential equation of hypergeometric type

$$\sigma(x)y''(x) + \tau(x)y'(x) + \lambda_n y(x) = 0, \tag{1.1}$$

where σ and τ are polynomials of at most degree 2 and 1, respectively, and $\lambda_n = -n\tau' - 1/2n(n-1)\sigma''$ is a constant. Therefore, if $y(x)$ is solution of the above equation then its derivatives of any order $v_n(x) = y^{(n)}(x)$ are solution of an equation of the same type, namely

$$\sigma(x)v_n''(x) + (\tau(x) + n\sigma'(x))v_n'(x) + (\lambda_n + n\tau' + \frac{1}{2}n(n-1)\sigma'')v_n = 0.$$

One of the advantages of the approach from the second order differential equation (1.1) is the ease of explicitly computing the coefficients of the three term recurrence relation satisfied by the family of orthogonal polynomials, which in the monic case reads as

$$xp_n(x) = p_{n+1}(x) + B_n p_n(x) + C_n p_{n-1}(x), \quad n \geq 1, \tag{1.2}$$

with $p_0(x) = 1$ and $p_1(x) = x - B_0$. If

$$y_n(x) = x^n + K_n x^{n-1} + \varpi_n x^{n-2} + \cdots,$$

then, after substituting into (1.1) it is found upon equating the coefficients of x^{n-1} and x^{n-2} that

$$K_n = \frac{n(b(n-1)+q)}{2a(n-1)+p}, \quad \varpi_n = \frac{(n-1)(K_n(b(n-2)+q)+cn)}{2(a(2n-3)+p)},$$

where $\sigma(x) = ax^2 + bx + c$, and $\tau(x) = px + q$. From the three term recurrence relation by equating again the coefficients,

$$\begin{aligned} B_n &= K_n - K_{n+1}, \\ C_n &= -B_n K_n + \varpi_n - \varpi_{n+1}. \end{aligned}$$

As a consequence, the coefficients of the three term recurrence are completely identified from the coefficients of the second order differential equation (1.1). Some important historical references about the characterization of classical orthogonal polynomials by means of the three term recurrence relation they satisfy can be found e.g. in [9, p. 182]. It is also important to recall that the orthogonality weight function $\varrho(x)$ satisfies the Pearson differential equation

$$(\sigma(x)\varrho(x))' = \tau(x)\varrho(x).$$

It is also important to recall that already in 1929 Bochner [10] posed the problem of finding all linear second order differential equations of Sturm–Liouville type with polynomial coefficients having orthogonal polynomial solutions.

The so-called classical orthogonal polynomials of a discrete variable [44] can be introduced as solutions of a second order linear difference equation of hypergeometric type

$$\sigma(x)\Delta\nabla y(x) + \tau(x)y(x) + \lambda_n y(x) = 0, \tag{1.3}$$

where the forward and backward difference operators are respectively defined as

$$\Delta y(x) = y(x+1) - y(x), \qquad \nabla y(x) = y(x) - y(x-1).$$

The above equation appears if we discretize (1.1) up to the second order in e.g. h and later set $h = 1$. These polynomials y_n, as stated in [1, p. 10], share with the classical continuous families a number of properties that indeed are characterizations: they all satisfy a second order linear difference equation as (1.3), their differences Δy_n are again a sequence of orthogonal polynomials, they possess a Rodrigues-type formula, the orthogonality weight function satisfies a Pearson-type difference equation

$$\Delta(\sigma(x)\varrho(x)) = \tau(x)\varrho(x),$$

and e.g. they all satisfy a structure relation. Essentially, these are the Charlier, Meixner, Kravchuk, and Hahn orthogonal polynomials. The first two families are infinite, while the third and fourth families are just orthogonal up to a certain positive integer N which gives the orthogonality domain $[0, N]$. As in the classical continuous case, it is possible to identify the orthogonality weight functions with well-known weights in Statistics [23]: Charlier $\leftrightarrow$ Poisson, Meixner $\leftrightarrow$ Pascal, Kravchuk $\leftrightarrow$ binomial, Hahn $\leftrightarrow$ hypergeometric —or Pólya. Moreover, in [44, p. 34] another family of polynomials appear as an analytic continuation of Hahn polynomials with respect to the parameters α and β from the domain $\alpha > -1$, $\beta > -1$ into the domain $\alpha < 1 - N$, $\beta < 1 - N$. This fifth family of classical orthogonal polynomials of a discrete variable has been shown to be very useful in certain connection problems related with Bernstein bases [50]. In this framework the orthogonal polynomials are defined in the real line but the orthogonality relation is evaluated on a set of finite or infinite positive integers —linear lattice $x(s) = s$.

As it happens in the univariate continuous situation, if $y_n(x)$ is a solution to (1.3) and if $v_m(x) = \Delta^m y(x)$ then $v_m(x)$ is solution of an equation of the same type:

$$\sigma(x)\Delta\nabla v_m(x) + \tau_m(x)\Delta v_m(x) + \mu_m v_m(x) = 0,$$

where $\tau_m(x) = \tau_{m-1}(x+1) + \Delta\sigma(x)$, with $\tau_0(x) = \tau(x)$, and $\mu_m = \mu_{m-1} + \Delta\tau_{m-1}(x)$, with $\mu_0 = \lambda_n$. If we denote again $\sigma(x) = ax^2 + bx + c$ and $\tau(x) = px + q$, then $\tau_m(x) = (2ma + p)x + m(ma + b + p) + q$. A similar approach can be followed to obtain the coefficients of the three term recurrence relation (1.2) from the coefficients of the polynomials in the second-order difference equation (1.3).

Equation (1.1) can be also discretized by using a q-linear lattice, $x(s) = q^s$; in doing so, the q-difference operator appears in a natural way

$$D_q y(x) = \frac{y(qx) - q(x)}{(q-1)x}, \quad (q \neq 1, \quad x \neq 0).$$

Then we get following second-order q-difference equation

$$\sigma(x)D_q D_{q^{-1}} y(x) + \tau(x)D_q y(x) + \lambda_n y(x) = 0. \tag{1.4}$$

Detailed analysis of the polynomial solutions of the above equation was done in the original works Hahn and Lesky (see e.g. [21] or [33]), and are nicely compiled in [27]. For this class of univariate orthogonal polynomials (those belonging to the so-called q-Hahn class) there exist also a number of properties that characterize them [1]. It is important to mention here that in [1, p. 11] five of these properties are stated, but the last one is not completely correct: The moments of q-classical orthogonal polynomials are hypergeometric if the moments are either the power moments (moments against x^k), the q-shifted factorial moments (moments against $(x;q)_k$) or moments against $(xq;q)_k$—which is the case of Al-Salam-Carlitz I orthogonal polynomials [4, 43].

As happens in the univariate continuous situation and also for a linear lattice $x(s) = s$, if $y_n(x)$ is a solution to (1.4) then $v_m(x) = D_q^m y_n(x)$ is solution to an equation of the same type. Again, the same idea can be used to obtain the coefficients of the three term recurrence relation (1.2) from the coefficients of the polynomials in the second-order q-difference equation (1.4).

The final class of univariate orthogonal polynomials to be presented in this review appears if we discretize (1.1) by using a nonuniform lattice $x(s)$ [44], giving rise to [44, Eq. (3.1.3)]

$$\begin{aligned}\sigma(x(s))\frac{1}{x(s+h/2)-x(s-h/2)}\left[\frac{y(s+h)-y(s)}{x(s+h)-x(s)} - \frac{y(s)-y(s-h)}{x(s)-x(s-h)}\right]\\ + \frac{\tau(x(s))}{2}\left[\frac{y(s+h)-y(s)}{x(s+h)-x(s)} + \frac{y(s)-y(s-h)}{x(s)-x(s-h)}\right] + \lambda y(s) = 0.\end{aligned} \tag{1.5}$$

The above equation approximates (1.1) to second order in h and it can be rewritten in many ways by using appropriate divided-difference operators. For our purposes, we shall consider the following divided-difference operators $\mathbb{D}$ and $\mathbb{S}$ [14, 36, 37] defined by

$$\mathbb{D}f(s) = \frac{f(s+1/2)-f(s-1/2)}{x(s+1/2)-x(s-1/2)}, \quad \mathbb{S}f(s) = \frac{f(s+1/2)+f(s-1/2)}{2}. \tag{1.6}$$

Notice that the above operators transform polynomials of degree n in the lattice $x(s)$ into polynomials of respectively degree $n-1$ and n in the same variable $x(s)$. With these notations, classical orthogonal polynomials on a nonuniform lattice $x(s)$ are solution of the following second-order divided-difference equation of hypergeometric type

$$\phi(x(s))\mathbb{D}^2 y_n(s) + \tau(x(s))\mathbb{S}\mathbb{D}y_n(s) + \lambda_n y_n(s) = 0, \tag{1.7}$$

where ϕ and τ are polynomials of at most degree 2 and 1, respectively, and λ_n is a constant. In [9] the classical orthogonal polynomials on nonuniform lattices are analyzed in detail, including e.g. the Pearson-type divided-difference equation, the Rodrigues-type formula, integral representations, orthogonality property and the moment property against appropriate generalized bases. Some other recent works on characterizations of classical orthogonal polynomials on nonuniform lattices can be found in e.g. [15, 42] and references therein. In that case [44] it is also possible to obtain the coefficients of the three term recurrence relation

$$x(s)p_n(x(s)) = p_{n+1}(x(s)) + \tilde{B}_n p_n(x(s)) + \tilde{C}_n p_{n-1}(x(s))$$

from the coefficients of the polynomials appearing in the second-order divided-difference equation (1.7). Moreover, in [16] suitable bases that enable the direct series representation of orthogonal polynomial systems on nonuniform lattices are presented that allow to derive solutions of holonomic divided-difference equations on nonuniform lattices. The q-Racah polynomials and the Askey–Wilson polynomials are solution of a divided-difference equation as (1.7) in a q-quadratic lattice and in a quadratic lattice, respectively.

The limit transitions between classical univariate orthogonal polynomials is well-known and can be read as the limits of univariate distributions [23]. All the limit relations have been presented in [27], and they include limits inside the same class (continuous, discrete, q-analogues and non-uniform lattices) or between polynomials in different classes. The remainders in the limits have been analyzed eg. in [13, 19, 49]. Also, a interesting presentation of the Askey scheme has been given in [29].

As recalled in [9, p. 189], in [2] the following definition of univariate classical orthogonal polynomials was suggested:

> An orthogonal polynomial sequence is classical if it is a special case or a limiting case of the ${}_4\phi_3$ polynomials given by the q-Racah polynomials or the Askey–Wilson polynomials.

Let us recall some of the aforementioned special cases and limiting cases. Let us introduce univariate Racah polynomials in terms of hypergeometric series as [27, page 190]

$$
\begin{aligned}
r_n(\alpha,\beta,\gamma,\delta;s) = r_n(s) &= (\alpha+1)_n\,(\beta+\delta+1)_n\,(\gamma+1)_n \\
&\times\ {}_4F_3\left(\left.\begin{matrix}-n, n+\alpha+\beta+1, -s, s+\gamma+\delta+1\\ \alpha+1, \beta+\delta+1, \gamma+1\end{matrix}\right|1\right), \quad n=0,1,\ldots,N,
\end{aligned}
\tag{1.8}
$$

where $r_n(\alpha,\beta,\gamma,\delta;s)$ is a polynomial of degree $2n$ in s and of degree n in the quadratic lattice [9, 44]

$$x(s) = s(s+\gamma+\delta+1), \tag{1.9}$$

Univariate Racah polynomials are solution of the second-order linear divided-difference equation [14]

$$\phi(x(s))\mathbb{D}^2 r_n(s) + \tau(x(s))\mathbb{S}\mathbb{D} r_n(s) + \lambda_n r_n(s) = 0, \tag{1.10}$$

where ϕ is a polynomial of degree 2 in the lattice $x(s)$ given by

$$
\begin{aligned}
\phi(x(s)) = -(x(s))^2 &+ \frac{1}{2}(-\alpha(2\beta+\delta+\gamma+3)+\beta(\delta-\gamma-3)-2(\delta\gamma+\delta+\gamma+2))x(s) \\
&- \frac{1}{2}(\alpha+1)(\gamma+1)(\beta+\delta+1)(\delta+\gamma+1),
\end{aligned}
$$

τ is a polynomial of degree 1 in the lattice $x(s)$ given by

$$\tau(x(s)) = -(\alpha+\beta+2)x(s) - (\alpha+1)(\gamma+1)(\beta+\delta+1),$$

the eigenvalues λ_n are given by

$$\lambda_n = n(\alpha+\beta+n+1),$$

and the difference operators $\mathbb{D}$ and $\mathbb{S}$ are defined in (1.6). The second-order divided-difference equation (1.10) can be written in terms of rational functions as [27, Eq. (9.2.5)]

$$n(n+\alpha+\beta+1)r_n(s) = B(s)r_n(s+1) - (B(s)+D(s))r_n(s) + D(s)r_n(s-1),$$

where

$$B(s) = \frac{(\alpha + s + 1)(\gamma + s + 1)(\beta + \delta + s + 1)(\delta + \gamma + s + 1)}{(\delta + \gamma + 2s + 1)(\delta + \gamma + 2s + 2)},$$

$$D(s) = \frac{s(\delta + s)(-\beta + \gamma + s)(-\alpha + \delta + \gamma + s)}{(\delta + \gamma + 2s)(\delta + \gamma + 2s + 1)}.$$

If we set $\alpha = a+b-1, \beta = c+d-1, \gamma = a+d-1, \delta = a-d, x \to -a+ix$, then the Racah polynomials become the Wilson polynomials. The Wilson polynomials can be also introduced in terms of hypergeometric series as

$$w_n(s^2; a, b, c, d) = (a+b)_n(a+c)_n(a+d)_n \times {}_4F_3\left(\begin{matrix} -n, a+b+c+d+n-1, a-is, a+is \\ a+b, a+c, a+d \end{matrix} \middle| 1 \right), \tag{1.11}$$

where $i^2 = -1$, and $(A)_n = A(A+1)\cdots(A+n-1)$ denotes the Pochhammer symbol. These polynomials are solution of a second-order linear divided-difference equation of hypergeometric type in the lattice $x(s) = s^2$, which can be obtained from the equation for Racah polynomials by introducing the aforementioned changes:

$$(x(s)^2 + (cd - b(c+d) - a(b+c+d))x(s) + abcd)\mathbf{D}^2 y(s)$$
$$+((a+b+c+d)x(s) - bcd - a(cd + b(c+d)))\mathbf{SD}y(s) - n(n+a+b+c+d-1)y(s) = 0, \tag{1.12}$$

where the divided-difference operators in the lattice $x(s) = s^2$ are given by

$$\mathbf{D}f(x) = \frac{f(x+i/2) - f(x-i/2)}{2ix}, \quad \mathbf{S}f(x) = \frac{f(x+i/2) + f(x-i/2)}{2}.$$

This equation for Wilson polynomials can be also written in expanded form as [27, Eq. (9.1.6)]

$$n(n+a+b+c+d-1)y(s) = B_w(s)y(s+i) - (B_w(s) + D_w(s))y(s) + D_w(s)y(s-i), \tag{1.13}$$

where

$$B_w(s) = \frac{(a-is)(b-is)(c-is)(d-is)}{2is(2is-1)},$$

$$D_w(s) = \frac{(a+is)(b+is)(c+is)(d+is)}{2is(1+2is)}.$$

The continuous dual Hahn polynomials are also polynomials in the lattice $x(s) = s^2$, which can be introduced in terms of hypergeometric series as

$$d_n(a,b,c|s^2) = (a+b)_n(a+c)_n\, {}_3F_2\left(\left.\begin{matrix} -n, a+is, a-is \\ a+b, a+c \end{matrix}\right| 1\right). \tag{1.14}$$

The continuous dual Hahn polynomials are the limit of Wilson polynomials as

$$\lim_{d\to\infty} \frac{w_n(s^2;a,b,c,d)}{(a+d)_n} = d_n(a,b,c|s^2).$$

If we consider the limit in the divided-difference equation (1.12) or in the difference equation (1.13) as $d \to \infty$ we obtain the difference equation(s) for continuous dual Hahn polynomials

$$(-(a+b+c)x(s)+abc)\mathbf{D}^2 y(s) + (x(s)-ab-ac-bc)\mathbf{SD}y(s) - ny(s) = 0,$$

and

$$ny(s) = B_C(s)y(s+i) - (B_C(s)+D_C(s))y(s) + D_C(s)y(s-i),$$

where

$$B_C(s) = \frac{(a-is)(s+ib)(s+ic)}{2s(2s+i)}, \quad D_C(s) = \frac{(a+is)(s-ib)(s-ic)}{2s(2s-i)}.$$

2 Hypergeometric Multivariate Orthogonal Polynomials

As it has been shown in the previous section, in the univariate case "classical" is used for families that are solution of a second-order linear differential (difference, q-difference, or divided-difference) equation of hypergeometric type. From that property it is simple to deduce that the derivative (difference, q-difference, or divided-difference) of a classical orthogonal polynomial sequence is again an orthogonal polynomial sequence (Hahn's characterization). Also, the orthogonality weight function appears in a natural way as the symmetrization factor of the differential equation. Moreover, the Rodrigues formula can be also deduced from the differential equation. In the multivariate case, as it will be shown in this section, this is essentially the situation in the continuous, discrete and q-analogues situations; nevertheless, the multivariate orthogonal polynomials on nonuniform lattices are solution of a fourth-order divided difference equation, which turns out to be of hypergeometric type.

The basic techniques will be shown for bivariate continuous orthogonal polynomials, solutions to a class of partial differential equations. Similar ideas can be followed for partial difference or q-difference equations. Bivariate orthogonal polynomials on nonuniform lattices will be analyzed in a different section.

Once this analysis has been done, as in the univariate case one can suggest in the multivariate setting that:

> A multivariate orthogonal polynomial sequence is classical if it is a special case or a limiting case of the multivariate Racah polynomials or the multivariate Askey–Wilson polynomials.

Before entering in the details, a brief summary of multivariate orthogonal polynomials will be presented. In 1926 it appeared the basic reference [3] which contains a number of properties about generalizations of Hermite polynomials to several variables. The domain of orthogonality in the bidimensional case, for these generalizations is the full plane $\mathbf{R}^2$. Orthogonal polynomials on a triangular region (simplex) were introduced by Proriol [45] which were applied to the problem of solving the Schrödinger equation for the Helium atom [40, 41]. The same class was independently obtained by Karlin and McGregor a few years later [25, 26] in view of applications to genetics, as indicated by Koornwinder [28], which is probably the first systematic study on bivariate orthogonal polynomials. In the latter reference it appears a general method of generating orthogonal polynomials of two variables from orthogonal polynomials of one variable. This method was also discussed by Dunkl and Xu in their monograph [11]. As for the analysis of multivariate orthogonal polynomials starting from second-order partial differential equations we might refer to Krall and Sheffer [32] and Engelis [12]. Later, Krall and Sheffer introduced the multivariable Hahn polynomials, i.e. an extension of orthogonal polynomials of a discrete variable to the multidimensional situation. Griffiths [20] introduced a generalization of Kravchuk polynomials by considering orthogonal polynomials on the multinomial distribution. In 1989 Tratnik [55] presented a multivariable biorthogonal generalization for Meixner, Kravchuk and Meixner–Pollaczek. Tratnik showed that these families of polynomials are orthogonal with respect to subspaces of lower degree and biorthogonal within a given subspace. Moreover, in [56, 57] extensions of the remaining families of the Askey scheme were given.

If we focus on the partial differential equation, Krall and Sheffer [32] studied the problem of finding all polynomial eigenfunctions of second-order linear differential operators in two variables having polynomial coefficients of degree equal to the order of derivative under certain further restrictions relating to its symmetrizability and the orthogonality of their eigenfunctions. They classified all possible normal forms of the operators satisfying the required properties. Engelis [12] gave a detailed list of second-order linear partial differential equations for which orthogonal polynomial in two variables are solutions. This question was afterwards studied and systematically described by Suetin in his book [52] (first published in 1988 and translated into English in 1999). A second-order linear partial differential equation is said to be admissible if there exists a sequence $\{\lambda_n\}$ $(n = 0, 1, \ldots)$ such that for $\lambda = \lambda_n$, there are precisely $n + 1$ linearly independent solutions in the form of polynomials of total degree n and has no non-trivial solutions in the set of polynomials whose total degree is less than n. This concept was introduced by Krall and Sheffer in the case of second-order partial differential equations and also by Xu in the case of second-order partial difference equations [58]. As a generalization of the concept of differential equation of hypergeometric type Lyskova [34, 35] considered a special class of linear partial differential equations, called basic class,

$$\sum_{i,j=1}^{n} \tilde{a}_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} \tilde{b}_i(x)\frac{\partial u}{\partial x_i} + \lambda u = 0,$$

where $\tilde{a}_{ij}(x) = \tilde{a}_{ji}(x)$ and the coefficients $\tilde{a}_{ij}(x)$ and $\tilde{b}_i(x)$ are chosen so that the derivatives of any order of the solutions of the equation are also solutions of an equation of the same type.

2.1 Bivariate Orthogonal Polynomials: Continuous, Discrete and q-Analogues

The class of admissible potentially self-adjoint linear second-order partial differential equations of hypergeometric type has been analyzed in [6]. Later, the discrete and q-analogues have been studied in [7, 8, 46–48].

We shall consider monomials of the form $x^{n-k}y^k$ of total degree n, and the column vector of all these monomials will be denoted by

$$\mathbf{x}^n = (x^{n-k}y^k), \quad 0 \le k \le n, \quad n \in \mathbf{N}_0.$$

A polynomial of total degree n can contain many monomials of the form $x^{n-k}y^k$; it is said to be monic if it contains only one monomial of total degree n.

Let us consider the following second-order linear partial differential equation

$$\tilde{a}_{11}(x,y)\frac{\partial^2 u(x,y)}{\partial^2 x} + \tilde{a}_{12}(x,y)\frac{\partial^2 u(x,y)}{\partial x \partial y} + \tilde{a}_{22}(x,y)\frac{\partial^2 u(x,y)}{\partial^2 y} + \tilde{b}_1(x,y)\frac{\partial u(x,y)}{\partial x} + \tilde{b}_2(x,y)\frac{\partial u(x,y)}{\partial y} + \lambda u(x,y) = 0,$$

which is assumed to be of hypergeometric type (basic class following [34, 35]) and admissible. Then, the equation has the form

$$\begin{aligned}(ax^2 + b_1x + c_1)\partial_{xx}u(x,y) &+ 2(axy + b_3x + c_3y + d_3)\partial_{xy}u(x,y)\\ &+ (ay^2 + b_2y + c_2)\partial_{yy}u(x,y)\\ &+ (ex + f_1)\partial_x u(x,y) + (ey + f_2)\partial_y u(x,y) + \lambda_n u(x,y) = 0, \qquad (2.1)\end{aligned}$$

where $\lambda_n = -n((n-1)a + e)$ and the coefficients a, b_j, c_j, d_j, e, f_j are arbitrary fixed real numbers, but the numbers a and e are such that the condition

$$\varpi_k := ak + e \neq 0 \tag{2.2}$$

holds true for any non-negative integer k. Then, it is possible to follow the ideas in [52, Chapter 5] to define an orthogonality weight function $\varrho(x, y)$ over a certain domain $D \subset \mathbf{R}^2$. Since the equation is admissible, then for each n it has $n + 1$ linearly independent solutions which are polynomials of total degree n; following Kowalski [30, 31] $\mathbb{P}_n$ will denote the (column) polynomial vector

$$\mathbb{P}_n = (P^n_{n,0}(x, y), P^n_{n-1,1}(x, y), \ldots, P^n_{1,n-1}(x, y), P^n_{0,n}(x, y))^{\mathrm{T}}. \tag{2.3}$$

Then, the orthogonality condition reads as

$$\iint_D \mathbf{x}^m \mathbb{P}_n^{\mathrm{T}} \varrho(x, y) dx dy = \begin{cases} 0 \in \mathcal{M}^{(m+1,n+1)}, & n > m, \\ H_n \in \mathcal{M}^{(n+1,n+1)}, & n = m, \end{cases}$$

where $\mathcal{M}^{(m+1,n+1)}$ denotes a matrix of size $(m + 1) \times (n + 1)$ and H_n is a non singular matrix of size $(n + 1) \times (n + 1)$. Following [52] it is possible to compute the orthogonality weight function from the polynomials in the partial differential equation. Moreover, since the equation is assumed to be of hypergeometric type, [6] it is possible to obtain the orthogonality weight for the partial derivatives of the orthogonal polynomial solutions, as well as to derive a Rodrigues-type formula [6, Eq. (36)]. These conditions lead to

Theorem 2.1 *For $n \geq 0$, there exist unique matrices $A_{n,j}$ of size $(n+1) \times (n+2)$, $B_{n,j}$ of size $(n + 1) \times (n + 1)$,and $C_{n,j}$ of size $(n + 1) \times n$, such that*

$$x_j \mathbb{P}_n = A_{n,j} \mathbb{P}_{n+1} + B_{n,j} \mathbb{P}_n + C_{n,j} \mathbb{P}_{n-1}, \quad j = 1, 2, \tag{2.4}$$

with the initial conditions $\mathbb{P}_{-1} = 0$ and $\mathbb{P}_0 = 1$. Here the notation $x_1 = x$, $x_2 = y$ is used.

As it was mentioned in the univariate situation, the next step is to obtain explicitly these matrices. If we expand

$$\mathbb{P}_n = G_{n,n}\mathbf{x}^n + G_{n,n-1}\mathbf{x}^{n-1} + \cdots + G_{n,0}\mathbf{x}^0,$$

then we have [6]

Theorem 2.2 *The explicit expressions of the matrices $A_{n,j}$, $B_{n,j}$ and $C_{n,j}$ ($j = 1, 2$) appearing in (2.4) in terms of the values of the leading coefficients $G_{n,n}$,*

$G_{n,n-1}$ and $G_{n,n-2}$ are given by

$$\begin{cases} A_{n,j} = G_{n,n} L_{n,j} G_{n+1,n+1}^{-1}, \quad n \geq 0, \\ B_{0,j} = -A_{0,j} G_{1,0}, \\ B_{n,j} = (G_{n,n-1} L_{n-1,j} - A_{n,j} G_{n+1,n}) G_{n,n}^{-1}, \quad n \geq 1, \\ C_{1,j} = -(A_{1,j} G_{2,0} + B_{1,j} G_{1,0}), \\ C_{n,j} = (G_{n,n-2} L_{n-2,j} - A_{n,j} G_{n+1,n-1} - B_{n,j} G_{n,n-1}) G_{n-1,n-1}^{-1}, \quad n \geq 2\,. \end{cases} \tag{2.5}$$

Finally, the last step is to compute the matrices $G_{n,n}$, $G_{n,n-1}$ and $G_{n,n-2}$ in terms of the coefficients of the partial differential equation. We shall show the results in the specific case of monic polynomial solutions

$$\hat{\mathbb{P}}_n = \mathbf{x}^n + \hat{G}_{n,n-1} \mathbf{x}^{n-1} + \hat{G}_{n,n-2} \mathbf{x}^{n-2} + \cdots\,. \tag{2.6}$$

In order to substitute the latter expression in the partial differential equation, some basic properties of the monomials $\mathbf{x}^n$ are needed. The multiplication by x and y can be done as

$$\begin{cases} x\, \mathbf{x}^n = L_{n,1} \mathbf{x}^{n+1}, \\ y\, \mathbf{x}^n = L_{n,2} \mathbf{x}^{n+1}, \end{cases} \tag{2.7}$$

where the matrices $L_{n,j}$ of size $(n+1) \times (n+2)$ are defined by

$$L_{n,1} = \begin{pmatrix} 1 & & \bigcirc & 0 \\ & \ddots & & \vdots \\ \bigcirc & & 1 & 0 \end{pmatrix} \quad \text{and} \quad L_{n,2} = \begin{pmatrix} 0 & 1 & & \bigcirc \\ \vdots & & \ddots & \\ 0 & \bigcirc & & 1 \end{pmatrix}. \tag{2.8}$$

Notice that for $j = 1, 2$ we have $L_{n,j}\, L_{n,j}^{\mathrm{T}} = I_{n+1}$, where I_{n+1} denotes the identity matrix of size $n+1$.

Moreover, for $n \geq 1$,

$$\begin{cases} \partial_x \mathbf{x}^n = \mathbb{E}_{n,1}\, \mathbf{x}^{n-1}, \\ \partial_y \mathbf{x}^n = \mathbb{E}_{n,2}\, \mathbf{x}^{n-1}, \end{cases} \tag{2.9}$$

where the matrices $\mathbb{E}_{n,j}$ of size $(n+1) \times n$ are given by

$$\mathbb{E}_{n,1} = \begin{pmatrix} n & & & \bigcirc \\ & n-1 & & \\ & & \ddots & \\ & \bigcirc & & 1 \\ 0 & \ldots & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbb{E}_{n,2} = \begin{pmatrix} 0 & \ldots & & 0 \\ 1 & & & \bigcirc \\ & 2 & & \\ & & \ddots & \\ \bigcirc & & & n \end{pmatrix}. \tag{2.10}$$

If we substitute the monic expansion (2.6) into the partial differential equation (2.1) then equating coefficients we obtain

$$\widehat{G}_{n,n-1} = \begin{pmatrix} \tilde{g}_{1,1} & & & & \bigcirc \\ \tilde{g}_{2,1} & \tilde{g}_{2,2} & & & \\ & \ddots & \ddots & & \\ & & \tilde{g}_{n-1,n-2} & \tilde{g}_{n-1,n-1} & \\ & & & \tilde{g}_{n,n-1} & \tilde{g}_{n,n} \\ \bigcirc & & & 0 & \tilde{g}_{n+1,n} \end{pmatrix}, \quad (n \geq 1)\,, \tag{2.11}$$

where, for $1 \leq i \leq n$,

$$\tilde{g}_{i,i} = \frac{(n+1-i)((n-i)b_1 + 2(i-1)c_3 + f_1)}{\varpi_{2n-2}}\,,$$

$$\tilde{g}_{i+1,i} = \frac{i((i-1)b_2 + 2(n-i)b_3 + f_2)}{\varpi_{2n-2}}\,,$$

and

$$\widehat{G}_{n,n-2} = \begin{pmatrix} g_{1,1} & & & \bigcirc \\ g_{2,1} & g_{2,2} & & \\ g_{3,1} & g_{3,2} & g_{3,3} & \\ \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ & g_{n-1,n-3} & g_{n-1,n-2} & g_{n-1,n-1} \\ \bigcirc & & g_{n,n-2} & g_{n,n-1} \\ & & 0 & g_{n+1,n-1} \end{pmatrix}, \quad (n \geq 2)\,, \tag{2.12}$$

where, for $1 \leq i \leq n-1$,

$$\begin{aligned} g_{i,i} &= \frac{(n-i)(n+1-i)}{2\varpi_{2n-2}\varpi_{2n-3}} \\ &\quad \times (\varpi_{2n-2}c_1 + ((n-i)b_1 + 2(i-1)c_3 + f_1)((n-i-1)b_1 + 2(i-1)c_3 + f_1)), \\ g_{i+1,i} &= \frac{i(n-i)}{\varpi_{2n-2}\varpi_{2n-3}}(f_1 f_2 + d_3\varpi_{2n-2} + b_3(2(n-2+2(i-2)(n-i-1))c_3 \\ &\quad + (2n-2i-1)f_1) + (2i-1)c_3 f_2 + (i-1)b_2((2i-1)c_3 + f_1) \\ &\quad + (n-1-i)((i-1)b_2 + (2n-2i-1)b_3 + f_2)), \end{aligned}$$

$$g_{i+2,i} = \frac{i(i+1)}{2\varpi_{2n-2}\varpi_{2n-3}} \times (\varpi_{2n-2}c_2 + ((i-1)b_2 + 2(n-i-1)b_3 + f_2)(ib_2 + 2(n-i-1)b_3 + f_2)),$$

where $\varpi_n = na + e \neq 0$—see (2.2).

As an example of the bivariate continuous case, let us consider the following potentially self-adjoint second order partial differential equation of hypergeometric type —tiny variation of [52, Chapter III]

$$\begin{aligned} x(x-1)\frac{\partial^2}{\partial x^2}f(x,y) + y(y-1)\frac{\partial^2}{\partial y^2}f(x,y) + 2xy\frac{\partial^2}{\partial x \partial y}f(x,y) \\ + ((3+\alpha+\beta+\gamma)x - 1 - \alpha)\frac{\partial}{\partial x}f(x,y) + ((3+\alpha+\beta+\gamma)y - 1 - \beta)\frac{\partial}{\partial y}f(x,y) \\ - n(n+\alpha+\beta+\gamma+2)f(x,y) = 0. \end{aligned} \tag{2.13}$$

Following the approach described in this section, in the monic case, the recursion coefficients $\widehat{B}_{n,j}$ are given by

$$\widehat{B}_{n,1} = \begin{pmatrix} b_{0,0} & 0 & & & \bigcirc \\ b_{1,0} & b_{1,1} & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-1,n-2} & b_{n-1,n-1} & 0 \\ \bigcirc & & & b_{n,n-1} & b_{n,n} \end{pmatrix}, \tag{2.14}$$

where

$$b_{i,i} = \frac{(i-n)(\alpha-i+n)}{\alpha+\beta+\gamma+2n+1} + \frac{(n-i+1)(\alpha-i+n+1)}{\alpha+\beta+\gamma+2n+3}, \quad 0 \le i \le n,$$

$$b_{i+1,i} = -\frac{2(i+1)(\beta+i+1)}{(\alpha+\beta+\gamma+2n+1)(\alpha+\beta+\gamma+2n+3)}, \quad 0 \le i \le n-1,$$

and

$$\widehat{B}_{n,2} = \begin{pmatrix} \tilde{b}_{0,0} & \tilde{b}_{0,1} & & & \bigcirc \\ 0 & \tilde{b}_{1,1} & \tilde{b}_{1,2} & & \\ & \ddots & \ddots & \ddots & \\ & & & \tilde{b}_{n-1,n-1} & \tilde{b}_{n-1,n} \\ \bigcirc & & & 0 & \tilde{b}_{n,n} \end{pmatrix}, \tag{2.15}$$

with

$$\tilde{b}_{i,i} = \frac{(i+1)(\beta+i+1)}{\alpha+\beta+\gamma+2n+3} - \frac{i(\beta+i)}{\alpha+\beta+\gamma+2n+1}, \quad 0 \leq i \leq n,$$

$$\tilde{b}_{i,i+1} = \frac{2(i-n)(\alpha-i+n)}{(\alpha+\beta+\gamma+2n+1)(\alpha+\beta+\gamma+2n+3)}, \quad 0 \leq i \leq n-1.$$

Moreover, the recursion coefficients $C_{n,i}$ are given by

$$\widehat{C}_{n,1} = \begin{pmatrix} c_{0,0} & & & & \bigcirc \\ c_{1,0} & c_{1,1} & & & \\ c_{2,0} & c_{2,1} & c_{2,2} & & \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & c_{n-1,n-3} & c_{n-1,n-2} & c_{n-1,n-1} \\ & & & c_{n,n-2} & c_{n,n-1} \end{pmatrix}, \tag{2.16}$$

where for $0 \leq i \leq n-1$ we have

$$c_{i,i} = \frac{(n-i)(\alpha-i+n)(\beta+\gamma+i+n+1)(\alpha+\beta+\gamma+i+n+1)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)},$$

$$c_{i+1,i} = -\frac{(i+1)(\beta+i+1)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)},$$
$$\times \left(\alpha^2+\alpha(\beta+\gamma+2n+1)-2(i-n)(\beta+\gamma+i+n)-\beta-\gamma-4i+2n-2\right)$$

and for $0 \leq i \leq n-2$,

$$c_{i+2,i} = \frac{(i+1)(i+2)(\beta+i+1)(\beta+i+2)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)};$$

and

$$\widehat{C}_{n,2} = \begin{pmatrix} \tilde{c}_{0,0} & \tilde{c}_{0,1} & & & \bigcirc \\ \tilde{c}_{1,0} & \tilde{c}_{1,1} & \tilde{c}_{1,2} & & \\ & \ddots & \ddots & \ddots & \\ & \tilde{c}_{n-2,n-3} & \tilde{c}_{n-2,n-2} & \tilde{c}_{n-2,n-1} & \\ & & \tilde{c}_{n-1,n-2} & \tilde{c}_{n-1,n-1} & \\ \bigcirc & & & \tilde{c}_{n,n-1} & \end{pmatrix}, \tag{2.17}$$

with

$$\tilde{c}_{i,i} = \frac{(i-n)(\alpha-i+n)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)},$$
$$\times\left(\alpha(\beta+2i+1)+\beta^2+\beta\gamma+2n(\beta+2i+1)+\beta+2\gamma i+\gamma-2i^2\right),$$
$$\tilde{c}_{i+1,i} = \frac{(i+1)(\beta+i+1)(\alpha+\gamma-i+2n)(\alpha+\beta+\gamma-i+2n)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)},$$

for $0 \le i \le n-1$ and

$$\tilde{c}_{i,i+1} = \frac{(n-i-1)(n-i)(\alpha-i+n-1)(\alpha-i+n)}{(\alpha+\beta+\gamma+2n)(\alpha+\beta+\gamma+2n+1)^2(\alpha+\beta+\gamma+2n+2)},$$

for $0 \le i \le n-2$. Moreover, the monic polynomials can be expressed as

$$\begin{aligned}\widehat{P}_{n,m}^{(\alpha,\beta,\gamma)}(x,y) &= \frac{(-1)^{m+n}(\alpha+1)_n(\beta+1)_m}{(\alpha+\beta+\gamma+m+n+2)_{m+n}}\\ &\quad\times F^{1:1;1}_{0:1;1}\left(\begin{matrix}\alpha+\beta+\gamma+m+n+2:-n;-m\\ -:\alpha+1;\beta+1\end{matrix}\,\middle|\, x,y\right),\end{aligned} \tag{2.18}$$

so that

$$\widehat{\mathbb{P}}_n = \left(\widehat{P}_{n-m,m}^{(\alpha,\beta,\gamma)}(x,y)\right)_{m=0}^{n}.$$

These polynomials are orthogonal in

$$A = \{(x,y)\in\mathbf{R}^2 \mid 0\le y\le 1-x, \quad 0\le x\le 1\}$$

with respect to $\varrho(x,y) = x^\alpha y^\beta (1-x-y)^\gamma$.

One disadvantage of this approach is that it is too restrictive and many families of orthogonal polynomials are not considered here (see [52] for a number of partial differential equations that are not of hypergeometric type or admissible or potentially self-adjoint). On the other hand, this method allows to give a general approach to this class of orthogonal polynomials and to extend it to the discrete and q-analogues. As a byproduct, we can easily obtain the relations between "different" solutions of the same partial differential (difference or q-difference equation). In order to show this last property, we shall consider the following second-order admissible potentially self-adjoint partial difference equation of hypergeometric type [7]

$$\begin{aligned}&(p_1-1)x\Delta_1\nabla_1 u(x,y)+p_1 y\Delta_1\nabla_2 u(x,y)+p_2 x\Delta_2\nabla_1 u(x,y)+(p_2-1)y\Delta_2\nabla_2 u(x,y)\\ &+(x-Np_1)\Delta_1 u(x,y)+(y-Np_2)\Delta_2 u(x,y)-(n_1+n_2)u(x,y)=0,\end{aligned} \tag{2.19}$$

where the forward and backward difference operators are defined by

$$\Delta_1 u(x,y) = u(x+1,y) - u(x,y), \quad \Delta_2 u(x,y) = u(x,y+1) - u(x,y),$$
$$\nabla_1 u(x,y) = u(x,y) - u(x-1,y), \quad \nabla_2 u(x,y) = u(x,y) - u(x,y-1).$$

There exist at least four polynomial solutions of (2.19):

1. The monic bivariate Kravchuk polynomials [55, 57], defined as a generalized Kampé de Fériet series [51] by means of

$$\hat{K}^{p_1,p_2}_{n_1,n_2}(x,y;N) = (-1)^{n_1+n_2} p_1^{n_1} p_2^{n_2} (N-n_1-n_2+1)_{n_1+n_2} \times F^{0:2;2}_{1:0;0}\left(\begin{matrix} - : -n_1, -x; -n_2, -y \\ -N : -; - \end{matrix} \,\middle|\, \frac{1}{p_1}, \frac{1}{p_2} \right); \tag{2.20}$$

2. the non-monic bivariate Kravchuk polynomials [55, 57], defined also as a generalized Kampé de Fériet series,

$$K^{p_1,p_2}_{n_1,n_2}(x,y;N) = (x+y-N)_{n_1+n_2} \times F^{0:2;2}_{1:0;0}\left(\begin{matrix} - : -n_1, -x; -n_2, -y \\ -n_1-n_2-x-y+N+1 : -; - \end{matrix} \,\middle|\, \frac{p_1+p_2-1}{p_1}, \frac{p_1+p_2-1}{p_2} \right); \tag{2.21}$$

   These polynomials are exactly the same obtained from the Rodrigues-type formula [47, Eq. (60)].

3. the non-monic bivariate Kravchuk polynomials [46] defined as a product of univariate Kravchuk polynomials

$$\tilde{K}^{p_1,p_2}_{n_1,n_2}(x,y;N) = \frac{(N-n_1)!}{N!\,(n_1-N)_{n_2}} \times K_{n_1}(x;p_1/(p_1+p_2), x+y)\, K_{n_2}(x+y-n_1; p_1+p_2, N-n_1)\,, \tag{2.22}$$

   where for $0 < p < 1$ and $n = 0, 1, \ldots, N$, the univariate Kravchuk polynomials are normalized as [44]

$$K_n(x;p,N) = (-N)_n\, {}_2F_1\left(\begin{matrix} -n, -x \\ -N \end{matrix} \,\middle|\, \frac{1}{p} \right); \tag{2.23}$$

4. and the non-monic bivariate Kravchuk polynomials [17, Eq. (5.19)] defined also as a product of univariate Kravchuk polynomials

$$K_2(n_1, n_2; x, y; p_1, p_2; N) = \frac{1}{(-N)_{n_1+n_2}} K_{n_1}(x; p_1, N - n_2)\, K_{n_2}(y; p_2/(1 - p_1), N - x). \qquad (2.24)$$

The above polynomials have been analyzed in [39, Eq. (44)] for specific values of p_1 and p_2 as eigenfunctions of a certain isotropic Hamiltonian.

From the hypergeometric approach [7] applied to (2.19) we obtain that these four families of polynomials are orthogonal with respect to the trinomial distribution [24]

$$\varrho^{N,p_1,p_2}(x, y) = \frac{N!}{x!\, y!\, (N - x - y)!} p_1^x\, p_2^y\, (1 - p_1 - p_2)^{N-x-y}, \qquad (2.25)$$

in the triangular domain G defined by $x \geq 0$, $y \geq 0$, and $0 \leq x + y \leq N$, where N is a positive integer and p_1 and p_2 are real numbers satisfying $p_1 > 0$, $p_2 > 0$, $0 < p_1 + p_2 < 1$.

Let $\hat{\mathbb{P}}_n$ be the column vector of the monic bivariate Kravchuk polynomials (2.20). Let $\mathbb{P}_n^i$ and $\mathbb{P}_n^j$ be any two other families. Then, we have

$$\mathbb{P}_n^i = G_{n,n}^i\, \hat{\mathbb{P}}_n, \qquad \mathbb{P}_n^j = G_{n,n}^j\, \hat{\mathbb{P}}_n.$$

As a consequence, we obtain the following formula relating the two families of bivariate Kravchuk polynomials

$$\mathbb{P}_n^i = G_{n,n}^i (G_{n,n}^j)^{-1}\, \mathbb{P}_n^j,\ n \geq 0.$$

The latter expression can also be written as

$$P_{n-j,j}(x, y) = \sum_{l=0}^{n} b_{j,l}\, \tilde{P}_{n-l,l}(x, y), \quad j = 0, 1, \ldots, n,$$

where $P_{n-j,j}(x, y)$ and $\tilde{P}_{n-l,l}(x, y)$ stand for the elements of the ith and jth family respectively, and $(b_{j,0}, b_{j,1}, \ldots, b_{j,n})$ is the jth row of the matrix $G_{n,n}^i (G_{n,n}^j)^{-1}$.

For instance, for the fourth family we have

$$G_{n,n}^4 = \left((-1)^n \binom{r}{s} \frac{(1 - p_1)^s\, p_2^{-s}\, p_1^{r-n}}{(N - n + 1)_n} \right)_{r,s=0}^{n}.$$

Therefore,

$$(G^4_{n,n})^{-1} = \left(\binom{r}{s}(-1)^{n-r-s}(1-p_1)^{-r}p_2^r(N-n+1)_n p_1^{n-s}\right)_{r,s=0}^{n}.$$

As an example, for $n = 3$, the link between the third family (2.22) and the fourth family (2.24) is given by

$$\mathbb{P}_3^3 = \begin{pmatrix} p_2^3K^3 & -3p_2^3K^3 & 3p_2^3K^3 & -p_2^3K^3 \\ p_1p_2^2\Lambda K^2 & -p_2^2K^2\,(p_2+2p_1\Lambda) & p_2^2K^2\,(2p_2+p_1\Lambda) & -p_2^3K^2 \\ p_1^2p_2\Lambda^2K & -p_1p_2\Lambda K\,(2p_2+p_1\Lambda) & p_2^2K\,(p_2+2p_1\Lambda) & -p_2^3K \\ p_1^3\Lambda^3 & -3p_1^2p_2\Lambda^2 & 3p_1p_2^2\Lambda & -p_2^3 \end{pmatrix}\mathbb{P}_3^4,$$

where $\Lambda = p_1 + p_2 - 1$ and $K = p_1 + p_2$.

Limit relations from bivariate Kravchuk polynomials to bivariate Hermite polynomials have been studied in [5], as well as some limit relations between bivariate Hahn to bivariate Appell polynomials.

2.2 Bivariate Orthogonal Polynomials on Nonuniform Lattices

As we have seen in the univariate case all "classical" orthogonal polynomials are solution of a second-order differential (difference, q-difference, or divided-difference) equation of hypergeometric type. In the multivariate situation we have analyzed admissible potentially self-adjoint second-order partial differential (difference or q-difference) equations of hypergeometric type having orthogonal polynomial solutions. But, if we try to analyze the equation in nonuniform lattices this is no longer true.

Let us consider the bivariate Racah polynomials considered introduced by Tratnik in [57] and deeply analyzed by Geronimo and Iliev in [17], where they construct a commutative algebra $\mathcal{A}_x$ of difference operators in $\mathbf{R}^2$, depending on 5 parameters, which is diagonalized by the multivariable Racah polynomials considered by Tratnik. The bivariate Racah polynomials are defined in terms of univariate Racah polynomials (1.8) as

$$\begin{aligned} R_{n,m}(s,t;\beta_0,\beta_1,\beta_2,\beta_3,N) &= r_n(\beta_1-\beta_0-1,\beta_2-\beta_1-1,-t-1,\beta_1+t;s) \\ \times\, r_m(2n+\beta_2-\beta_0-1,&\,\beta_3-\beta_2-1,n-N-1,n+\beta_2+N;t-n), \end{aligned} \tag{2.26}$$

which are polynomials in the lattices $x(s) = s(s+\beta_1)$ and $y(t) = t(t+\beta_2)$. These polynomials coincide with the bivariate Racah polynomials of parameters a_1, a_2, a_3, γ, and η introduced by Tratnik [57, Eq. (2.1)] after the substitutions

$$\beta_0 = a_1-\eta-1,\quad \beta_1 = a_1,\quad \beta_2 = a_1+a_2,\quad \beta_3 = a_1+a_2+a_3,\ \text{ and }\ N = -\gamma-1. \tag{2.27}$$

The equation for bivariate Racah polynomials given in [18] has 9 rational coefficients, compared to 6 polynomial coefficients for the operator corresponding to bivariate big q-Jacobi polynomials. After some manipulations the equation obtained by Geronimo and Iliev can be expressed as [53]

$$
\begin{aligned}
&f_1(x(s),y(t))\mathbb{D}_x^2\mathbb{D}_y^2R_{n,m}(s,t)+f_2(x(s),y(t))\mathbb{S}_x\mathbb{D}_x\mathbb{D}_y^2R_{n,m}(s,t)\\
&\quad+f_3(x(s),y(t))\mathbb{S}_y\mathbb{D}_y\mathbb{D}_x^2R_{n,m}(s,t)+f_4(x(s),y(t))\mathbb{S}_x\mathbb{D}_x\mathbb{S}_y\mathbb{D}_yR_{n,m}(s,t)\\
&\quad+f_5(x(s))\mathbb{D}_x^2R_{n,m}(s,t)+f_6(y(t))\mathbb{D}_y^2R_{n,m}(s,t)\\
&\quad+f_7(x(s))\mathbb{S}_x\mathbb{D}_xR_{n,m}(s,t)+f_8(y(t))\mathbb{S}_y\mathbb{D}_yR_{n,m}(s,t)\\
&\quad+(m+n)(\beta_3-\beta_0+m+n-1)R_{n,m}(s,t)=0,
\end{aligned}
\tag{2.28}
$$

where $R_{n,m}(s,t) := R_{n,m}(s,t;\beta_0,\beta_1,\beta_2,\beta_3,N)$, and the coefficients f_i, $i = 1,\dots,8$ are polynomials in the lattices $x(s)$ and $y(t)$ given by

$$
\begin{aligned}
f_8(y(t))&=(\beta_0-\beta_3)y(t)-N(\beta_0-\beta_2)(\beta_3+N),\\
f_7(x(s))&=(\beta_0-\beta_3)x(s)-N(\beta_0-\beta_1)(\beta_3+N),\\
f_6(y(t))&=-(y(t))^2+\frac{1}{2}(2N^2+2\beta_3(\beta_0+N)-\beta_2(\beta_3+\beta_0))y(t)\\
&\quad-\frac{1}{2}N\beta_2(\beta_0-\beta_2)(\beta_3+N),\\
f_5(x(s))&=-(x(s))^2+\frac{1}{2}(2\beta_3(N+\beta_0)+2N^2-\beta_1(\beta_3+\beta_0))x(s)\\
&\quad-\frac{1}{2}N\beta_1(\beta_0-\beta_1)(\beta_3+N),\\
f_4(x(s),y(t))&=-2x(s)y(t)+(2N^2+\beta_2(1-\beta_0)+\beta_3(\beta_0-1+2N))x(s)\\
&\quad+(\beta_0-\beta_1)(\beta_3+1)y(t)-N(\beta_0-\beta_1)(\beta_2+1)(\beta_3+N),\\
f_3(x(s),y(t))&=(\beta_2-\beta_3)(x(s))^2\\
&\quad+x(s)\Big(-(1+\beta_1+\beta_3-2\beta_0)y(t)+(1+\beta_1-2\,\beta_0+\beta_2)\,N^2\\
&\quad-\beta_3\,(-\beta_1-\beta_2-1+2\,\beta_0)\,N+\frac{1}{2}\,(\beta_2-\beta_3)\,(\beta_1\beta_0-2\,\beta_0+\beta_1)\Big)\\
&\quad+\frac{1}{2}\,\beta_1\,(\beta_3+1)\,(\beta_0-\beta_1)\,y(t)-\frac{1}{2}\,\beta_1N\,(\beta_2+1)\,(\beta_3+N)\,(\beta_0-\beta_1)\,,\\
f_2(x(s),y(t))&=(\beta_0-\beta_1)(y(t))^2\\
&\quad+x(s)\Big((\beta_0+\beta_2-2\beta_3-1)y(t)+(1-\beta_0+\beta_2)\,N^2
\end{aligned}
$$

$$
\begin{aligned}
&- \beta_3\,(-1+\beta_0-\beta_2)\,N + \frac{1}{2}\,\beta_2\,(\beta_2-\beta_3)\,(\beta_0-1)\Big)\\
&- \frac{1}{2}\,(\beta_0-\beta_1)\Big(2\,\beta_3 N - \beta_3\beta_2 + 2\,N^2 - 2\,\beta_3 + \beta_2\Big)\,y(t)\\
&- \frac{1}{2}\,(\beta_0-\beta_1)\,N\beta_2\,(\beta_2+1)\,(\beta_3+N)\,,
\end{aligned}
$$

$$
\begin{aligned}
f_1(x(s),y(t)) = {}& -(x(s))^2 y(t) - x(s)(y(t))^2 + (N^2+\beta_3 N - \frac{1}{2}\,\beta_2\,(\beta_2-\beta_3))(x(s))^2\\
&+ \frac{1}{2}\,\beta_1\,(\beta_0-\beta_1)\,(y(t))^2 + \Big(\Big(\frac{1}{2}\,\beta_1+\frac{1}{2}-\frac{1}{2}\,\beta_3-\beta_0\Big)\,\beta_2-\beta_3\\
&- \frac{1}{2}\,\beta_1 + 2\,\beta_0\beta_3 + \beta_0 + N^2 - \beta_1\beta_3 + \beta_3 N - \frac{1}{2}\,\beta_1\beta_0\Big)x(s)y(t)\\
&+ \Big(\Big(\frac{1}{2}\,\beta_1\beta_0 + \frac{1}{2}\,{\beta_2}^2 + \frac{1}{2}\,\beta_1\beta_2 + \frac{1}{2}\,\beta_2 - \beta_0\beta_2 + \frac{1}{2}\,\beta_1 - \beta_0\Big)\,N^2\\
&+ \frac{1}{2}\,\beta_3\,\Big(\beta_1\beta_0 + {\beta_2}^2 + \beta_1\beta_2 + \beta_2 - 2\,\beta_0\beta_2 + \beta_1 - 2\,\beta_0\Big)\,N\\
&- \frac{1}{4}\,\beta_2\,(\beta_2-\beta_3)\,(\beta_1\beta_0+\beta_1-2\,\beta_0))x(s)\\
&- \frac{1}{4}\,\beta_1\,\Big(\beta_2 - 2\,\beta_3 + 2\,\beta_3 N + 2\,N^2 - \beta_2\beta_3\Big)\,(\beta_0-\beta_1)\,y(t)\\
&- \frac{1}{4}\,N\beta_1\beta_2\,(\beta_2+1)\,(\beta_0-\beta_1)\,(\beta_3+N)\,.
\end{aligned}
$$

From the above presentation of the equation obtained by Geronimo and Iliev it is possible to prove that the divided-difference(s) of bivariate Racah polynomials are solution of an equation of the same type and hence [53]

$$
\begin{aligned}
\mathbb{D}_x R_{n,m}(s,t;\beta_0,\beta_1,\beta_2,\beta_3,N) &= n(n-\beta_0+\beta_2-1)\\
&\quad\times R_{n-1,m}(s-1/2,t-1;\beta_0,\beta_1+1,\beta_2+2,\beta_3+2,N-1).
\end{aligned}
$$

Similarly, if we consider the following second family of the bivariate Racah polynomials obtained from [57, Equation (2.12)] using the transformations (2.27)

$$
\begin{aligned}
&\bar{R}_{n,m}(s,t;\beta_0,\beta_1,\beta_2,\beta_3,N)\\
&= r_n(2m-\beta_1+\beta_3-1,\beta_1-\beta_0-1,m-N-1,m-N-\beta_1,N-m-s)\\
&\quad\times r_m(\beta_3-\beta_2-1,\beta_2-\beta_1-1,s-N-1,-\beta_2-N-s,N-t), \qquad (2.29)
\end{aligned}
$$

it follows that

$$\begin{aligned}&\mathbb{D}_y \bar{R}_{n,m}(s, t; \beta_0, \beta_1, \beta_2, \beta_3, N)\\&\quad = m(m + \beta_3 - \beta_1 - 1)\bar{R}_{n,m-1}(s, t - 1/2; \beta_0, \beta_1, \beta_2 + 1, \beta_3 + 2, N - 1).\end{aligned}$$

As it happens in the bivariate Kravchuk case, both families of bivariate Racah polynomials $R_{n,m}(s, t; \beta_0, \beta_1, \beta_2, \beta_3, N)$ and $\bar{R}_{n,m}(s, t; \beta_0, \beta_1, \beta_2, \beta_3, N)$ are solution of the same divided-difference equation (2.28).

Another advantage of presenting the equation for bivariate Racah polynomials in the form (2.28) is that it is possible to compute explicitly the matrices of the three-term recurrence relation satisfied by a "monic" family of bivariate Racah polynomials, as they has been obtained in [54].

A similar approach has been followed for bivariate q-Racah and bivariate Askey–Wilson polynomials. For these families of bivariate orthogonal polynomials the difference equation(s) were also obtained by Geronimo and Iliev [18] and later presented in a different form in [54] which have allowed to obtain some properties of the bivariate orthogonal polynomials, as well as to explicitly compute the matrices of the three-term recurrence relation satisfied by a family of monic polynomials.

In the bivariate situation, it is possible to consider appropriate limit relations (or choice of the parameters) between the families of orthogonal polynomials. Some of them have been already mentioned as the limit from bivariate Kravchuk to bivariate Hermite polynomials. This is also valid if we consider orthogonal polynomials on nonuniform lattices. For instance, if we consider the change of variables [18, p. 443]

$$\left\{\begin{array}{l}\beta_0 = a - b, \quad \beta_1 = 2a, \quad \beta_2 = 2a + 2e_2, \quad \beta_3 = 2a + 2e_2 + c + d,\\ s = -a + ix, \quad t = -a - e_2 + iy, \quad N = -a - d - e_2,\end{array}\right. \tag{2.30}$$

the bivariate Racah polynomials (2.26) become the bivariate Wilson polynomials (similarly to the univariate case [27, p. 196])

$$\begin{aligned}&W_{n,m}(x, y; a, b, c, d; e_2)\\&= w_n(x^2; a, b, e_2 + iy, e_2 - iy)w_m(y^2; n + a + e_2, n + b + e_2, c, d),\end{aligned} \tag{2.31}$$

where $w_n(x^2; a, b, c, d)$ are the Wilson polynomials defined by (1.11). As indicated in [53], the change of variable (2.30) transforms the fourth-order linear partial divided-difference equation satisfied by bivariate Racah polynomials (2.28) into another fourth-order linear partial divided-difference equation satisfied by the bivariate Wilson polynomials (2.31):

$$\begin{aligned}&f_1(x, y)\mathbf{D}_x^2\mathbf{D}_y^2 W_{n,m}(x, y) + f_2(x, y)\mathbf{S}_x\mathbf{D}_x\mathbf{D}_y^2 W_{n,m}(x, y)\\&\qquad + f_3(x, y)\mathbf{S}_y\mathbf{D}_y\mathbf{D}_x^2 W_{n,m}(x, y) + f_4(x, y)\mathbf{S}_x\mathbf{D}_x\mathbf{S}_y\mathbf{D}_y W_{n,m}(x, y)\end{aligned}$$

$$+ f_5(x)\mathbf{D}_x^2 W_{n,m}(x,y) + f_6(y)\mathbf{D}_y^2 W_{n,m}(x,y)$$
$$+ f_7(x)\mathbf{S}_x\mathbf{D}_x W_{n,m}(x,y) + f_8(y)\mathbf{S}_y\mathbf{D}_y W_{n,m}(x,y)$$
$$+ (m+n)(2e_2+a+b+c+d+m+n-1)W_{n,m}(x,y) = 0, \qquad (2.32)$$

where

$$\begin{aligned}
f_8(y) &= (-a-b-2e_2-c-d)\,y^2 + (c+d)\,e_2^2 + (ad+ca+db+bc+2\,dc)\,e_2 \\
&\quad + adc + dba + bac + dbc, \\
f_7(x) &= (-a-b-2\,e_2-c-d)\,x^2 + (a+b)\,e_2^2 + (bc+db+ad+2\,ba+ca)\,e_2 \\
&\quad + bac + dbc + adc + dba, \\
f_6(y) &= -y^4 + \Big(be_2 + ba + 2\,ce_2 + ca + ae_2 + e_2^2 + bc + dc + db + 2e_2 d + ad\Big)\,y^2 \\
&\quad - dc\,(e_2+b)\,(e_2+a)\,, \\
f_5(x) &= -x^4 + \Big(e_2^2 + 2\,ae_2 + e_2\,d + ad + 2\,be_2 + ba + bc + ce_2 + dc + db + ca\Big)\,x^2 \\
&\quad - ba\,(e_2+d)\,(e_2+c)\,, \\
f_4(x,y) &= -2x^2y^2 + (d + c + 2\,ce_2 + ca + bc + 2\,dc + db + 2\,e_2\,d + ad)\,x^2 \\
&\quad + \Big(2\,ae_2 + ca + ad + a + 2\,be_2 + 2\,ba + bc + db + b\Big)y^2 \\
&\quad - (c+d)\,(a+b)\,e_2^2 + (-2\,dba - 2\,adc - ad - 2\,dbc - ca - db \\
&\quad - bc - 2\,bac)\,e_2 - 2\,dbac - adc - dba - bac - dbc, \\
f_3(x,y) &= (c+d)x^4 - ba\,(2\,e_2 + d + c + 1)\,y^2 + (1 + 2\,a + 2\,e_2 + c + d + 2\,b)x^2y^2 \\
&\quad + ba\Big(\,(c+d)\,e_2^2 + (d + c + 2\,dc)\,e_2 + dc\Big) \\
&\quad + \Big(\,(-c-d)\,e_2^2 + (-2\,ad - 2\,bc - 2\,ca - 2\,db - c - d - 2\,dc)\,e_2 \\
&\quad - db - ca - bc - 2\,adc - dc - dba - bac - ad - 2\,dbc\Big)x^2, \\
f_2(x,y) &= (a+b)y^4 - dc\,(1 + a + b + 2\,e_2)\,x^2 + (a + b + 2\,e_2 + 2\,c + 2\,d + 1)x^2y^2 \\
&\quad + dc\Big(\,(a+b)\,e_2^2 + (2\,ba + a + b)\,e_2 + ba\Big) \\
&\quad + \Big(\,(-a-b)\,e_2^2 + (-a - 2\,ba - 2\,ad - 2\,bc - b - 2\,ca - 2\,db)\,e_2 \\
&\quad - dbc - ba - ca - ad - 2\,dba - bc - adc - 2\,bac - db\Big)y^2,
\end{aligned}$$

$$
\begin{aligned}
f_1(x,y) &= x^4y^2 + x^2y^4 - cdx^4 - aby^4 + \Big((-2\,c - 2\,b - 2\,d - 1 - 2\,a)\,e_2 \\
&\quad - e_2^2 - a - b - d - c - dc - ba - 2\,ca - 2\,db - 2\,bc - 2\,ad \Big) x^2y^2 \\
&\quad + dc\left(e_2^2 + (2\,b + 2\,a + 1)\,e_2 + b + ba + a\right)x^2 \\
&\quad + ba\left(e_2^2 + (2\,c + 2\,d + 1)\,e_2 + c + d + dc\right)y^2 - adbe_2\,c\,(1 + e_2)\,.
\end{aligned}
$$

Moreover, if we divide (2.31) by b^{n+m} and take the limit as $b \to \infty$ we obtain (after redefining $c \to b$ and $d \to c$) the bivariate continuous dual Hahn polynomials [56]

$$
D_{n,m}(a,b,c,e_2;x,y) = d_n(a, e_2{+}ix, e_2{-}iy|x^2)d_m(n{+}a{+}e_2, b, c|y^2), \tag{2.33}
$$

where the univariate continuous dual Hahn polynomials are defined in (1.14). As in the univariate case for the second-order divided-difference equation for univariate continuous dual Hahn polynomials, the fourth-order divided-difference equation for bivariate continuous dual Hahn polynomials can be obtained by taking the limit as $b \to \infty$ from the fourth-order divided-difference equation for bivariate Wilson polynomials—and redefining $c \to b$ and $d \to c$—namely

$$
\begin{aligned}
& f_1(x,y)\mathbf{D}_x^2\mathbf{D}_y^2 D_{n,m}(x,y) + f_2(x,y)\mathbf{S}_x\mathbf{D}_x\mathbf{D}_y^2 D_{n,m}(x,y) \\
&\quad + f_3(x,y)\mathbf{S}_y\mathbf{D}_y\mathbf{D}_x^2 D_{n,m}(x,y) + f_4(x,y)\mathbf{S}_x\mathbf{D}_x\mathbf{S}_y\mathbf{D}_y D_{n,m}(x,y) \\
&\quad + f_5(x)\mathbf{D}_x^2 D_{n,m}(x,y) + f_6(y)\mathbf{D}_y^2 D_{n,m}(x,y) \\
&\quad + f_7(x)\mathbf{S}_x\mathbf{D}_x D_{n,m}(x,y) + f_8(y)\mathbf{S}_y\mathbf{D}_y D_{n,m}(x,y) + (m+n)D_{n,m}(x,y) = 0,
\end{aligned} \tag{2.34}
$$

where

$$
\begin{aligned}
f_8(y) &= -y^2 + (c + b)\,e_2 + cb + ab + ca, \\
f_7(x) &= -x^2 + e_2^2 + (c + 2\,a + b)\,e_2 + ca + cb + ab, \\
f_6(y) &= -cb\,(a + e_2) + (c + b + e_2 + a)\,y^2, \\
f_5(x) &= -a\,(e_2 + c)\,(e_2 + b) + (2\,e_2 + a + b + c)\,x^2, \\
f_4(x,y) &= (-b - c)\,{e_2}^2 + (-2\,ca - 2\,ab - b - 2\,cb - c)\,e_2 - ca - 2\,bac \\
&\quad - cb - ab + (1 + 2\,e_2 + 2\,a + b + c)\,y^2 + (c + b)\,x^2, \\
f_3(x,y) &= a\left(e_2\,b + 2\,be_2\,c + ce_2 + cb + c{e_2}^2 + b{e_2}^2\right) + 2\,x^2y^2 \\
&\quad - a\,(2\,e_2 + c + 1 + b)\,y^2 + (-b - ab - 2\,e_2\,b - 2\,cb - c - 2\,ce_2 - ca)\,x^2,
\end{aligned}
$$

$$f_2(x,y) = cb\left(2\,ae_2 + a + e_2 + {e_2}^2\right) + x^2y^2 - x^2cb + y^4 + \left(-e_2 - a - b - {e_2}^2 - 2\,e_2\,b - 2\,ae_2 - 2\,ca - c - 2\,ce_2 - cb - 2\,ab\right)y^2,$$

$$f_1(x,y) = -be_2\,ca\,(1+e_2) + (-1 - 2\,c - 2\,e_2 - 2\,b - a)\,x^2y^2 + cb\,(1 + 2\,e_2 + a)\,x^2 - ay^4 + a\left(cb + e_2 + b + 2\,e_2\,b + {e_2}^2 + 2\,ce_2 + c\right)y^2.$$

Hence, as mentioned at the beginning of this section, we suggest:

Definition 2.3 A multivariate orthogonal polynomial sequence is classical if it is a special case or a limiting case of the multivariate Racah polynomials or the multivariate Askey–Wilson polynomials.

Finally, the following questions arise in a natural way. How to:

1. State a general fourth-order divided-difference equation of hypergeometric type and study its polynomial solutions, both for quadratic and q-quadratic lattices?
2. Study if the fourth-order divided-difference equation is admissible?
3. Study if the fourth-order divided-difference equation is potentially self-adjoint?
4. Obtain a Pearson-type system from the fourth-order divided difference equation, as it has been done for second-order partial differential (difference, or q-difference) equations of hypergeometric type?
5. Obtain a Rodrigues-type representation for orthogonal polynomial solutions of the fourth-order divided-difference equation?
6. Obtain a hypergeometric representation of monic bivariate Racah and monic bivariate Askey–Wilson polynomials? It might be interesting to notice that e.g. both (2.20) and (2.22) are orthogonal polynomial solutions of (2.19); so, we have bivariate Kravchuk polynomials in terms of one hypergeometric series (2.20) and also as product of two univariate hypergeometric functions (2.22). The question is related with giving another orthogonal polynomial solution to (2.28) in terms of one hypergeometric series.

The above open problems are now under consideration and will be submitted elsewhere.

Acknowledgements The authors thanks the reviewers for their careful reading of the manuscript. The author also wishes to thank Prof. Amílcar Branquinho for his valuable comments which improved a preliminary version of this manuscript.

The author thanks the financial support from the Agencia Estatal de Investigación (AEI) of Spain under grant MTM2016–75140–P, co-financed by the European Community fund FEDER.

References

1. W. Al-Salam, Characterization theorems for orthogonal polynomials, in *Orthogonal Polynomials: Theory and Practice*, ed. by P. Nevai (Kluwer Academic, Dordrecht, 1990), pp. 1–24

2. G.E. Andrews, R. Askey, Classical orthogonal polynomials, in *Polynômes Orthogonaux et Applications*, ed. by C. Brezinski et al. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 36–62
3. P. Appell, J. Kampé de Fériet, *Fonctions Hypergéométriques et Hypersphériques. Polynômes d'Hermite* (Gauthier-Villars, Paris, 1926)
4. I. Area, Polinomios ortogonales de variable discreta: pares coherentes. Problemas de conexión. Ph.D. Thesis, Universidade de Vigo (1999)
5. I. Area, E. Godoy, On limit relations between some families of bivariate hypergeometric orthogonal polynomials. J. Phys. A Math. Theor. **46**, 035202 (2013)
6. I. Area, E. Godoy, A. Ronveaux, A. Zarzo, Bivariate second-order linear partial differential equations and orthogonal polynomial solutions. J. Math. Anal. Appl. **387**, 1188–1208 (2012)
7. I. Area, E. Godoy, J. Rodal, On a class of bivariate second-order linear partial difference equations and their monic orthogonal polynomial solutions. J. Math. Anal. Appl. **389**, 165–178 (2012)
8. I. Area, N.M. Atakishiyev, E. Godoy, J. Rodal, Linear partial q-difference equations on q-linear lattices and their bivariate q-orthogonal polynomial solutions. Appl. Math. Comput. **223**, 520–536 (2013)
9. N.M. Atakishiyev, M. Rahman, S.K. Suslov, On classical orthogonal polynomials. Constr. Approx. **11**, 181–223 (1995)
10. S. Bochner, Über Sturm-Liouvillesche Polynomsysteme. Math. Z. **29**, 730–736 (1929)
11. C.F. Dunkl, Y. Xu, *Orthogonal Polynomials of Several Variables.* Encyclopedia of Mathematics and Its Applications, vol. 81 (Cambridge University Press, Cambridge, 2001)
12. G.K. Engelis, On some two–dimensional analogues of the classical orthogonal polynomials (in Russian). Latviĭskiĭ Matematičeskiĭ Ežegodnik **15**, 169–202 (1974)
13. C. Ferreira, J.L. López, P.J. Pagola, Asymptotic approximations between the Hahn-type polynomials and Hermite, Laguerre and Charlier polynomials. Acta Appl. Math. **103**, 235–252 (2008)
14. M. Foupouagnigni, On difference equations for orthogonal polynomials on nonuniform lattices. J. Differ. Equ. Appl. **14**, 127–174 (2008)
15. M. Foupouagnigni, M. Kenfack Nangho, S. Mboutngam, Characterization theorem for classical orthogonal polynomials on non-uniform lattices: the functional approach. Integral Transform. Spec. Funct. **22**, 739–758 (2011)
16. M. Foupouagnigni, W. Koepf, M. Kenfack-Nangho, S. Mboutngam, On solutions of holonomic divided-difference equations on nonuniform lattices. Axioms **3**, 404–434 (2014)
17. J. Geronimo, P. Iliev, Bispectrality of multivariable Racah-Wilson polynomials. Constr. Approx. **3**1, 417–457 (2010)
18. J.S. Geronimo, P. Iliev, Multivariable Askey-Wilson function and bispectrality. Ramanujan J. **24**, 273–287 (2011)
19. E. Godoy, A. Ronveaux, A. Zarzo, I. Area, On the limit relations between classical continuous and discrete orthogonal polynomials. J. Comput. Appl. Math. **91**, 97–105 (1998)
20. R.C. Griffiths, Orthogonal polynomials on the multinomial distribution. Austral. J. Stat. **13**, 27–35 (1971). Corrigenda (1972) Austral. J. Stat. 14, 270
21. W. Hahn, Über Orthogonalpolynome, die q-Differenzengleichungen genügen. Math. Nachr. **2**, 4–34 (1949)
22. Ch. Hermite, Extrait d'une lettre *) de M. Hermite à M. Borchardt. J. Crelle **64**, 294–296 (1865). As indicated in the publication *)="Une exposition plus détaillée du sujet traité dans cette lettre se trouve dans les Comptes Rendus de l'Académie des Sciences de Paris, année 1865, séances du 20 et 27 février, du 6 et 16 mars". https://babel.hathitrust.org/cgi/pt?id=mdp.39015036985383
23. N.L. Johnson, S. Kotz, A.W. Kemp, *Univariate Discrete Distributions* (Wiley, New York, 1992)
24. N.L. Johnson, S. Kotz, N. Balakrishnan, *Discrete Multivariate Distributions* (Wiley, New York, 1997)

25. S. Karlin, J. McGregor, On some stochastic models in genetics, in *Stochastic Models in Medicine and Biology*, ed. by J. Gurland (University of Wisconsin Press, Madison, 1964), pp. 245–271
26. S. Karlin, J. McGregor, Linear growth models with many types and multidimensional Hahn polynomials, in *Theory and Application of Special Functions*. Proceedings of an Advanced Seminar, The University of Wisconsin, Madison (1975), pp. 261–288
27. R. Koekoek, P.A. Lesky, R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010)
28. T. Koornwinder, Two–variable analogues of the classical orthogonal polynomials, in *Theory and Application of Special Functions*, ed. by R. Askey. Proceedings of an Advanced Seminar, The University of Wisconsin–Madison (Academic, Cambridge, 1975), pp. 435–495
29. T.H. Koornwinder, The Askey scheme as a four-manifold with corners. Ramanujan J. **20**, 409–439 (2009)
30. M.A. Kowalski, The recursion formulas for orthogonal polynomials in n variables. SIAM J. Math. Anal. **13**, 309–315 (1982)
31. M.A. Kowalski, Orthogonality and recursion formulas for polynomials in n variables. SIAM J. Math. Anal. **13**, 316–323 (1982)
32. H.L. Krall, I.M. Sheffer, Orthogonal polynomials in two variables. Ann. Mat. Pura Appl. **4**, 325–376 (1967)
33. P. Lesky, Zweigliedrige Rekursionen für die Koeffizienten von Polynomlösungen Sturm-Liouvillescher q-Differenzengleichungen. Z. Angew. Math. Mech. **74**, 497–500 (1994)
34. A.S. Lyskova, Orthogonal polynomials in several variables. Sov. Math. Dokl. **43**, 264–268 (1991)
35. A.S. Lyskova, On some properties of orthogonal polynomials in several variables. Russ. Math. Surv. **52**, 840–841 (1997)
36. A.P. Magnus, Associated Askey-Wilson polynomials as Laguerre-Hahn orthogonal polynomials, in *Orthogonal Polynomials and Their Applications (Segovia, 1986)*, ed. by M. Alfaro et al., Lecture Notes in Mathematics, vol. 1329 (Springer, Berlin, 1988), pp. 261–278
37. A.P. Magnus, Special nonuniform lattice (SNUL) orthogonal polynomials on discrete dense sets of points. J. Comput. Appl. Math. **65**, 253–265 (1995)
38. F. Marcellán, A. Branquinho, J. Petronilho, Classical orthogonal polynomials: a functional approach. Acta Appl. Math. **34**(3), 283–303 (1994)
39. H. Miki, S. Post, L. Vinet, A. Zhedanov, A superintegrable finite oscillator in two dimensions with SU(2) symmetry. J. Phys. A: Math. Gen. **46**, 1–13 (2013)
40. G. Munschy, Résolution de l'équation de Schrödinger des atomes à deux électrons III. J. Phys. Radium **8**, 552–558 (1957)
41. G. Munschy, P. Pluvininage, Résolution de l'équation de Schrödinger des atomes à deux électrons II. J. Phys. Radium **8**, 157–160 (1957)
42. M. Njinkeu Sandjon, A. Branquinho, M. Foupouagnigni, I. Area, Characterizations of classical orthogonal polynomials on quadratic lattices. J. Differ. Equ. Appl. **23**, 983–1002 (2017)
43. P. Njionou Sadjang, W. Koepf, M. Foupouagnigni, On moments of classical orthogonal polynomials. J. Math. Anal. Appl. **424**, 122–151 (2015)
44. A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable* (Springer, Berlin, 1991)
45. J. Proriol, Sur une famille de polynômes à deux variables orthogonaux dans un triangle. C.R. Acad. Sci. Paris **245**, 2459–2461 (1957)
46. J. Rodal, I. Area, E. Godoy, Orthogonal polynomials of two discrete variables on the simplex. Integral Transform. Spec. Funct. **16**, 263–280 (2005)
47. J. Rodal, I. Area, E. Godoy, Linear partial difference equations of hypergeometric type: orthogonal polynomial solutions in two discrete variables. J. Comput. Appl. Math. **200**, 722–748 (2007)
48. J. Rodal, I. Area, E. Godoy, Structure relations for monic orthogonal polynomials in two discrete variables. J. Math. Anal. Appl. **340**, 825–844 (2008)

49. A. Ronveaux, A. Zarzo, I. Area, E. Godoy, Transverse limits in the Askey tableau. J. Comput. Appl. Math. **99**, 327–335 (1998)
50. A. Ronveaux, A. Zarzo, I. Area, E. Godoy, Bernstein bases and Hahn-Eberlein orthogonal polynomials. Integral Transform. Spec. Funct. **7**(1–2), 87–96 (1998)
51. M.H. Srivastava, P.W. Karlsson, *Multiple Gaussian Hypergeometric Series.* Ellis Horwood Series: Mathematics and its Applications (Ellis Horwood, Chichester, 1985)
52. P.K. Suetin, *Orthogonal Polynomials in Two Variables* (Gordon and Breach Science Publishers, Amsterdam, 1999)
53. D.D. Tcheutia, Y. Guemo Tefo, M. Foupouagnigni, E. Godoy, I. Area, Linear partial divided-difference equation satisfied by multivariate orthogonal polynomials on quadratic lattices. Math. Model. Nat. Pheno. **12**, 14–43 (2017)
54. D.D. Tcheutia, M. Foupouagnigni, Y. Guemo Tefo, I. Area, Divided-difference equation and three-term recurrence relations of some systems of bivariate q-orthogonal polynomials. J. Differ. Equ. Appl. **23**, 2004–2036 (2017)
55. M.V. Tratnik, Multivariable Meixner, Krawtchouk, and Meixner–Pollaczek polynomials. J. Math. Phys. **30**, 2740–2749 (1989)
56. M.V. Tratnik, Some multivariable orthogonal polynomials of the Askey tableau—continuous families. J. Math. Phys. **32**, 2065–2073 (1991)
57. M.V. Tratnik, Some multivariable orthogonal polynomials of the Askey tableau—discrete families. J. Math. Phys. **32**, 2337–2342 (1991)
58. Y. Xu, Second order difference equations and discrete orthogonal polynomials of two variables. Intl. Math. Res. Not. **8**, 449–475 (2005)

Signal Processing, Orthogonal Polynomials, and Heun Equations

Geoffroy Bergeron, Luc Vinet, and Alexei Zhedanov

Abstract A survey of recents advances in the theory of Heun operators is offered. Some of the topics covered include: quadratic algebras and orthogonal polynomials, differential and difference Heun operators associated to Jacobi and Hahn polynomials, connections with time and band limiting problems in signal processing.

Keywords Bispectral problems · Heun equation · Askey scheme · Orthogonal polynomials · Time and band limiting

Mathematics Subject Classification (2000) Primary 33C80; Secondary 94A11

1 Introduction

This lecture aims to present an introduction to the algebraic approach to Heun equation. To offer some motivation, we shall start with an overview of a central problem in signal treatment, namely that of time and band limiting. Our stepping stone will be the fact that Heun type operators play a central role in this analysis thanks to the work of Landau, Pollack and Slepian [19], see also the nice overview in [6]. After reminding ourselves of the standard Heun equation, we shall launch into our forays. We shall recall that all polynomials of the Askey scheme are solutions to bispectral problems and we shall indicate that all their properties can be encoded into quadratic algebras that bear the name of these families. We shall use the Jacobi polynomials as example. We shall then discuss the tridiagonalization procedure designed to move from lower to higher families of polynomials in the Askey hierarchy. This will be illustrated by obtaining the Wilson/Racah polynomials

G. Bergeron · L. Vinet (✉)
Centre de recherches mathématiques, Université de Montréal, Montréal, QC, Canada
e-mail: bergerog@crm.umontreal.ca; vinet@crm.umontreal.ca; luc.vinet@umontreal.ca

A. Zhedanov
Department of Mathematics, School of Information, Renmin University of China, Beijing, China

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_11

from the Jacobi ones or equivalently by embedding the Racah algebra in the Jacobi algebra. We shall then show that the standard Heun operator can be obtained from the most general tridiagonalization of the hypergeometric (the Jacobi) operator. This will lead us to recognize that an algebraic Heun operator can be associated to each entry of the Askey tableau. We shall then proceed to identify the Heun operator associated to the Hahn polynomials. It will be seen to provide a difference version of the standard Heun operator. We shall have a look at the algebra this operator forms with the Hahn operator and of its relation to the Racah algebra. We shall then loop the loop by discussing the finite version of the time and band limiting problem and by indicating how the Heun-Hahn operator naturally provides a tridiagonal operator commuting with the non-local limiting operators. We shall conclude with a summary of the lessons we will have learned.

2 Motivation and Background

2.1 Time and Band Limiting

A central problem in signal processing is that of the optimal reconstruction of a signal from limited observational data. Several physical constraints arise when sampling a signal. We will here focus on those corresponding to a limited time window and to a cap on the detection of frequencies. Consider a signal represented as a function of time by

$$f : \mathbb{R} \longrightarrow \mathbb{R},$$

and suppose f can only be observed for a finite time interval

$$W = [-T, T] \subset \mathbb{R}.$$

This time limiting can be expressed as multiplication by a step function χ_W defined by

$$\chi_W(t) = \begin{cases} 1, & \text{if } -T \leq t \leq T, \\ 0, & \text{otherwise.} \end{cases}$$

Now, suppose the measurements are limited in their bandwidth. This corresponds to an upper bound on accessible frequencies. Let us express this band limiting as multiplication by a step function χ_N of the Fourier transform of the signal f, where

$$\chi_N(n) = \begin{cases} 1, & \text{if } 0 \leq n \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

This defines the time limiting operator χ_W

$$\chi_W : \mathcal{C}(\mathbb{R}) \longrightarrow \mathcal{C}(\mathbb{R}),$$

acting by multiplication on functions of time and the band limiting operator χ_N

$$\chi_N : \mathcal{C}(\mathbb{R}) \longrightarrow \mathcal{C}(\mathbb{R}),$$

acting by multiplication on functions of frequencies. Thus, the available data on f is limited to $\chi_N\, F\, \chi_W\, f$, where F denotes the Fourier transform. The time and band limiting problem consists in the optimal reconstruction of f from the limited available data $\chi_N\, F\, \chi_W\, f$.

In this context, the best approximation of f requires finding the singular vectors of the operator

$$E = \chi_N\, F\, \chi_W,$$

which amounts to the eigenvalue problems for the following operators

$$E^*E = \chi_W F^{-1} \chi_N F \chi_W, \quad \text{and} \quad EE^* = \chi_N F \chi_W F^{-1} \chi_N.$$

For F the standard Fourier transform, one has

$$\begin{aligned}
\left[EE^* \tilde{f}\right](l) &= \chi_N \int_{-T}^{T} e^{ilt} \left(\int_0^N \tilde{f}(k) e^{-ikt} dk \right) dt \\
&= \chi_N \int_0^N \tilde{f}(k) \left(\int_{-T}^{T} e^{i(l-k)t} dt \right) dk, \\
&= \int K_T(l,k) \tilde{f}(k) dk, \qquad (2.1)
\end{aligned}$$

where

$$K_T(l,k) = \int_{-T}^{T} e^{i(l-k)T} dt = \frac{\sin(l-k)T}{(l-k)},$$

which is the integral operator with the well-known *sinc kernel*. It is known, that non local operators such as E^*E have spectra that are not well-suited to numerical analysis. This makes difficult obtaining solutions to the time and band limiting problem. However, a remarkable observation of Landau, Pollak and Slepian [12–14, 18, 20] is that there is a differential operator D with a well-behaved spectrum that commutes with the integral operator E^*E. This reduces the time and band limiting problem to the numerically tractable eigenvalue problem for D. In the above example, this operator D is a special case of the Heun operator. The algebraic approach presented here will give indications (in the discrete-discrete case in particular) as to why this "miracle" happens.

2.2 The Heun Operator

Let us first remind ourselves of basic facts regarding the usual Heun operator [9]. The Heun equation is the Fuchsian differential equation with four regular singularities. The standard form is obtained through homographic transformations by placing the singularities at $x = 0, 1, d$ and ∞ and is given by

$$\frac{d^2}{dx^2}\psi(x) + \left(\frac{\gamma}{x} + \frac{\delta}{x-1} + \frac{\epsilon}{x-d}\right)\frac{d}{dx}\psi(x) + \frac{\alpha\beta x - q}{x(x-1)(x-d)}\psi(x) = 0,$$

where

$$\alpha + \beta - \gamma - \delta + 1 = 0,$$

to ensure regularity of the singular point at $x = \infty$. This Heun equation can be written in the form

$$M\psi(x) = \lambda\psi(x)$$

with M the Heun operator given by

$$M = x(x-1)(x-d)\frac{d^2}{dx^2} + (\rho_2 x^2 + \rho_1 x + \rho_0)\frac{d}{dx} + r_1 x + r_0, \tag{2.2}$$

with

$$\begin{aligned} \rho_2 &= -(\gamma + \delta + \epsilon), & \rho_1 &= (\gamma + \delta)d + \gamma + \epsilon, \\ \rho_0 &= -\gamma d, & & \\ r_1 &= -\alpha\beta, & r_0 &= q + \lambda. \end{aligned}$$

One can observe that M sends any polynomial of degree n to a polynomial of degree $n+1$. Indeed, the Heun operator can be defined as the most general second order differential operator with this property.

3 The Askey Scheme and Bispectral Problems

A pair of linear operators X and Y is said to be bispectral if there is a two-parameter family of common eigenvectors $\psi(x, n)$ such that one has

$$X\psi(x,n) = \omega(x)\psi(x,n)$$
$$Y\psi(x,n) = \lambda(n)\psi(x,n),$$

where, X acts on the variable n and Y, on the variable x. These relations define two representations for the operators X and Y, the "x" and the "n" representations depending on which side of the equations is adopted (see below). It is understood that the same representation is used when computing products of operators. For the band-time limiting problem associated the sinc kernel, one has the two-parameter family of eigenfunctions given by $\psi(t,n) = e^{itn}$ with the bispectral pair identified as

$$X = -\frac{d^2}{dn^2}, \qquad \omega(t) = t^2,$$
$$Y = -\frac{d^2}{dt^2}, \qquad \lambda(n) = n^2.$$

In this case, in each of the representation, one of the operator is differential. In general, bispectral pairs can be realized in terms of continuous and discrete operators.

A key observation is that each family of hypergeometric polynomials of the Askey scheme defines a bispectral problem. Indeed, these polynomials are the solution to both a recurrence relation and a differential or difference equation. By associating X with the recurrence relation and Y with the differential or difference equation, one forms a bispectral problem as follows. In the x-representation, X acts a multiplication by the variable and Y as the differential or difference operator while in the n-representation, X acts as a three-term difference operator over n and Y as multiplication by the eigenvalue. The family of common eigenvectors are the orthogonal polynomials.

As a relevant example, consider the (monic) Jacobi polynomials $\hat{P}_n^{(\alpha,\beta)}(x)$ defined as follows [10]

$$\hat{P}_n^{(\alpha,\beta)}(x) = \frac{(-1)^n(\alpha+1)_n}{(\alpha+\beta+n+1)_n}\, {}_2F_1\left(\begin{matrix}-n, n+\alpha+\beta+1\\ \alpha+1\end{matrix}; x\right).$$

These polynomials are the eigenvectors of the hypergeometric operator D_x given by

$$D_x \equiv x(x-1)\frac{d^2}{dx^2} + (\alpha+1-(\alpha+\beta+2)x)\frac{d}{dx}, \tag{3.1}$$

such that

$$D_x \, \hat{P}_n^{(\alpha,\beta)}(x) = \lambda_n \, \hat{P}_n^{(\alpha,\beta)}(x),$$

with eigenvalues given by $\lambda_n = -n(n+\alpha+\beta+1)$. They form an orthogonal set:

$$\int_0^1 \hat{P}_n^{(\alpha,\beta)}(x)\hat{P}_m^{(\alpha,\beta)}(x)x^\alpha(1-x)\beta dx = h_n\delta_{n,m}, \tag{3.2}$$

where

$$h_n = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta_2)}u_1u_2\cdots u_n.$$

The Jacobi polynomials also satisfy the three-term recurrence relation given by

$$x\,\hat{P}_n^{(\alpha,\beta)}(x) = \hat{P}_{n+1}^{(\alpha,\beta)}(x) + b_n\hat{P}_n^{(\alpha,\beta)}(x) + u_n\,\hat{P}_{n-1}^{(\alpha,\beta)}(x), \tag{3.3}$$

where

$$u_n = \frac{n(n+\alpha)(n+\beta)(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)},$$
$$b_n = \frac{1}{2} + \frac{\alpha^2-\beta^2}{4}\left(\frac{1}{2n+\alpha+\beta} - \frac{1}{2n+\alpha+\beta+2}\right).$$

Taking

$$X = x, \qquad Y = D_x,$$

for the x-representation and

$$X = T_n^+ + b_n \cdot 1 + u_nT_n^-, \qquad Y = \lambda_n, \qquad \text{where} \qquad T_n^\pm f_n = f_{n\pm 1},$$

for the n-representation, the Jacobi polynomials provide a two-parameter set of common eigenvectors of X and Y and hence of the bispectral problem they define. This construction arises similarly for all the orthogonal polynomials in the Askey scheme.

3.1 An Algebraic Description

The properties of the orthogonal polynomials of the Askey scheme can be encoded in an algebra as follows. For any such polynomials, take the X operator to be the multiplication by the variable and the Y operator as the differential or difference equation they satisfy. Consider then the associative algebra generated by K_1, K_2 and K_3 where

$$K_1 \equiv X, \qquad K_2 \equiv Y, \qquad K_3 \equiv [K_1, K_2]. \tag{3.4}$$

Upon using these definitions for the generators, one can derive explicitly the commutation relations to obtain that $[K_2, K_3]$ and $[K_3, K_1]$ are quadratic expressions in K_1 and K_2. Once these relations have been identified, the algebra can be posited abstractly and the properties of the corresponding polynomials follow from representation theory.

Sitting at the top of the Askey scheme, the Wilson and Racah polynomials [10] are the most general ones and the algebra encoding their properties encompasses the others. As the algebraic description is insensitive to truncation, both the Wilson and Racah polynomials are associated to the same algebra. This algebra is known as the Racah-Wilson or *Racah algebra* and is defined [4] as the associative algebra over $\mathbb{C}$ generated by $\{K_1, K_2, K_3\}$ with relations

$$[K_1, K_2] = K_3 \tag{3.5}$$

$$[K_2, K_3] = a_1\{K_1, K_2\} + a_2K_2^2 + bK_2 + c_1K_1 + d_1I \tag{3.6}$$

$$[K_3, K_1] = a_1K_1^2 + a_2\{K_1, K_2\} + bK_1 + c_2K_2 + d_2I, \tag{3.7}$$

where $a_1, a_2, b, c_1, c_2, d_1$ and d_2 are structure parameters and where $\{A, B\} = AB + BA$ denotes the anti-commutator. One can show that the Jacobi identity is satisfied. The Racah algebra naturally arises in the study of classical orthogonal polynomials but has proved useful in the construction of integrable models and in representation theory [3, 4].

Other polynomials of the Askey scheme can be obtained from the Racah or Wilson polynomials by limits and specializations. The associated algebras can be obtained from the Racah algebra in the same way. In particular, the *Jacobi algebra* [5] constitutes one such specialization where $a_1, c_1, d_1, d_2 \to 0$. Indeed, taking

$$\begin{aligned} A_1 &= Y = D_x \equiv x(x-1)\frac{d^2}{dx^2} + (\alpha + 1 - (\alpha+\beta+2)x)\frac{d}{dx}, \\ A_2 &= X = x, \quad A_3 \equiv [A_1, A_2] = 2x(x-1)\frac{d}{dx} - (\alpha+\beta+2)x + \alpha + 1, \end{aligned} \tag{3.8}$$

one finds the following relations for the Jacobi algebra

$$[A_1, A_2] = A_3 \tag{3.9}$$

$$[A_2, A_3] = a_2A_2^2 + dA_2 \tag{3.10}$$

$$[A_3, A_1] = a_2\{A_1, A_2\} + dA_1 + c_2A_2 + e_2, \tag{3.11}$$

where $a_2 = 2$, $d = -2$, $c_2 = -(\alpha + \beta)(\alpha + \beta + 2)$ and $e_2 = (\alpha + 1)(\alpha + \beta)$.

3.2 Duality

The bispectrality of the polynomials in the Askey scheme is related to a notion of duality where the variable and the degree are exchanged. In the algebraic description, this corresponds to exchanging the X and Y operator. Let us make details explicit in the finite-dimensional case where the polynomials satisfy both a second order difference equation and a three-term recurrence relation [8].

In finite dimension, both the X and Y operator will admit a finite eigenbasis. Let us denote the eigenbasis of X by $\{e_n\}$ and the one of Y by $\{d_n\}$ for $n = 0, 1, 2, \ldots, N$. One first notices that Y will be tridiagonal in the X eigenbasis and likewise for X in the Y eigenbasis. Explicitly, one has

$$\begin{aligned} Xe_n &= \lambda_n e_n, & Yd_n &= \mu_n d_n, \\ Xd_n &= a_{n+1}d_{n+1} + b_n d_n + a_n d_{n-1}, & Ye_n &= \xi_{n+1}e_{n+1} + \eta_n e_n + \xi_n e_{n-1}, \\ & & & n = 0, 1, \ldots, N \end{aligned} \tag{3.12}$$

where $\{a_n\}$, $\{b_n\}$, $\{\xi_n\}$ and $\{\eta_n\}$ for $n = 0, 1, \ldots, N$ are scalar coefficients. As both the X and Y eigenbases span the same space, one can expand one basis onto the other as follows

$$e_s = \sum_{n=0}^{N} \sqrt{w_s}\phi_n(\lambda_s)d_n, \tag{3.13}$$

where $\phi_n(x)$ are the polynomials associated to the algebra defined by the following recurrence relation

$$a_{n+1}\phi_{n+1}(x) + b_n\phi_n(x) + a_n\phi_{n-1}(x) = x\phi_n(x), \quad \phi_{-1} = 0, \quad \phi_0 = 1,$$

which verify the orthogonality relation

$$\sum_{s=0}^{N} w_s\phi_n(\lambda_s)\phi_m(\lambda_s) = \delta_{n,m},$$

so that the reverse expansion is easily seen to be

$$d_n = \sum_{s=0}^{N} \sqrt{w_s}\phi_n(\lambda_s)e_s.$$

Consider now the *dual* set of polynomials $\chi_n(x)$ defined by the following recurrence relation

$$\xi_{n+1}\chi_{n+1}(x) + \eta_n\chi_n(x) + \xi_n\chi_{n-1}(x) = x\chi_n(x), \quad \chi_{-1} = 0, \ \chi_0 = 1,$$

which are orthogonal with respect to the dual weights $\tilde{w}_s$:

$$\sum_{s=0}^{N} \tilde{w}_s\chi_n(\mu_s)\chi_m(\mu_s) = \delta_{n,m}. \tag{3.14}$$

These dual polynomials provide an alternative expansion of one basis onto the other. One has

$$d_s = \sum_{n=0}^{N} \sqrt{\tilde{w}_s}\chi_n(\mu_s)e_n. \tag{3.15}$$

One readily verifies this expansion by applying Y to obtain

$$\begin{aligned} Yd_s &= \sum_{n=0}^{N} \sqrt{\tilde{w}_s}\chi_n(\mu_s)Ye_n = \sum_{n=0}^{N} \sqrt{\tilde{w}_s}\chi_n(\mu_s)[\xi_{n+1}e_{n+1} + \eta_n e_n + \xi_n e_{n-1}] \\ &= \sum_{n=0}^{N} \sqrt{\tilde{w}_s}[\xi_{n+1}\chi_{n+1}(\mu_s) + \eta_n\chi_n(\mu_s) + \xi_n\chi_{n-1}(\mu_s)]e_n = \mu_s d_s. \end{aligned}$$

Using the orthogonality of the polynomials $\{\chi_n(\mu_s)\}$ given by (3.14), the expansion (3.15) is inverted as

$$e_n = \sum_{s=0}^{N} \sqrt{\tilde{w}_s}\chi_N(\mu_s)d_s.$$

Comparing the above with the first expansion in (3.13), knowing the $\{d_n\}$ to be orthogonal, one obtains

$$\sqrt{w_s}\phi_n(\lambda_s) = \sqrt{\tilde{w}_n}\chi_s(\mu_n), \tag{3.16}$$

a property known as Leonard duality [15], see also [21] for an introduction to Leonard pairs.

4 Tridiagonalization of the Hypergeometric Operator

Tridiagonalization enables one to construct orthogonal polynomials with more parameters from simpler ones and thus to build a bottom-up characterization of the families of the Askey scheme from this bootstrapping. In particular, properties of the Wilson and Racah polynomials can be found from the tridiagonalization of the hypergeometric operator [5]. Moreover, by considering the most general tridiagonalization, one recovers the complete Heun operator [7].

4.1 The Wilson and Racah Polynomials from the Jacobi Polynomials

In the canonical realization of the Jacobi algebra in terms of differential operators presented in (3.8), one of the generators is the hypergeometric operator (3.1) and the other is the difference operator in the degree corresponding to the recurrence relation (3.3). We consider the construction of an operator in the algebra which is tridiagonal in the eigenbases of both operators.

Let $Y = D_x$ be the hypergeometric operator and $X = x$ be multiplication by the variable. Define M in the Jacobi algebra as follows

$$M = \tau_1 XY + \tau_2 YX + \tau_3 X + \tau_0, \tag{4.1}$$

where τ_i, $i = 0, 1, 2, 3$ are scalar parameters. Knowing that X leads to the three-term recurrence relation of the Jacobi polynomials $\hat{P}_n^{(\alpha,\beta)}(x)$:

$$X\,\hat{P}_n^{(\alpha,\beta)}(x) = x\,\hat{P}_n^{(\alpha,\beta)}(x) = \hat{P}_{n+1}^{(\alpha,\beta)}(x) + b_n\,\hat{P}_n^{(\alpha,\beta)}(x) + u_n\,\hat{P}_{n-1}^{(\alpha,\beta)}(x),$$

and is obviously tridiagonal, it is clear from (4.1) that M will also be tridiagonal in the eigenbasis of Y that the Jacobi polynomials form. One has

$$M\,\hat{P}_n^{(\alpha,\beta)}(x) = \xi_{n+1}\,\hat{P}_{n+1}^{(\alpha,\beta)}(x) + \eta_n\,\hat{P}_n^{(\alpha,\beta)}(x) + b_n u_n\,\hat{P}_{n-1}^{(\alpha,\beta)}(x), \tag{4.2}$$

where

$$\begin{aligned} \xi_n &= \tau_1 \lambda_{n-1} + \tau_2 \lambda_n + \tau_3, \\ \eta_n &= (\tau_1 + \tau_2)\lambda_n b_n + \tau_3 b_n, \\ b_n &= \tau_1 \lambda_n + \tau_2 \lambda_{n-1} + \tau_3. \end{aligned}$$

If $\tau_1 + \tau_2 = 0$, then M simplifies to $M = \tau_1[X, Y] + \tau_3 X$, which is a first order differential operator. In order for M to remain a second order operator, one demands that $\tau_1 + \tau_2 \neq 0$. In this case, normalizing M so that $\tau_1 + \tau_2 = 1$, one obtains explicitly

$$M = x^2(x-1)\frac{d^2}{dx^2} + x[\alpha + 1 - 2\tau_2 - (\alpha + \beta - 2\tau_2)x]\frac{d}{dx} \\ - [\tau_2(\alpha + \beta + 2) - \tau_3]x + (\alpha + 1)\tau_2 + \tau_0. \qquad (4.3)$$

We now construct a basis in which M is diagonal. In the realization (3.8), where the algebra acts on functions of x, X is multiplication by x and its inverse is defined by

$$X^{-1} : f(x) \longmapsto \frac{1}{x} f(x).$$

With this definition, one can invert the expression for M given by (4.1) to obtain

$$Y = \tau_1 X^{-1} M + \tau_2 M X^{-1} + (2\tau_1\tau_2 - \tau_0)X^{-1} - (2\tau_1\tau_2 + \tau_3). \qquad (4.4)$$

Observing that (4.4) has the same structure as (4.1) under the transformation $X \mapsto X^{-1}$, the eigenfunctions of M can be constructed as follows. Introduce the variable $y = 1/x$ and conjugate M and Y by a monomial in y to obtain

$$\tilde{Y} = y^{\nu-1} Y y^{1-\nu}, \qquad \tilde{M} = y^{\nu-1} M y^{1-\nu}.$$

Then, by demanding that

$$\tau_3 = (4 + \alpha + \beta - \nu)(\tau_2 + \nu - 1) - \nu\tau_2,$$

the conjugated operators take the following form,

$$-\tilde{Y} = y^2(y-1)\frac{d^2}{dy^2} + y(a_1 y + b_1)\frac{d}{dy} + c_1 y + d_1,$$

$$-\tilde{M} = y(y-1)\frac{d^2}{dy^2} + (a_2 y + b_2)\frac{d}{dy} + d_2,$$

with all the new parameters being simple expressions in terms of α, β, τ_0, τ_2 and ν. Up to a global sign, one recognizes $\tilde{M}$ as the hypergeometric operator in terms of the variable y, while $\tilde{Y}$ is similar to M. As the Jacobi polynomials diagonalizes the hypergeometric operator, the eigenvectors satisfying

$$M\psi_n(x) = \tilde{\lambda_n}\psi_n(x) \qquad (4.5)$$

are easily found to be

$$\psi_n(x) = x^{\nu-1}\hat{P}_n^{(\tilde{\alpha},\tilde{\beta})}(1/x), \qquad \tilde{\lambda_n} = n(n+\tilde{\alpha}+\tilde{\beta}+1),$$
$$\tilde{\beta} = \beta, \qquad \tilde{\alpha} = 2(\tau_2+\nu) - \alpha - \beta - 7.$$

It follows from the recurrence relation of the Jacobi polynomials (3.3) that X^{-1} is tridiagonal in the basis $\psi_n(x)$ as it corresponds to multiplication by the variable. Thus, a glance at (4.4) confirms that Y is tridiagonal in the $\psi_n(x)$ basis.

In order to relate this result with the Wilson and Racah orthogonal polynomials, consider the expansion of $\psi_n(x)$ in terms of $\hat{P}_k^{(\alpha,\beta)}(x)$. One has

$$\psi_n(x) = \sum_{k=0}^{\infty} G_k(n)\hat{P}_k^{(\alpha,\beta)}(x). \tag{4.6}$$

By factoring the expansion coefficients as $G_k(n) = G_0(n)\Xi_k Q_k(n)$, one finds using (4.2) and (4.5) that, for a unique choice of Ξ_k, Q_k satisfies the following three-term recurrence relation

$$\tilde{\lambda_n} Q_k(n) = B_k Q_{k+1}(n) + U_k Q_k(n) + F_k Q_{k-1}(n),$$

where

$$\begin{aligned} B_k &= u_{k+1}(\tau_1\lambda_{k+1} + \tau_2\lambda_k + \tau_3), \\ U_k &= \lambda_k b_k + \tau_3 b_k, \\ F_k &= \tau_1\lambda_{k-1} + \tau_2\lambda_k + \tau_3. \end{aligned} \tag{4.7}$$

The recurrence relation allows to identify the factor $Q_k(n)$ of the expansion coefficient in (4.6) as the four-parameter Wilson polynomials $W_n(x; k_1, k_2, k_3, k_4)$. In this construction, two of these parameters are inherited from the Jacobi polynomials while, after scaling, the tridiagonalisation introduced two free parameters.

The Racah polynomials occur in this setting when a supplementary restriction is introduced. Indeed, a glance at (4.3) shows that the generic M operator maps polynomials of degree n into polynomials of degree $n+1$. However, one can see from (4.2) that if

$$\xi_{N+1} = \tau_1\lambda_N + \tau_2\lambda_{N+1} + \tau_3 = 0,$$

both Y and M preserve the space of polynomials of degree less than or equal to N. This truncation condition is satisfied when $\nu = N+1 = 2-2\tau_2$. In this case, the eigenvectors of M are

$$\psi_n(x) = x^N \hat{P}_n^{(N-\alpha-\beta-4,\beta)}(1/x),$$

which are manifestly polynomials of degree $N - n$. One then considers again the expansion of the basis element $\psi_n(x)$ into $\hat{P}_k^{(\alpha,\beta)}(x)$ to obtain

$$\psi_n(x) = \sum_{k=0}^{N} R_{n,k} \hat{P}_k^{(\alpha,\beta)}(x),$$

where the expansion coefficients $R_{n,k}$ can be shown to be given in terms of the Racah polynomials. Using the orthogonality of the Jacobi polynomials given in (3.2), one obtains

$$R_{n,k} h_k = \int_0^1 \psi_n(x) \hat{P}_k^{(\alpha,\beta)}(x)\, x^\alpha (1-x)^\beta dx,$$

an analog of the Jacobi-Fourier transform of Koornwinder [11], giving an integral representation of the Racah polynomials.

It was stated earlier that the properties of the orthogonal polynomials in the Askey scheme are encoded in their associated algebras. This can be seen from the construction of the Wilson and Racah polynomials from the Jacobi polynomials by the tridiagonalization procedure which corresponds algebraically to an embedding of the Racah algebra in the Jacobi algebra. This is explicitly given by

$$K_1 = A_1, \qquad K_2 = \tau_1 A_2 A_1 + \tau_2 A_1 A_2 + \tau_3 A_2, \tag{4.8}$$

where A_1, A_2 are the Jacobi algebra generators as in (3.8). One shows that K_1 and K_2 as defined in (4.8) verify the relations (3.5) of the Racah algebra assuming that A_1 and A_2 verify the Jacobi relations as given in (3.9). Thus, the embedding (4.8) encodes the tridiagonalization result abstractly.

The tridiagonalisation (4.1) used to derive higher polynomials from the Jacobi polynomials is not the most general tridiagonal operator that can be constructed from the Jacobi algebra generators. Indeed, consider the addition in (4.1) of a linear term in Y, given by (3.1):

$$M = \tau_1 XY + \tau_2 YX + \tau_3 X + \tau_4 Y + \tau_0. \tag{4.9}$$

It is straightforward to see that M as given by (4.9) is equal to the Heun operator (2.2). Expressed as in (4.9), the Heun operator is manifestly tridiagonal on the Jacobi polynomials, which offers a simple derivation of a classical result. For the finite dimensional situation see [16].

5 The Algebraic Heun Operator

The emergence of the standard Heun operator from the tridiagonalization of the hypergeometric operator suggests that Heun-type operators can be associated to bispectral problems. In particular, knowing all polynomials in the Askey scheme to define bispectral problems, there should be Heun-like operators associated to each of these families of polynomials. Guided by this observation, consider a set of polynomials in the Askey scheme and let X and Y be the generators of the associated algebra as in (3.4). As before, X is the recurrence operator and Y, the difference or differential operator. The corresponding Heun-type operator W is defined as

$$W = \tau_1 XY + \tau_2 YX + \tau_3 X + \tau_4 Y + \tau_0, \tag{5.1}$$

and will be referred to as an *algebraic Heun operator* [8]. The operator W associated to a polynomial family will have features similar to those of the standard Heun operator which arises in the context of the Jacobi polynomials. To illustrate this, a construction that parallels the one made for the Jacobi polynomials is presented next.

5.1 A Discrete Analog of the Heun Operator

The standard Heun operator can be defined as the most general degree increasing second order differential operator. In analogy with this, one defines the *difference Heun operator* as:

Definition (Difference Heun Operator) The difference Heun operator is the most general second order difference operator on a uniform grid which sends polynomials of degree n to polynomials of degree $n + 1$.

We now obtain an explicit expression for the difference Heun operator on the finite grid $G = \{0, 1, \ldots, N\}$. Let $T^{\pm}$ be shift operators defined by

$$T^{\pm} f(x) = f(x \pm 1), \tag{5.2}$$

and take W to be a generic second order difference operator with

$$W = A_1(x)T^{+} + A_2(x)T^{-} + A_0(x)I. \tag{5.3}$$

By demanding that W acting on $1, x$ and x^2 yields polynomials of one degree higher, one obtains that

$$A_0(x) = \tilde{\pi}_1(x) - \tilde{\pi}_3(x), \qquad A_1(x) = \frac{\tilde{\pi}_3(x) - \tilde{\pi}_2(x)}{2}, \qquad A_2(x) = \frac{\tilde{\pi}_3(x) + \tilde{\pi}_2(x)}{2}, \tag{5.4}$$

where the $\tilde{\pi}_i(x)$ are arbitrary polynomials of degree i for $i = 1, 2, 3$. Thus, in general, $A_i(x)$ for $i = 0, 1, 2$ are third degree polynomials with $A_1(x)$ and $A_2(x)$ having the same leading coefficient. Moreover, the restriction of the action of W to the finite grid G implies that A_1 has $(x - N)$ as a factor and A_2 has x as a factor. Hence, one has

$$\begin{aligned} A_1(x) &= (x - N)(\kappa x^2 + \mu_1 x + \mu_0), \\ A_2(x) &= x(\kappa x^2 + \nu_1 x + \nu_0), \\ A_0(x) &= -A_1(x) - A_2(x) + r_1 x + r_0, \end{aligned}$$

for $\mu_0, \mu_1, \nu_0, \nu_1, r_0, r_1$ and κ arbitrary parameters. Then, it is easy to see that

$$W[x^n] = \sigma_n \, x^{n+1} + O(x^n),$$

for a certain σ_n depending on the parameters. We shall see next that this difference Heun operator coincides with the algebraic Heun operator associated to the Hahn algebra.

5.2 The Algebraic Heun Operator of the Hahn Type

The Hahn polynomials P_n are orthogonal polynomials belonging to the Askey scheme. As such, an algebra encoding their properties is obtained as a specialization of the Racah algebra (3.5) by taking $a_2 \to 0$. One obtains the Hahn algebra, generated by $\{K_1, K_2, K_3\}$ with the following relations

$$\begin{aligned} [K_1, K_2] &= K_3, \\ [K_2, K_3] &= a\{K_1, k_2\} + bK_2 + c_1 K_1 + d_1 I, \\ [K_3, K_1] &= aK_1^2 + bK_1 + c_2 K_2 + d_2 I. \end{aligned} \tag{5.5}$$

A natural realization of the Hahn algebra is given in terms of the bispectral operators associated to the Hahn polynomials P_n, namely,

$$\begin{aligned} X &= K_1 = x, \\ Y &= K_2 = B(x)T^+ + D(x)T^- - (B(x) - D(x))I, \end{aligned} \tag{5.6}$$

with

$$B(x) = (x - N)(x + \alpha + 1), \qquad D(x) = x(x - \beta - N - 1),$$

and where $T^{\pm}$ is as in (5.2). The action of Y is diagonal in the basis given by the Hahn polynomial P_n and is

$$YP_n(x) = \lambda P_n(x), \qquad \lambda_n = n(n + \alpha + \beta + 1).$$

One checks that X and Y satisfy the Hahn algebra relations (5.5) with the structure constants expressed in terms of α, β and N.

Upon identifying the algebra associated to the Hahn polynomials, one can introduce the algebraic Heun operator W of the Hahn type [22] using the generic definition (5.1). In this realization, one finds that W can be written as

$$W = A_1(x)T^+ + A_2(x)T^- + A_0(x)I,$$

where

$$\begin{aligned}
A_1(x) &= (x - N)(x + \alpha + 1)((\tau_1 + \tau_2)x + \tau_2 + \tau_4),\\
A_2(x) &= x(x - \beta - N - 1)((\tau_1 + \tau_2)x + \tau_4 - \tau_2),\\
A_0(x) &= -A_1(x) - A_2(x) + ((\alpha + \beta + 2)\tau_2 + \tau_3)x + \tau_0 - N(\alpha + 1)\tau_2.
\end{aligned}$$

As announced, the operator defined above coincides, upon identification of parameters, with the difference Heun operator W given in (5.3) and (5.4) and defined through its degree raising action on polynomials. That the difference Heun operator is tridiagonal on the Hahn polynomials then follows as a direct result. This parallels the construction in the Jacobi algebra that led to a simple proof of the standard Heun operator being tridiagonal on the Jacobi polynomials. Moreover, in the limit $N \to \infty$, the difference Heun operator W goes to the standard Heun operator, which further supports the appropriateness of the abstract definition (5.1) for the algebraic Heun operator.

To conclude this algebraic analysis, let us consider the algebra generated by Y and W in the context of the Hahn algebra. By introducing a third generator given by $[W, Y]$ and using the relation of the Hahn algebra in (5.5), one finds that the algebra thus generated closes as a cubic algebra with relations given by

$$\begin{aligned}
[Y, [W, Y]] &= g_1Y^2 + g_2\{Y, W\} + g_3Y + g_4W + g_5I,\\
[[W, Y], W] &= e_1Y^2 + e_2Y^3 + g_2W^2 + g_1\{Y, W\} + g_3W + g_6Y + g_7I,
\end{aligned}$$

where the structure constants depend on the parameters of the Hahn polynomials and the parameters of the tridiagonalization (5.1). One can recognize the above as a generalization of the Racah algebra (3.5) with the following two additional terms:

$$e_1Y^2 + e_2Y^3.$$

The conditions for these terms to vanish are given by

$$\tau_1 + \tau_2 = 0, \qquad \tau_2 \pm \tau_4 = 0.$$

When these equalities are satisfied, the operator W simplifies to W_+ or W_- with

$$W_\pm = \pm\frac{1}{2}[X, Y] \pm \gamma X - \frac{Y}{2} \pm \epsilon I.$$

Moreover, any pair from the set $\{Y, W_+, W_-\}$ satisfies the Racah algebra relations given by (3.5). Thus, the choice of a pair of operators specifies an embedding of the Racah algebra in the Hahn algebra, which is analogous to the embedding given in (4.8). These embeddings encode abstractly the construction of the Racah polynomials starting from the Hahn polynomials and provide another example where higher polynomials are constructed from simpler ones.

6 Application to Time and Band Limiting

We now return to the problem of time and band limiting. Consider a finite dimensional bispectral problem as the one associated to the Hahn polynomials. Denote by $\{e_n\}$ and $\{d_n\}$ for $n = 1, 2, \dots, N$ the two eigenbases of this bispectral problem such that

$$X : \{e_n\} \to \{e_n\}, \qquad Xe_n = \lambda_n e_n,$$
$$Y : \{d_n\} \to \{d_n\}, \qquad Yd_n = \mu_n d_n.$$

In this context, X can be thought of being associated to discrete time and Y to frequencies. Suppose now that the spectrum of both X and Y are restricted. These restrictions can be modelled as limiting operators in the form of two projections π_1 and π_2 given by

$$\pi_1 e_n = \begin{cases} e_n & \text{if } n \le J_1, \\ 0 & \text{if } n > J_1, \end{cases} \qquad \pi_2 d_n = \begin{cases} d_n & \text{if } n \le J_2, \\ 0 & \text{if } n > J_2, \end{cases} \tag{6.1}$$
$$\pi_1^2 = \pi_1, \qquad \pi_2^2 = \pi_2.$$

Simultaneous restrictions on the eigensubspaces of X and Y accessible to sampling lead to the two limiting operators

$$V_1 = \pi_1\pi_2\pi_1 = E_1E_2, \quad V_2 = \pi_2\pi_1\pi_2 = E_2E_1,$$

with

$$E_1 = \pi_1 \pi_2, \qquad E_2 = \pi_2 \pi_1.$$

Here, the limiting operator V_1 and V_2 are symmetric and are diagonalizable. A few limit cases are simple. When there are no restriction, $J_1 = J_2 = N$, in which case $V_1 = V_2 = I$. If the restriction is on only one of the spectra, for instance if $J_2 = N$, then $V_1 = V_2 = \pi_1$ having $J_1 + 1$ unit eigenvalues and the other $N - J_1$ equal to zero. However, the case where J_1 and J_2 are arbitrary is much more complicated.

In the generic case, the eigenbasis expansions (3.13) and (3.15) can be used to evaluate the action of π_2 on an eigenvector of X. One has,

$$\pi_2 e_n = \sum_{s=0}^{J_2} \sqrt{w_n} \phi_s(\lambda_n) d_s = \sum_{s=0}^{J_2} \sum_{t=0}^{N} \sqrt{w_n \tilde{w}_s} \phi_s(\lambda_n) \chi_t(\mu_s) e_t.$$

Similarly, one can evaluate the action of π_1 on eigenvectors of Y and obtain

$$V_1 e_n = \pi_1 \pi_2 \pi_1 e_n = \sum_{t=0}^{J_1} \sum_{s=0}^{J_2} \sqrt{w_n \tilde{w}_s} \phi_s(\lambda_n) \chi_t(\mu_s) e_t = \sum_{t=0}^{J_1} K_{t,n} e_t, \tag{6.2}$$

with

$$\begin{aligned} K_{t,n} &= \sum_{s=0}^{J_2} \sqrt{w_n \tilde{w}_s} \phi_s(\lambda_n) \chi_t(\mu_s) \\ &= \sum_{s=0}^{J_2} \sqrt{w_n w_t} \phi_s(\lambda_n) \phi_s(\lambda_t) \\ &= \sum_{s=0}^{J_2} \sqrt{\tilde{w}_s} \chi_n(\mu_s) \chi_t(\mu_s), \end{aligned} \tag{6.3}$$

where the Leonard duality relation (3.16) has been used to obtain the last two equalities. The operator V_1 in (6.2) is the discrete analog of the integral operator (2.1) that restricts both in time and frequency, with (6.3) being the discrete kernel. As in the continuous case, V_1 and V_2 are non-local operator and the problem of finding their eigenvectors is numerically difficult. However, if there exists a tridiagonal matrix M that commutes with both V_1 and V_2, then M would admit eigenvectors that are shared with V_1 and V_2. This renders the discrete time and band limiting problem well controlled. In this context, the tridiagonal matrix M is the discrete analog of a second order differential operator and plays the role of the differential operator found by Landau, Pollak and Slepian for the continuous time and band limiting problems.

Tridiagonal matrices that commute with the limiting operators π_1 and π_2 in (6.1) will also commute with V_1 and V_2. One then wants to find for M such that

$$[M, \pi_1] = [M, \pi_2] = 0. \tag{6.4}$$

Taking M to be an algebraic Heun operator with

$$M = \tau_1 XY + \tau_2 YX + \tau_3 X + \tau_4 Y,$$

and using (6.4), one finds the following conditions

$$\tau_2 = \tau_1, \qquad \tau_1(\lambda_{J_1} + \lambda_{J_1+1}) + \tau_4 = 0, \qquad \tau_1(\mu_{J_2} + \mu_{J_2+1}) + \tau_3 = 0.$$

Except for the Bannai-Ito spectrum, it is always possible to find τ_3 and τ_4 satisfying the above [8], see also [17]. Hence, the algebraic Heun operator provides the commuting operator that enables efficient solutions to the time and band limiting.

7 Conclusion

This lecture has offered an introduction to the concept of algebraic Heun operators and their applications. This construct stems from the observation that the standard Heun operator can be obtained from the tridiagonalization of the hypergeometric operator.The key idea is to focus on operators that are bilinear in the generators of the quadratic algebras associated to orthogonal polynomials. The Heun type operators obtained in this algebraic fashion, coincide with those arising from the definition that has Heun operators raising by one the degree of arbitrary polynomials. This has been illustrated for the discrete Heun operator in its connection to the Hahn polynomials. This notion of algebraic Heun operators tied to bispectral problems has moreover been seen to shed light on the occurence of commuting operators in band and time limiting analyses. The exploration of these algebraic Heun operators and the associated algebras has just begun [1, 2, 22] but the results found so far let us believe that it could lead to significant new advances.

Acknowledgements One of us (Luc Vinet) is very grateful to Mama Foupouagnigni, Wolfram Koepf, AIMS (Cameroun) and the Volkswagen Stiftung for the opportunity to lecture in Douala. Geoffroy Bergeron benefitted from a NSERC postgraduate scholarship. The research of Luc Vinet is supported by a NSERC discovery grant and that of Alexei Zhedanov by the National Science Foundation of China (Grant No. 11711015).

References

1. P. Baseilhac, L. Vinet, A. Zhedanov, The q-Heun operator of big q-Jacobi type and the q-Heun algebra (2018). Preprint. arXiv:1808.06695
2. P. Baseilhac, S. Tsujimoto, L. Vinet, A. Zhedanov, The Heun-Askey-Wilson algebra and the Heun operator of Askey-Wilson type (2018). Preprint. arXiv:1811.11407
3. V.X. Genest, L. Vinet, A. Zhedanov, The equitable Racah algebra from three $\mathfrak{su}(1, 1)$ algebras. J. Phys. A **47**(2), 025203 (2013)
4. V.X. Genest, L. Vinet, A. Zhedanov, The Racah algebra and superintegrable models. J. Phys. Conf. Ser. **512**(1), 012011 (2014)
5. V.X. Genest, M.E.H. Ismail, L. Vinet, A. Zhedanov, Tridiagonalization of the hypergeometric operator and the Racah-Wilson algebra. Proc. Am. Math. Soc. **144**, 4441–4454 (2016)
6. F.A. Grünbaum, Time-band limiting and the bispectral problem. Commun. Pure Appl. Math. **47**, 307–328 (1994)
7. F.A. Grünbaum, L. Vinet, A. Zhedanov, Tridiagonalization and the Heun equation. J. Math. Phys. **58**, 31703 (2017)
8. F.A. Grünbaum, L. Vinet, A. Zhedanov, Algebraic Heun operator and band-time limiting. Commun. Math. Phys. **364**(3), 1041–1068 (2018)
9. G. Kistenson, *Second Order Differential Equations* (Springer, New York, 2010)
10. R. Koekoek, R.F. Swarttouw, The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue. Delft University of Technology, Faculty of Information Technology and Systems, Department of Technical Mathematics and Informatics, Report no. 98–17 (1998)
11. T.H. Koornwinder, Special orthogonal polynomial systems mapped onto each other by the Fourier-Jacobi transform, in *Polynômes Orthogonaux et Applications*. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 174–183
12. H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty - II. Bell Syst. Tech. J. **40**(1), 65–84 (1961)
13. H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty - III: the dimension of the space of essentially time- and band-limited signals. Bell Syst. Tech. J. **40**(4), 1295–1336 (1962)
14. H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty - V: the discrete case. Bell Syst. Tech. J. **57**(5), 1371–1430 (1978)
15. D.A. Leonard, Orthogonal polynomials, duality and association schemes. SIAM J. Math. Anal. **13**(4), 656–663 (1981)
16. K. Nomura, P. Terwilliger, Linear transformations that are tridiagonal with respect to both eigenbases of a Leonard pair. Linear Algebra Appl. **420**(1), 198–207 (2007)
17. R.K. Perline, Discrete time-band limiting operators and commuting tridiagonal matrices. SIAM. J. Algebr. Discrete Methods **8**(2), 192–195 (1987)
18. D. Slepian, Prolate spheroidal wave functions, fourier analysis and uncertainty - IV: extensions to many dimensions; generalized prolate spheroidal functions. Bell Syst. Tech. J. **43**(6), 3009–3058 (1964)
19. D. Slepian, Some comments on fourier analysis, uncertainty and modeling. SIAM Rev. **25**(3), 379–393 (1983)
20. D. Slepian, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty - I. Bell Syst. Tech. J. **40**(1), 43–64 (1961)
21. P. Terwilliger, Introduction to Leonard pairs. J. Comput. Appl. Math. **153**(1–2), 463–475 (2003)
22. L. Vinet, A. Zhedanov, The Heun operator of the Hahn type (2018). Preprint. arXiv:1808.00153

Some Characterization Problems Related to Sheffer Polynomial Sets

Hamza Chaggara, Radhouan Mbarki, and Salma Boussorra

Abstract In this work, we show some properties of Sheffer polynomials arising from quasi-monomiality. We survey characterization problems dealing with d-orthogonal polynomial sets of Sheffer type. We revisit some families in the literature and we state an explicit formula giving the exact number of Sheffer type d-orthogonal sets. We investigate, in detail, the $(d+1)$-fold symmetric case as well as the particular cases $d = 1, 2, 3$.

Keywords Sheffer polynomials · d-orthogonal polynomials · Generating functions · $(d+1)$-fold symmetric polynomials · Quasi-monomiality

Mathematics Subject Classification (2000) 33C45, 44A20, 65D20, 41A58

1 Introduction

Let $\mathcal{P}$ be the vector space of polynomials with coefficients in $\mathbb{C}$ and let $\mathcal{P}'$ be its algebraic dual. A polynomial sequence $\{P_n\}_{n\geq 0}$ in $\mathcal{P}$ is called a *polynomial set* (PS, for short) if and only if $\deg P_n = n,\ n = 0, 1, 2, \ldots$. we denote by $\langle u, f\rangle$ the effect of the functional $u \in \mathcal{P}'$ on the polynomial $f \in \mathcal{P}$.

H. Chaggara (✉)
King Khalid University College of Sciences, Department of Mathematics, Abha, Kingdom of Saudi Arabia

École Supérieure des Sciences et de la Technologie, Université de Sousse, Sousse, Tunisia
e-mail: hamza.chaggara@ipeim.rnu.tn; hshaggara@kku.edu.sa

R. Mbarki
Institut Préparatoire aux Études d'Ingénieur de Monastir, Université de Monastir, Monastir, Tunisia

S. Boussorra
École Supérieure des Sciences et de la Technologie, Université de Sousse, Sousse, Tunisia

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_12

A set of polynomials $\{P_n\}_{n\geq 0}$ is monic if the leading coefficient of each P_n is unitary.

A PS is said to be of Sheffer type (or of Sheffer A-type zero [22, 26]) if it has a generating function of the form:

$$G(x,t) = A(t)\exp(xC(t)) = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} t^n, \tag{1.1}$$

where A and C are given by

$$A(t) = \sum_{k=0}^{\infty} a_k t^k, \quad a_0 \neq 0, \tag{1.2}$$

$$C(t) = \sum_{k=0}^{\infty} c_k t^{k+1}, \quad c_0 \neq 0. \tag{1.3}$$

This means that $A(t)$ is invertible and $C(t)$ has a compositional inverse.

We note that the Sheffer polynomials have many applications and have been extensively investigated. They contain a large class of polynomial sequences that include Laguerre polynomials, Meixner polynomials, Bernoulli polynomials, Stirling polynomials and many others. For $C(t) = t$, we meet the definition of Appell polynomial sets.

In [29], Sheffer wrote a remarkable paper on some properties of polynomials of type zero. Rainville, based on the manipulation of formal series, found important properties of Sheffer type polynomials. These properties are reported in his book [26].

One of the most successful languages commonly used for handling Sheffer sequences is the so-called Umbral Calculus [22, 27].

We are interested, in this work, in studying Sheffer sequences when they are d-orthogonal. We will give the general properties of these sequences and a classification in some particular cases.

An important generalization of the notion of orthogonality was introduced by Van Iseghem [31] and Maroni [24], that is the d-orthogonality (d being a non-negative integer):

A PS $\{P_n\}_{n\geq 0}$ is said to be a *d-orthogonal polynomial set* (d-OPS) if there exists a functional vector, $\mathfrak{U} = {}^t(u_0, u_1, \ldots, u_{d-1})$ such that

$$\begin{cases} \langle u_k, P_m P_n \rangle = 0, \quad \text{if} \quad m > nd + k, \\ \\ \langle u_k, P_n P_{nd+k} \rangle \neq 0, \quad n \geq 0, \quad k \in \{0, 1, \ldots, d-1\}. \end{cases} \tag{1.4}$$

The notions of d-dimensional orthogonality for polynomials [24], vectorial orthogonality as defined and studied in [30] or simultaneous orthogonality as in [1, 14] and even more generally the multiple orthogonality [22] are obviously generalizations of the notion of ordinary orthogonality for polynomials.

The d-orthogonality has been the subject of numerous investigations and applications. In particular, it is related to the study of vector Padé approximants [30, 31], simultaneous Padé approximants [14], vectorial continued fractions, polynomial solutions of higher-order differential equations, spectral study of multidiagonal nonsymmetric operators [2] and infinite dynamical systems [13].

A generalized spectral theorem characterizing d-orthogonal polynomials by a recurrence relation which was given by Van Iseghem and Maroni in [24, 30].

A monic PS $\{P_n\}_{n\geq 0}$ is a d-OPS if and only if it fulfils a $(d+1)$-order recurrence relation, that is, a relation between $d+2$ consecutive polynomials of the form:

$$P_{n+1}(x) = (x+\alpha_{n+1})P_n(x) + \sum_{k=1}^{d}\binom{n}{k}\beta_k^{(n+1)}P_{n-k}(x) \quad ; \quad \beta_d^{(n+1)} \neq 0, \quad \forall n \geq 0, \tag{1.5}$$

with $P_{-n} = 0$ for $n \geq 1$ and $\binom{n}{k} = 0$ if $k > n$.

In [20], Maroni extended the notion of symmetric polynomials ($P_n(-x) = (-1)^n P_n(x)$) into $(d+1)$-fold symmetric ones as follows:

A PS $\{P_n\}_{n\geq 0}$ is called $(d+1)$-fold symmetric (d-SOPS) if it fulfils:

$$P_n(\omega x) = \omega^n P_n(x), \quad \omega = \exp\Big(\frac{2i\pi}{d+1}\Big). \tag{1.6}$$

Douak and Maroni characterized $(d+1)$-fold symmetric d-orthogonal polynomials by means of a specific recurrence relation [20]:

A monic d-OPS $\{P_n\}_{n\geq 0}$ is $(d+1)$-fold symmetric if and only if it satisfies a $(d+1)$-order recurrence relation of the form:

$$\begin{cases} P_{n+1}(x) = xP_n(x) + \binom{n}{d}\beta_d^{(n+1)}P_{n-d}(x),\ n \geq d,\ \beta_d^{(n+1)} \neq 0, \\ P_n(x) = x^n,\ 0 \leq n < d. \end{cases} \tag{1.7}$$

Most of the known d-orthogonal families were introduced as solutions of characterization problems. Such problems consist in finding all d-OPSs having a given property: with a particular generating function [10], $(d+1)$-fold symmetry property [5], classical and semi-classical character [11, 28], Appell type property [6, 18], specific hypergeometric form [8], etc.

Characterization problems related to Sheffer PSs have a deep history. Many authors have worked on solving the following problem:

Find all d-orthogonal polynomials of Sheffer type.

In a classical paper [25], Meixner considered the case $d = 1$ (orthogonal PS of Sheffer type) and determined all orthogonal polynomials including Hermite, Charlier, Laguerre, Meixner, and Meixner-Pollaczek polynomials. It is the so-called classical Meixner class. Kubo [23] reconsidered Meixner's classification from the viewpoint of multiplicative renormalization method and concerned the effects of affine transforms and multipliers.

The case $d = 2$ has been treated by Boukhemis and Maroni [12]. The general case was recently studied by Ben Cheikh and Gam [7], by Chaggara and Mbarki [16] and also by Varma [32].

The organization of this lecture will be as follows. In Sect. 2, we collect the basic results used in the rest of this work, namely, the quasi-monomiality principle. In Sect. 3, we show how to derive some corresponding properties of Sheffer Sequences using the quasi-monomiality principle (operational rules related to appropriate operators). We recover some well known results which characterize the polynomial sets of the Sheffer type. In Sect. 4, we survey characterization theorems dealing with polynomial sets which are d-OPS and, as a particular case, the d-SOPS families [16]. We treat in detail the cases $d = 1, 2, 3$.

2 Preliminary Results

For $j \in \mathbb{Z}$, we denote by $\Lambda^{(j)}$ the space of operators acting on analytic functions that augment (respectively reduce) the degree of every polynomial by exactly j if $j \geq 0$ (respectively $j \leq 0$). That includes the fact that if $\lambda \in \Lambda^{(-1)}$, $\lambda(1) = 0$. We present, in this part, some basic definitions and results which we need in the rest of this work.

2.1 Operators

Definition 2.1 ([4]) A PS $\{P_n\}_{n\geq 0}$ is called *quasi-monomial* if it is possible to define two operators λ and ρ, independent of n, such that:

$$\lambda P_n = nP_{n-1}, \tag{2.1}$$

$$\rho P_n = P_{n+1}. \tag{2.2}$$

Here λ and ρ play the roles analogous, respectively, to the derivative and multiplication (by x) operators on monomials $\{x^n\}_{n\geq 0}$.

Definition 2.2 Let $\lambda \in \Lambda^{(-1)}$. A PS $\{B_n\}_{n\geq 0}$ is called a sequence of basic polynomials for λ if

1. $B_n(0) = \delta_{n,0}$, $\forall n \in \mathbb{N}$, where $\delta_{n,0}$ the Kronecker symbol.
2. $\lambda B_n = nB_{n-1}$, $\forall n \in \mathbb{N}^*$.

In [3], it was shown that every $\lambda \in \Lambda^{(-1)}$ has a unique sequence of basic polynomials.

Theorem 2.3 ([4]) *Let $\{P_n\}_{n\geq 0}$ be a PS, then there exists an unique triplet $(\lambda, \rho, \tau) \in \Lambda^{(-1)} \times \Lambda^{(1)} \times \Lambda^{(0)}$ of operators on $\mathcal{P}$ such that:*

$$\begin{cases} \lambda P_n = n P_{n-1}, \\ \rho P_n = P_{n+1}, \\ \tau B_n = P_n, \end{cases}$$

where $\{B_n\}_{n\geq 0}$ is the sequence of basic polynomials for λ.

We refer to the λ, ρ and τ operators as, respectively, the *lowering*, the *raising* and the *transfer* operators associated with the PS $\{P_n\}_{n\geq 0}$.

Some properties of $\{P_n\}_{n\geq 0}$ can be deduced from the structure of these operators. Namely,

1. The two operators λ and ρ, satisfy the commutation relation

$$[\lambda, \rho] = \lambda\rho - \rho\lambda = 1.$$

   and thus display a Weyl group structure.
2. P_n is an eigenfunction of the operators $\lambda\rho$ and $\rho\lambda$ associated, respectively, to n and $n+1$ as eigenvalues.
   That is to say:

$$\rho\lambda P_n = n P_n \quad \text{and} \quad \lambda\rho P_n = (n+1) P_n. \tag{2.3}$$

   which can be viewed, if λ and ρ have a differential or a difference realization, as a differential or a difference equation satisfied by P_n.
3. If P_0 is given, then the polynomial $P_n(x)$ can be explicitly constructed as:

$$P_n = \rho^n P_0. \tag{2.4}$$

4. The point 3 implies that a generating function of $\{P_n\}_{n\geq 0}$ can always be brought in the form:

$$e^{\rho t} P_0(x) = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} t^n. \tag{2.5}$$

In our investigation, we need the following technical lemma which is a special case for a more general formula known as Crofton type operational rule [15].

Lemma 2.4 ([15, Lemma 3.1]) *Let h be a formal power series, then, for all $f \in \mathcal{P}$, we have:*

$$h(D)(xf) = \Big(xh(D) + h'(D)\Big)(f), \tag{2.6}$$

where $D = \frac{d}{dx}$ designates the derivative operator.

Proof We write $h(t) = \sum_{n=0}^{\infty} \alpha_n t^n$ and according to Leibniz' formula, we get:

$$\begin{aligned} h(D)(xf) &= \sum_{n=0}^{\infty} \alpha_n D^n(xf) = \sum_{n=0}^{\infty} \alpha_n \left(\sum_{k=0}^{n} \binom{n}{k} x^{(k)} f^{(n-k)} \right) \\ &= x \sum_{n=0}^{\infty} \alpha_n D^n(f) + \sum_{n=1}^{\infty} n\alpha_n D^{n-1}(f) = xh(D)(f) + h'(D)(f). \end{aligned}$$

□

Proposition 2.5 ([4]) *Let* $\{P_n\}_{n\geq 0}$ *be a PS generated by:*

$$G(x,t) = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} t^n. \tag{2.7}$$

Then we have the equivalences:

$$\lambda G(x,t) = tG(x,t) \Leftrightarrow \lambda P_n = nP_{n-1}, \quad \forall n \in \mathbb{N}^*, \tag{2.8}$$

$$\rho G(x,t) = \frac{\partial}{\partial t} G(x,t) \Leftrightarrow \rho P_n = P_{n+1}, \quad \forall n \in \mathbb{N}. \tag{2.9}$$

3 Properties of Sheffer Polynomials

The Sheffer polynomial sets generated by (1.1) have been shown to be quasi-monomials under the action of the operators [4]

$$\lambda = C^{-1}(D), \quad \rho = \frac{A'}{A}(\lambda) + xC'(\lambda) \quad \text{and} \quad \tau = A(\lambda). \tag{3.1}$$

If $C(t) = t$, we have $\lambda = D$. That corresponds to Appell polynomial sets.

If $A(t) = 1$, we obtain the basic sets of Sheffer type.

Proposition 3.1 (Sheffer Identity [27]) *A PS* $\{S_n\}_{n\geq 0}$ *is of Sheffer type if and only if it satisfies the identity:*

$$S_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} S_k(x) B_{n-k}(y), \tag{3.2}$$

where $\{B_n\}_{n\geq 0}$ *is a given sequence of basic Sheffer polynomials.*

Proof We have

$$G(x,t)=\sum_{n=0}^{\infty}\frac{S_n(x)}{n!}t^n=A(t)\exp(xC(t)).$$

Then

$$\begin{aligned}G(x+y,t)=A(t)\exp((x+y)C(t))&=\sum_{n=0}^{\infty}\frac{S_n(x)}{n!}t^n\sum_{n=0}^{\infty}\frac{B_n(y)}{n!}t^n\\&=\sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{S_k(x)}{k!}t^k\frac{B_{n-k}(y)}{(n-k)!}t^{n-k}=\sum_{n=0}^{\infty}\sum_{k=0}^{n}\binom{n}{k}S_k(x)B_{n-k}(y)\frac{t^n}{n!},\end{aligned}$$

which establishes the formula. □

Definition 3.2 (PS of Binomial Type [27]) A PS $\{B_n\}_{n\geq 0}$ is of binomial type if it satisfies the relation:

$$B_n(x+y)=\sum_{k=0}^{n}\binom{n}{k}B_k(x)B_{n-k}(y). \tag{3.3}$$

It follows immediately that the PSs of binomial type are exactly the basic Sheffer PSs.

Theorem 3.3 ([26]) *Let $\{P_n\}_{n\geq 0}$ be a Sheffer PS, generated by (1.1). Then there exist two sequences $(\alpha_k)_k$ and $(\beta_k)_k$, independent of x and n, such that for $n\geq 1$,*

$$xP_n'(x)-nP_n(x)=-n!\left[\sum_{k=0}^{n-1}\alpha_k\frac{P_{n-k-1}(x)}{(n-k-1)!}+x\sum_{k=0}^{n-1}\beta_k\frac{P_{n-k-1}'(x)}{(n-k-1)!}(x)\right], \tag{3.4}$$

where

$$\sum_{n=0}^{\infty}\alpha_n t^{n+1}=\frac{tA'(t)}{A(t)}\quad \textit{and}\quad 1+\sum_{n=0}^{\infty}\beta_n t^{n+1}=\frac{tC'(t)}{C(t)}. \tag{3.5}$$

Proof

$$\begin{aligned}\rho G(x,t)=\frac{\partial}{\partial t}G(x,t)\iff&\left[\frac{A'(t)}{A(t)}+xC'(t)\right]G(x,t)=\frac{\partial}{\partial t}G(x,t)\\\iff& xt\frac{C'(t)}{C(t)}\frac{\partial}{\partial x}G(x,t)-t\frac{\partial}{\partial t}G(x,t)=-t\frac{A'(t)}{A(t)}G(x,t)\end{aligned}$$

Substituting the expansions (3.5), we get:

$$\left[1+\sum_{n=0}^{\infty}\beta_n t^{n+1}\right]\left[\sum_{n=0}^{\infty}\frac{xP_n'(x)}{n!}t^n\right]-\sum_{n=0}^{\infty}\frac{nP_n(x)}{n!}t^n=-\left[\sum_{n=0}^{\infty}\alpha_n t^{n+1}\right]\left[\sum_{n=0}^{\infty}\frac{P_n(x)}{n!}t^n\right].$$

Therefore

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{xP_n'(x)-nP_n(x)}{n!}t^n&=-\left[\sum_{n=0}^{\infty}\alpha_n t^{n+1}\right]\left[\sum_{n=0}^{\infty}\frac{P_n(x)}{n!}t^n\right]-\left[\sum_{n=0}^{\infty}\beta_n t^{n+1}\right]\left[\sum_{n=0}^{\infty}\frac{xP_n'(x)}{n!}t^n\right]\\&=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\left(-\alpha_k\frac{P_{n-k}(x)}{(n-k)!}-\beta_k\frac{xP_{n-k}'(x)}{(n-k)!}\right)\right)t^{n+1}\\&=\sum_{n=1}^{\infty}\left(\sum_{k=0}^{n-1}\left(-\alpha_k\frac{P_{n-k-1}(x)}{(n-k-1)!}-\beta_k\frac{xP_{n-k-1}'(x)}{(n-k-1)!}\right)\right)t^n.\end{aligned}$$

By identification, we get the desired result. □

Theorem 3.4 ([26]) *Let $\{P_n\}_{n\geq 0}$ be a PS of Sheffer type, generated by (1.1). There exist two sequences $(\alpha_k)_{k\geq 0}$ and $(c_k)_{k\geq 0}$, independent of x and n, such that for $n\geq 1$,*

$$\sum_{k=0}^{n-1}(\alpha_k+x(k+1)c_k)\frac{n!}{(n-k-1)!}P_{n-1-k}(x)=nP_n(x), \tag{3.6}$$

where

$$\sum_{k=0}^{\infty}\alpha_k t^k=\frac{A'(t)}{A(t)},\quad \text{and}\quad \sum_{k=0}^{\infty}c_k t^k=C(t). \tag{3.7}$$

Proof We have, by virtue of (2.3),

$$nP_n=\rho\lambda P_n=\left[\frac{A'}{A}(\lambda)+xC'(\lambda)\right]\lambda P_n=\frac{A'}{A}(\lambda)\lambda P_n+xC'(\lambda)\lambda P_n,$$

therefore

$$nP_n=\sum_{k=0}^{\infty}\alpha_k\lambda^{k+1}P_n+x\sum_{k=0}^{\infty}(k+1)c_k\lambda^{k+1}P_n=\sum_{k=0}^{n-1}(\alpha_k+x(k+1)c_k)\frac{n!}{(n-k-1)!}P_{n-1-k}.$$

□

Theorem 3.5 ([26]) *Under the assumptions of Theorem 3.4, there exist two sequences* $(\mu_k)_{k\geq 0}$ *and* $(\nu_k)_{k\geq 0}$*, independent of x and n, such that for* $n \geq 1$*,*

$$\sum_{k=0}^{n-1}(\mu_k + x\nu_k)D^{k+1}P_n(x) = nP_n(x), \tag{3.8}$$

where

$$\sum_{k=0}^{\infty}\mu_k t^{k+1} = C^{-1}(t)\,\frac{A'}{A}(C^{-1}(t)), \quad \textit{and} \quad \sum_{k=0}^{\infty}\nu_k t^{k+1} = \frac{C^{-1}(t)}{\left(C^{-1}\right)'(t)}. \tag{3.9}$$

Proof We have:

$$(n+1)P_n = \lambda\rho P_n = \lambda\left[\frac{A'}{A}(\lambda) + xC'(\lambda)\right] = \lambda\frac{A'}{A}(\lambda)P_n + \lambda xC'(\lambda)P_n.$$

Since, $C'(t) = \sum_{k=0}^{\infty}(k+1)c_k t^k$, so

$$xC'(\lambda)P_n = x\sum_{k=0}^{\infty}(k+1)c_k\lambda^k P_n = \sum_{k=0}^{\infty}(k+1)c_k\frac{n!}{(n-k)!}xP_{n-k}. \tag{3.10}$$

According to (2.6), with $h = C^{-1}$,

$$\lambda(xP_{n-k}) = xC^{-1}(D)P_{n-k} + \left(C^{-1}\right)'(D)P_{n-k}. \tag{3.11}$$

Then, by (3.10) and (3.11), we obtain

$$\lambda(xC'(\lambda)P_n) = x\lambda C'(\lambda)P_n + \left(C^{-1}\right)'(D)C'(C^{-1}(D))P_n = x\lambda C'(\lambda)P_n + P_n,$$

since $\left(C^{-1}\right)'(t)C'(C^{-1}(t)) = (C\circ C^{-1})'(t) = 1$. Therefore

$$nP_n = \left(\lambda\frac{A'}{A}(\lambda)\right)P_n + x\lambda C'(\lambda)P_n.$$

Hence by (3.9), we get (3.8). □

Theorem 3.6 ([26]) *A PS* $\{P_n\}_{n\geq 0}$ *is of Sheffer type, if and only if there exists a sequence* $(c_k)_{k\geq 0}$*, independent of x and n, such that for* $n \geq 1$*,*

$$P_n' = \sum_{k=0}^{n-1}c_k\frac{n!}{(n-k-1)!}P_{n-1-k}. \tag{3.12}$$

Proof

$$\lambda P_n = C^{-1}(D)P_n \Leftrightarrow C(\lambda)P_n = P_n' \Leftrightarrow \sum_{k=0}^{\infty} c_k \lambda^{k+1} P_n = P_n'$$
$$\Leftrightarrow \sum_{k=0}^{n-1} c_k \frac{n!}{(n-k-1)!} P_{n-1-k} = P_n'.$$

□

4 Characterization Problems for Sheffer Sets

4.1 Characterization Theorem

The following theorem characterizes the d-OPSs of Sheffer type. The proof is based on recurrence relations and lowering operators. An alternative proof of a similar result based on suitable transform operators was already given firstly in [7] and then in [32].

Theorem 4.1 ([16, Theorem 2.1]) *Let $\{P_n\}_{n\geq 0}$ be a monic PS generated by (1.1). Then the following statements are equivalent:*

(a) $\{P_n\}_{n\geq 0}$ is a d-OPS.
(b) A and C satisfy two equations of the form:

$$C'(t) = \frac{1}{1+\sum_{k=0}^{d} \theta_k t^{k+1}}, \tag{4.1}$$

and

$$\frac{A'(t)}{A(t)} = \frac{\sum_{k=0}^{d} \sigma_k t^k}{1+\sum_{k=0}^{d} \theta_k t^{k+1}} = \left(\sum_{k=0}^{d} \sigma_k t^k\right) C'(t), \tag{4.2}$$

with the regularity conditions:

$$\sigma_d \neq 0 \quad \textit{and} \quad \frac{\theta_d}{\sigma_d} \notin \{\frac{1}{n}, n \in \mathbb{N}^*\}. \tag{4.3}$$

Proof $(a) \Rightarrow (b)$ By the hypothesis $\{P_n\}_{n\geq 0}$ is a d-OPS, therefore it satisfies the recurrence relation (1.5).

The proof of the necessary condition will be divided into two steps:

Step 1
Applying $\lambda = C^{-1}(D)$ to (1.5), with $\widehat{\beta}_k^{(n+1)} := \binom{n}{k}\beta_k^{(n+1)}$ and according to (2.6), we obtain $\forall n \in \mathbb{N}$,

$$(n+1)P_n(x) = n(x+\alpha_{n+1})P_{n-1}(x) + \left(C^{-1}\right)'(D)P_n(x) + \sum_{k=1}^{d}(n-k)\widehat{\beta}_k^{(n+1)}P_{n-k-1}(x). \tag{4.4}$$

By shifting the index $n \leftrightarrow n-1$ in (1.5) and by multiplication by n, we get:

$$nP_n(x) = n(x+\alpha_n)P_{n-1}(x) + n\sum_{k=1}^{d}\widehat{\beta}_k^{(n)}P_{n-k-1}(x). \tag{4.5}$$

The difference (4.4)–(4.5) gives:

$$P_n(x) = n(\alpha_{n+1}-\alpha_n)P_{n-1}(x) + \left(C^{-1}\right)'(D)P_n(x) + \sum_{k=1}^{d}\left((n-k)\widehat{\beta}_k^{(n+1)} - n\widehat{\beta}_k^{(n)}\right)P_{n-k-1}(x). \tag{4.6}$$

Therefore

$$\left(1-\left(C^{-1}\right)'(D)\right)P_n(x) = n(\alpha_{n+1}-\alpha_n)P_{n-1}(x) + \sum_{k=1}^{d}\frac{n!}{(n-k-1)!}\left(\gamma_k^{(n+1)}-\gamma_k^{(n)}\right)P_{n-k-1}(x), \tag{4.7}$$

where $\gamma_k^{(n+1)} = \frac{(n-k)!}{n!}\widehat{\beta}_k^{(n+1)} = \frac{\beta_k^{(n+1)}}{k!},\ 1 \le k \le d,\ n \ge 0.$

In (4.4), we replace n by $n+1$ and apply $C^{-1}(D)$. This gives

$$(n+2)P_n = n(x+\alpha_{n+2})P_{n-1} + 2\left(C^{-1}\right)'(D)P_n + \sum_{k=1}^{d}\frac{(n-k)(n+1-k)}{n+1}\widehat{\beta}_k^{(n+2)}P_{n-k-1}. \tag{4.8}$$

Subtracting (4.8) from (4.4) yields:

$$\left(1-\left(C^{-1}\right)'\right)(D)P_n = n(\alpha_{n+2}-\alpha_{n+1})P_{n-1} + \sum_{k=1}^{d}\frac{n!(\gamma_k^{(n+2)}-\gamma_k^{(n+1)})}{(n-k-1)!}P_{n-k-1}. \tag{4.9}$$

By identification of (4.7) and (4.9), we obtain for $k = 1, 2, \dots, d$ and $n \in \mathbb{N}$,

$$\alpha_{n+2} - \alpha_{n+1} = \alpha_{n+1} - \alpha_n \quad \text{and} \quad \gamma_k^{(n+2)} - \gamma_k^{(n+1)} = \gamma_k^{(n+1)} - \gamma_k^{(n)},$$

It follows that the two sequences $(\alpha_{n+1})_n$ and $(\gamma_k^{(n+1)})_n$ are arithmetic progressions.

Let $\theta_0 = \alpha_n - \alpha_{n+1}$ and $\theta_k = \gamma_k^{(n)} - \gamma_k^{(n+1)},\ 1 \le k \le d,\ n \ge 0.$

Then (4.7) becomes:

$$\begin{aligned}\left(1 - \left(C^{-1}\right)'(D)\right)P_n(x) &= -\left(\theta_0 C^{-1}(D)P_n(x) + \sum_{k=1}^{d} \frac{n!}{(n-k-1)!}\theta_k P_{n-k-1}(x)\right)\\ &= -\left(\theta_0 C^{-1}(D) + \sum_{k=1}^{d} \theta_k [C^{-1}(D)]^{k+1}\right)P_n(x),\end{aligned}$$

which means that $\left(C^{-1}\right)'(D) = 1 + \sum_{k=0}^{d} \theta_k [C^{-1}(D)]^{k+1}$. So

$$\left(C^{-1}\right)'(t) = 1 + \sum_{k=0}^{d} \theta_k [C^{-1}(t)]^{k+1}.$$

This is equivalent to:

$$C'(t) = \frac{1}{1 + \sum_{k=0}^{d} \theta_k t^{k+1}}.$$

Step 2
From the above, we can write:

$$\alpha_{n+1} - \alpha_1 = -n\theta_0; \quad \gamma_k^{(n+1)} - \gamma_k^{(k+1)} = -(n-k)\theta_k;$$

then

$$\alpha_{n+1} = -n\theta_0 + \alpha_1 \quad \text{and} \quad \widehat{\beta}_k^{(n+1)} = -\frac{n!}{(n-k-1)!}\theta_k + \binom{n}{k}\widehat{\beta}_k^{(k+1)}, \quad 1 \le k \le d,\ n \ge 0. \tag{4.10}$$

Substituting x by 0 in (1.5):

$$P_{n+1}(0) = (-\theta_0 n + \alpha_1)P_n(0) + \sum_{k=1}^{d}\left(-\frac{n!}{(n-k-1)!}\theta_k + \frac{n!}{(n-k)!}\frac{\widehat{\beta}_k^{(k+1)}}{k!}\right)P_{n-k}(0).$$

We have $A(t)=\sum_{n=0}^{\infty}\frac{P_n(0)}{n!}t^n$, then

$$\begin{aligned}A'(t)&=\sum_{n=0}^{\infty}\frac{P_{n+1}(0)}{n!}t^n\\&=-\theta_0\sum_{n=1}^{\infty}\frac{P_n(0)}{(n-1)!}t^n+\alpha_1\sum_{n=0}^{\infty}\frac{P_n(0)}{n!}t^n-\sum_{k=1}^{d}\theta_k\sum_{n=k+1}^{\infty}\frac{P_{n-k}(0)}{(n-k-1)!}t^n\\&\quad+\sum_{k=1}^{d}\frac{\beta_k^{(k+1)}}{k!}\sum_{n=k}^{\infty}\frac{P_{n-k}(0)}{(n-k)!}t^n\\&=-\theta_0tA'(t)+\alpha_1A(t)-\Big(\sum_{k=1}^{d}\theta_kt^{k+1}\Big)A'(t)+\Big(\sum_{k=1}^{d}\frac{\widehat{\beta_k}^{(k+1)}}{k!}t^k\Big)A(t).\end{aligned}$$

Putting $\sigma_0=\alpha_1$ and $\sigma_k=\frac{\widehat{\beta}_k^{(k+1)}}{k!}$, we deduce:

$$\frac{A'(t)}{A(t)}=\frac{\sum_{k=0}^{d}\sigma_kt^k}{1+\sum_{k=0}^{d}\theta_kt^{k+1}};\quad \sigma_d=\frac{\widehat{\beta}_d^{(d+1)}}{d!}\neq 0. \tag{4.11}$$

According to (4.10), $\widehat{\beta}_d^{(n+1)}=-\frac{n!}{(n-d-1)!}\theta_d+\binom{n}{d}\widehat{\beta}_d^{(d+1)}\neq 0$, then $\frac{\theta_d}{\sigma_d}\neq\frac{1}{k}$, $k\in\mathbb{N}^*$.
$(b)\Rightarrow(a)$ We have:

$$\rho=\frac{A'}{A}(\lambda)+xC'(\lambda)=\left(\sum_{k=0}^{d}\sigma_k\lambda^k\right)C'(\lambda)+xC'(\lambda),\ \sigma_d\neq 0,$$

then

$$\rho\left(1+\sum_{k=0}^{d}\theta_k\lambda^{k+1}\right)P_n=\rho P_n+\sum_{k=0}^{d}\theta_k\rho\lambda^{k+1}P_n=\sum_{k=0}^{d}\sigma_k\lambda^kP_n+xP_n.$$

with

$$\sigma_d\neq 0,\ \frac{\theta_d}{\sigma_d}\notin\left\{\frac{1}{k},\quad k\in\mathbb{N}^*\right\}.$$

Therefore

$$P_{n+1} = -\sum_{k=0}^{d} \theta_k \frac{n!}{(n-k-1)!} P_{n-k} + \sum_{k=0}^{d} \sigma_k \frac{n!}{(n-k)!} P_{n-k} + x P_n$$
$$= (x - n\theta_0 + \sigma_0)\, P_n + \sum_{k=1}^{d} n! \left(\frac{\sigma_k}{(n-k)!} - \frac{\theta_k}{(n-k-1)!} \right) P_{n-k},$$

with $n!\left(\frac{\sigma_d}{(n-d)!} - \frac{\theta_d}{(n-d-1)!}\right) = \frac{n!}{(n-d-1)!}\left(\frac{\sigma_d}{n-d} - \theta_d\right) \neq 0,\ n > d$, according to the regularity conditions (4.3). Therefore $\{P_n\}_{n\geq 0}$ is a d-OPS.

Hence, the converse statement of the theorem is established. □

In [16], the authors characterized $(d+1)$-fold symmetric d-OPS of Sheffer type as follows:

Corollary 4.2 *A Sheffer type PS $\{P_n\}_{n\geq 0}$ generated by (1.1) is a d-SOPS if and only if A and C satisfy the following equations:*

$$C'(t) = \frac{1}{1 + \theta_d t^{d+1}} \tag{4.12}$$

and

$$\frac{A'(t)}{A(t)} = \frac{\sigma_d t^d}{1 + \theta_d t^{d+1}}, \tag{4.13}$$

with the regularity conditions (4.3).

Proof If $\{P_n\}_{n\geq 0}$ is a d-SOPS of Sheffer type then by (4.10) and (4.11) we get:

$$\begin{cases} \sigma_d = \dfrac{\widehat{\beta}_d^{(d+1)}}{d!} \neq 0, \\ \widehat{\beta}_d^{(n+1)} = \frac{n!}{(n-d-1)!}\left(\frac{\sigma_d}{n-d} - \theta_d\right) \neq 0,\ n > d. \end{cases}$$

Using (1.6)–(1.7), we obtain that $(\alpha_{n+1})_n$ and $(\widehat{\beta}_k^{(n+1)})_n$ vanish for all $1 \leq k \leq d-1$. Hence $\theta_k = 0$ and $\sigma_k = 0$ for all $0 \leq k \leq d-1$, which gives (4.12) and (4.13). □

4.2 Examples

Some particular cases of d-OPSs represented in Theorem 4.1 and generalizing some classical families are worthy to note:

Hermite Type d-OPS In the case $\theta_k = 0\ \forall k \in \{0, 1, \dots, d\}$, we have:

$$C(t) = t \quad \text{and} \quad A(t) = \exp\left(H_{d+1}(t)\right) \quad \text{where} \quad H_{d+1}(t) = \sum_{k=0}^{d} \frac{\sigma_k}{k+1} t^{k+1},\ \sigma_d \neq 0.$$

It follows that

$$\sum_{n=0}^{\infty} \frac{P_n(x,d)}{n!} t^n = \exp\left(H_{d+1}(t)\right) \exp(xt). \tag{4.14}$$

Here, note that $\{P_n(\cdot, d)\}_{n\geq 0}$ was considered by Douak as the only d-OPS of Appell type [18, Theorem 3.1] for $\sigma_0 = 0$, $\sigma_d = -\frac{1}{(d+1)!}$ and $\sigma_1, \sigma_2, \dots, \sigma_{d-1}$ $d-1$ arbitrary constants. Among such polynomials, the d-symmetric ones are singled out. They are named d-orthogonal Hermite polynomials $\{H_n(\cdot, d)\}_{n\geq 0}$.

The Gould–Hopper PS [21] corresponds to the case $H_{d+1}(t) = \sigma_d t^{d+1}$. That reduces to Hermite PS if $d = 1$.

Charlier Type d-OPS For $\theta_k = 0\ \forall k \in \{1, \dots, d\}$, and $\theta_0 \neq 0$, we obtain

$$C(t) = \frac{1}{\theta_0} \ln(1 + \theta_0 t) \quad \text{and} \quad A(t) = \exp(Q_d(t))(1 + \theta_0 t)^{a_0},$$

where Q_d designates a polynomial of degree d and leading coefficient $\frac{\sigma_d}{\theta_0 d}$, $Q_d(0) = 0$ and $a_0 = \sum_{k=0}^{d} \sigma_k \frac{(-1)^k}{\theta_0^{k+1}}$.

$$G\left(\theta_0(x - a_0), \frac{1}{\theta_0} t\right) = \sum_{n=0}^{\infty} \frac{C_n(x,d)}{n!} t^n = \exp\left(\widetilde{Q}_d(t)\right)(1+t)^x\,; \tag{4.15}$$

where $\widetilde{Q}_d(t) = Q_d\left(\frac{1}{\theta_0} t\right)$. This family was characterized by Ben Cheikh and Zaghouani as the only d-OPS and Δ-Appell family [9, Theorem 1.1], where Δ designates the difference operator: $\Delta f(x) = f(x+1) - f(x)$. The Charlier type d-OPS represents one of the first examples of discrete d-OPSs. The case $d = 1$ corresponds to Charlier PS.

Laguerre Type d-OPS Consider the case $\theta_k = 0\ \forall k \in \{2, \dots, d\}$ and suppose that $1 + \theta_0 t + \theta_1 t^2 = (1 - \alpha t)^2$; $\alpha \neq 0$, then

$$C(t) = \frac{t}{1 - \alpha t} \quad \text{and} \quad A(t) = \exp\left(R_{d-1}(t) + \left(\sum_{k=0}^{d} \frac{\sigma_k}{\alpha^k}\right) \frac{t}{1 - \alpha t}\right)(1 - \alpha t)^{\sum_{k=1}^{d} k \frac{\sigma_k}{\alpha^{k+1}}},$$

where R_{d-1} is a polynomial of degree $d-1$ and leading coefficient $\frac{\sigma_d}{\alpha^2(d-1)}$ satisfying $R_{d-1}(0) = 0$.

If we choose the parameters in such a way that $\sum_{k=0}^{d} \frac{\sigma_k}{\alpha^k} = 0$, we get:

$$G\left(x, \frac{1}{\alpha}t\right) = \sum_{n=0}^{\infty} \frac{L_n^{(a)}(x,d)}{n!} t^n = \exp\left(\widetilde{R}_{d-1}(t)\right)(1-t)^a \exp\left(\frac{xt}{1-t}\right), \tag{4.16}$$

where $\widetilde{R}_{d-1}(t) = R_{d-1}(\frac{t}{\alpha})$ and $a = \sum_{k=1}^{d} \frac{k\,\sigma_k}{\alpha^{k+1}}$.

This family was already studied by Douak [19] in the special case $d = 2$ and then by Ben Cheikh and Zaghouani in the general case [10]. It was also investigated in the multiple orthogonality context by Coussement and Van Assche [17].

Meixner Type d-OPS If $\theta_k = 0, \quad \forall k \in \{2, \ldots, d\}$ and $1 + \theta_0 t + \theta_1 t^2 = (1-\alpha t)(1-\beta t)$; $\alpha \neq \beta \neq 0$, we obtain:

$$C(t) = \frac{1}{\beta - \alpha} \ln\left(\frac{1-\alpha t}{1-\beta t}\right) \quad \text{and} \quad A(t) = \exp(S_{d-1}(t)) \left[\frac{(1-\alpha t)^{\sum_{k=0}^{d} \frac{\sigma_k}{\alpha^k}}}{(1-\beta t)^{\sum_{k=0}^{d} \frac{\sigma_k}{\beta^k}}}\right]^{\frac{1}{\beta-\alpha}},$$

where S_{d-1} is a polynomial of degree $d-1$ and leading coefficient $\frac{\sigma_d}{\alpha\beta(d-1)}$ with $S_{d-1}(0) = 0$.

$$G\left((\beta-\alpha)(x-a), \frac{t}{\beta}\right) = \sum_{n=0}^{\infty} \frac{M_n^{\lambda,\mu}(x,d)}{n!} t^n = \exp\left(\widetilde{S}_{d-1}(t)\right)(1-t)^{-x+\lambda}\left(1-\frac{t}{\mu}\right)^x, \tag{4.17}$$

where

$$\widetilde{S}_{d-1}(t) = S_{d-1}(\frac{t}{\beta}),\ \ a = \frac{1}{\beta-\alpha}\sum_{k=0}^{d}\frac{\sigma_k}{\alpha^k},\ \ \lambda = \frac{1}{\beta-\alpha}\sum_{k=0}^{d}\sigma_k\left(\frac{1}{\alpha^k} - \frac{1}{\beta^k}\right) \text{ and } \mu = \frac{\beta}{\alpha}.$$

The Meixner type d-OPS are, in fact, introduced in [10] as an example of d-OPS family with suitable generating function. The corresponding functional vector was derived when $d = 2$ [10, Theorem 4.3].

4.3 Counting d-OPSs of Sheffer Type

Next, we give an explicit formula of the number of d-OPSs of Sheffer type.

Definition 4.3 A partition of a positive integer n is a non-increasing sequence of positive integers $p_1, p_2, \ldots, p_k$ whose sum is n.
$p(n)$ denotes the number of partitions of the integer n, $p(0)$ is defined to be 1.

Examples

$$
\begin{aligned}
&1 = 1 \Rightarrow p(1) = 1;\\
&2 = 1 + 1 \Rightarrow p(2) = 2;\\
&3 = 1 + 2 = 1 + 1 + 1 \Rightarrow p(3) = 3;\\
&4 = 1 + 3 = 2 + 2 = 1 + 1 + 2 = 1 + 1 + 1 + 1 \Rightarrow p(4) = 5;\\
&5 = 1 + 4 = 2 + 3 = 2 + 2 + 1 = 1 + 1 + 1 + 2\\
&\quad = 1 + 1 + 3 = 1 + 1 + 1 + 1 + 1 \Rightarrow p(5) = 7;\\
&6 = 5 + 1 = 4 + 2 = 3 + 3 = 2 + 3 + 1 = 2 + 2 + 2 = 4 + 1 + 1 = 3 + 1 + 1 + 1\\
&\quad = 2 + 2 + 1 + 1 = 2 + 1 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1 + 1 \Rightarrow p(6) = 11.
\end{aligned}
$$

To compute the number N_d of Sheffer type d-OPSs generated by (1.1), we need to determine the number of solutions of the Eq. (4.1) which depends on $\theta_0, \ldots, \theta_d$. More precisely, this corresponds to the number of ways that the polynomial function $\frac{1}{C'(t)} = 1 + \sum_{k=0}^{d} \theta_k t^{k+1}$ can be factorized according to the degree, and depending on whether the roots are real or complex.

A given formal power series C' corresponds to exactly one formal series $C(t)$ because $C(0) = 0$.

For fixed power series $C(t)$, and for a given polynomial of degree d, $F_d(t) = \sum_{k=0}^{d} \sigma_k t^k$ with regularity conditions (4.3), we get exactly one series $A(t)$ as solution of the differential system

$$
\begin{cases} A'(t) - F_d(t)C'(t)A(t) = 0,\\ A(0) = 1. \end{cases}
$$

The pair $(A(t), C(t))$ generates one solution of the problem.

Two d-orthogonal Sheffer type polynomials obtainable from each other by a linear change of variable or scaling factor cannot be associated to different pairs $(A(t), C(t))$. This is justified by the fact they both are monic (and therefore $A(0) = C'(0) = 1$), and satisfy (1.1).

Proposition 4.4 ([16]) *The number N_d of d-OPS generated by (1.1) is given by:*

$$
N_d = N_{d-1} + \sum_{k=0}^{[\frac{d+1}{2}]} p(d + 1 - 2k)p(k), \quad \forall d \in \mathbb{N}^*, \textit{with} \quad N_0 = 2 \tag{4.18}
$$

or, equivalently,

$$
N_d = 2 + \sum_{r=1}^{d} \sum_{k=0}^{[\frac{r+1}{2}]} p(r + 1 - 2k)p(k), \tag{4.19}
$$

where $p(n)$ denotes the number of partitions of n.

Proof We write

$$\frac{1}{C'(t)} = 1 + \sum_{k=0}^{d} \theta_k t^{k+1} = \prod_{k=0}^{d} (1 - v_k t) \tag{4.20}$$

- We can extend the number N_d for $d = 0$, by setting $N_0 = 2$, this is natural because if we return to Eqs. (4.1)–(4.2) with $d = 0$, we find two families of polynomials depending on whether $\theta_0 = 0$ or $\theta_0 \neq 0$.
  These two polynomials are, in fact, the only ones which satisfy the relation $P_{n+1}(x) = (x + \alpha_{n+1})P_n(x)$ with $\alpha_{n+1} \neq 0$.
  - If α_{n+1} is not depending on n (equal to α) then $P_n(x) = (x + \alpha)^n$ and $\Big(A(t), C(t)\Big) = \Big(e^{\alpha t}, t\Big)$.
  - Otherwise, we have $P_n(x) = \prod_{k=1}^{n}(x + \alpha_k)$. According to the proof of Theorem 4.1, $\alpha_{n+1} - \alpha_n = -\theta_0$ and $C'(t) = \frac{1}{1+\theta_0 t}$.
    It follows that $\Big(A(t), C(t)\Big) = \Big(\sum_{n=0}^{\infty} \frac{\prod_{k=1}^{n} \alpha_k}{n!} t^n, \frac{1}{\theta_0} \ln(1 + \theta_0 t)\Big)$.

  Next, we suppose that $d \geq 1$.
- If there exists $k \in \{0, 1, \ldots, d\}$ such that $v_k = 0$, we obtain the number N_{d-1} of $(d-1)$-OPSs.
- If $v_k \neq 0, \ \forall k \in \{0, 1, \ldots, d\}$, we count according to the real roots in (4.20).
  - If (4.20) has $d + 1$ real roots, then there are $p(d + 1)$ possibilities depending on multiplicities.
  - If (4.20) has exactly 2 conjugate complex roots, then there are $p(d+1-2)p(1)$ possibilities.
  - If (4.20) has exactly 4 two by two conjugate complex roots, then there are $p(d + 1 - 4)p(2)$.
  - Step by step, if (4.20) has exactly $2k$ two by two conjugate complex roots, then there are $p(d + 1 - 2k)p(k)$ possibilities, where $p(d + 1 - 2k)$ is the number of partitions of $(d + 1 - 2k)$ real roots and $p(k)$ is the number of partitions of k complex roots with their conjugates.

This gives (4.18).

By iteration, we deduce the overall number N_d and (4.19) follows. □

Examples

1. $N_1 = 5$ (Table 1 p. 235), $N_2 = 9$ (Table 2 p. 236), $N_3 = 18$ (Table 3 p. 238).
2. We will concretely treat the case $d = 4$:

$$\frac{1}{C'(t)} = 1+\sum_{k=0}^{4} \theta_k t^{k+1} = (1-v_0 t)(1-v_1 t)(1-v_2 t)(1-v_3 t)(1-v_4 t). \tag{4.21}$$

- If there exists $k \in \{0, 1, \ldots, 4\}$ such that $v_k = 0$, we obtain the number $N_3 = 18$ of 3-OPS (see Table 3).
- If $v_k \neq 0$, $\forall k \in \{0, 1, \ldots, 4\}$, we count according to the real roots in (4.21).
  - If (4.21) has 5 real roots, then there are $p(5) = 7$ possibilities:

$$\begin{gathered}(1 - v_0 t)^5;\\ (1 - v_0 t)^4(1 - v_1 t); \quad (1 - v_0 t)^3(1 - v_1 t)^2;\\ (1 - v_0 t)^3(1 - v_1 t)(1 - v_2 t); \quad (1 - v_0 t)^2(1 - v_1 t)^2(1 - v_2 t);\\ (1 - v_0 t)^2(1 - v_1 t)(1 - v_2 t)(1 - v_3 t);\\ (1 - v_0 t)(1 - v_1 t)(1 - v_2 t)(1 - v_3 t)(1 - v_4 t).\end{gathered}$$

  - If (4.21) has exactly 2 conjugate complex roots, then there are $p(5 - 2)p(1) = p(3)p(1) = 3$ possibilities:

$$\begin{gathered}(1 - v_0 t)^3(1 + at + bt^2), \quad a^2 - 4b < 0;\\ (1 - v_0 t)^2(1 - v_1 t)(1 + at + bt^2), \quad a^2 - 4b < 0;\\ (1 - v_0 t)(1 - v_1 t)(1 - v_2 t)(1 + at + bt^2), \quad a^2 - 4b < 0.\end{gathered}$$

  - If (4.21) has exactly 4 two by two conjugate complex roots, then there are $p(5 - 4)p(2) = p(1)p(2) = 2$ possibilities:

$$\begin{gathered}(1 - v_0 t)(1 + at + bt^2)^2, \quad a^2 - 4b < 0;\\ (1 - v_0 t)(1 + at + bt^2)(1 + a't + b't^2), \quad a^2 - 4b < 0, \quad a'^2 - 4b' < 0.\end{gathered}$$

Hence, we get $N_4 = N_3 + p(5) + p(5-2)p(1) + p(5-4)p(2) = 18+7+3+2 = 30$.

Computation of N_d with Maple The computation of N_d can be implemented using the package *combinat* of Maple system, and by virtue of Eq. (4.19):

$>$*with(combinat):*

$> p(n) := numbpart(n):$

$> N(d) := 2 + sum\left(sum\left(p(r + 1 - 2 \cdot k) \cdot p(k), k = 0..\text{floor}\left(\frac{r+1}{2}\right)\right), r = 1..d\right):$

$> seq([d, N(d)], d = 1..20);$

[1, 5], [2, 9], [3, 18], [4, 30], [5, 53], [6, 84], [7, 138], [8, 211], [9, 329], [10, 488], [11, 734], [12, 1063], [13, 1552], [14, 2203], [15, 3143], [16, 4385], [17, 6136], [18, 8434], [19, 11611], [20, 15753].

4.4 Classification of d-OPSs of Sheffer Type

4.4.1 Case $d = 1$

Both Meixner [25] and Sheffer [29] were interested in the same problem: what are all possible forms of PSs which are at the same time orthogonal and Sheffer polynomials?

The recurrence relation satisfied by an OPS $\{P_n\}_{n\geq 0}$ has the following form:

$$P_{n+1}(x) = (x + \alpha_{n+1})P_n(x) + \widehat{\beta}_1^{(n+1)} P_{n-1}(x). \tag{4.22}$$

From Theorem 4.1, the functions C and A verify the following equations:

$$C'(t) = \frac{1}{1 + \theta_0 t + \theta_1 t^2} \quad \text{and} \quad \frac{A'(t)}{A(t)} = \frac{\sigma_0 + \sigma_1 t}{1 + \theta_0 t + \theta_1 t^2}.$$

However the parameters in the recurrence relation (1.5) are

$$\alpha_{n+1} = -n\theta_0 + \sigma_0 \quad \text{and} \quad \widehat{\beta}_1^{(n+1)} = -n(n-1)\theta_1 + n\sigma_1; \quad \sigma_1 \neq 0, \quad \frac{\theta_1}{\sigma_1} \neq \frac{1}{k}, \ k \in \mathbb{N}^*.$$

If we discuss all possible cases in view of these two conditions, then we face with the known orthogonal PSs listed below:

- **(A):** $\alpha = \beta = 0$ $(\theta_0 = \theta_1 = 0)$.
- **(B):** $\alpha \neq 0$ and $\beta = 0$ $(\theta_0 = -\alpha,\ \theta_1 = 0)$.
- **(C):** $\alpha = \beta \neq 0$ $(\theta_1 \neq 0,\ \theta_0^2 = 4\theta_1)$.
- **(D$_1$):** $\alpha \neq \beta$, two non-zero real roots $(\theta_1 \neq 0,\ \theta_0^2 > 4\theta_1)$.
- **(D$_2$):** $\alpha \neq \beta$, two non-zero conjugate roots $(\theta_1 \neq 0,\ \theta_0^2 < 4\theta_1)$.

There exist exactly five Sheffer orthogonal families: Hermite polynomials, Laguerre polynomials, Charlier polynomials, Meixner polynomials (or Meixner I) and Meixner-Pollaczek polynomials (or Meixner II).

This class of polynomials is known in the literature as the *Meixner's Class*.

After affine transformations on x and t in the generating function $G(x, t)$, one obtains Table 1.

4.4.2 Case $d = 2$

In [11], Boukhemis investigated the 2-orthogonal polynomials of Sheffer type. General properties of the nine obtained families are given. Those which are classical were studied in greater detail. In some cases, an integral representation of the 2-dimensional vector assuring the 2-orthogonality was exhibited.

Table 1 OPSs of Sheffer type (Meixner [25])

Polynomial set	Generating function: $G(x,t)$
A: Hermite	$G(x,t) = e^{at^2}e^{xt};\ a = \frac{\sigma_1}{2} \neq 0,\ \sigma_0 = 0$
B: Charlier	$G\left(-\alpha x - \sigma_0 - \frac{\sigma_1}{\alpha}, \frac{t}{\alpha}\right) = e^{bt}(1+t)^x;\ b = \frac{\sigma_1}{\alpha^2} \neq 0$
C: Laguerre	$G\left(x, \frac{t}{\alpha}\right) = (1-t)^c e^{\frac{xt}{1-t}};\ c = \frac{\sigma_1}{\alpha^2} \neq 0,\ \sigma_0 = -\frac{\sigma_1}{\alpha}$
$\mathbf{D}_1$: Meixner I	$G\left((\beta-\alpha)(x-\sigma_0-\frac{\sigma_1}{\alpha}), \frac{t}{\beta}\right) = \left(1-\frac{t}{\eta}\right)^x (1-t)^{-x+d_1};$ $\mu = \frac{\beta}{\alpha} \neq 0,\ \lambda = \frac{\sigma_1}{\alpha\beta} \neq 0$
$\mathbf{D}_2$f: Meixner II	$G\left(\Im(\alpha)x - \frac{\sigma_0}{\Im(\alpha)}, \frac{t}{\Im(\alpha)}\right) = \left(\left(1+\delta t\right)^2 + t^2\right)^{d_2} e^{x\arctan\left(\frac{t}{1+\delta t}\right)};$ $\delta = -\frac{\Re(\alpha)}{\Im(\alpha)},\ d_2 = \frac{\sigma_1}{\lvert\alpha\rvert^2}$

Let $\{P_n\}_{n\geq 0}$ be a 2-OPS. Due to (1.5), the corresponding recurrence relation is:

$$P_{n+1}(x) = (x+\alpha_{n+1})P_n(x) + \widehat{\beta}_1^{(n+1)}P_{n-1}(x) + \widehat{\beta}_2^{(n+1)}P_{n-2}(x),\ n \in \mathbb{N}. \tag{4.23}$$

If $\{P_n\}_{n\geq 0}$ is also generated by (1.1), then according to Theorem 4.1, the functions C and A satisfy, respectively, the equations:

$$C'(t) = \frac{1}{(1-\alpha t)(1-\beta t)(1-\gamma t)} \quad \text{and} \quad \frac{A'(t)}{A(t)} = \frac{\sigma_0 + \sigma_1 t + \sigma_2 t^2}{(1-\alpha t)(1-\beta t)(1-\gamma t)}. \tag{4.24}$$

The coefficients of the recurrence relation are given by:
$\alpha_{n+1} = -n\theta_0 + \sigma_0$, $\widehat{\beta}_1^{(n+1)} = -n(n-1)\theta_1 + n\sigma_1$, and
$\widehat{\beta}_2^{(n+1)} = -n(n-1)(n-2)\theta_2 + n(n-1)\sigma_2$, with $\sigma_2 \neq 0$, $\frac{\theta_2}{\sigma_2} = -\frac{\alpha\beta\delta}{\sigma_2} \neq \frac{1}{k}$, $k \in \mathbb{N}^*$.

By considering all possibilities, we list below the different cases:

- **(A):** $\alpha = \beta = \gamma = 0$; $(\theta_0 = \theta_1 = \theta_2 = 0)$.
- **(B):** $\alpha \neq 0, \beta = \gamma = 0$; $(\theta_0 \neq 0, \theta_1 = \theta_2 = 0)$
- **(C):** $\alpha = \beta \neq 0$; $(\gamma = \theta_2 = 0)$.
- **(D):** $\alpha \neq \beta,\ \alpha \cdot \beta \neq 0, \gamma = 0$; $(\theta_0 \neq 0,\ \theta_1 \neq 0,\ \theta_2 = 0)$
  - **($\mathbf{D}_1$):** α and β are two real roots.
  - **($\mathbf{D}_2$):** α and β are two conjugate roots.
- **(E):** $\alpha = \beta = \gamma \neq 0$; $(\theta_2 \neq 0)$
- **(F):** $\alpha = \beta \neq \gamma$; $(\theta_2 \neq 0)$
- **(G):** $\alpha \neq \beta,\ \alpha \neq \gamma,\ \beta \neq \gamma$; $(\theta_2 \neq 0)$
  - **($\mathbf{G}_1$):** α, β, and γ are real roots.
  - **($\mathbf{G}_2$):** α, β are two conjugate roots and γ is a real root.

Table 2 2-OPS of Sheffer type

Case	Generating function: $G(x,t)$
A	$e^{a_1t^2+a_2t^3}\ e^{xt}$
B	$e^{b_1t+b_2t^2}\ (1+t)^x$
C	$(1-t)^{c_1}e^{c_2t}\ e^{\frac{xt}{1-t}}$
$\mathbf{D_1}$	$e^{d_1't}\left(1-\frac{t}{\eta}\right)^x(1-t)^{-x+d_1}$
$\mathbf{D_2}$	$\left(\left(1+\delta t\right)^2+t^2\right)^{d_2}\ e^{d_2't}\ e^{x\arctan\left(\frac{t}{1+\delta t}\right)}$
E	$(1-t)^{\varepsilon_1}\ e^{\varepsilon_2\frac{t}{(1-t)}}\ e^{x\frac{t(2-t)}{(1-t)^2}}$
F	$e^{(f_1x+f_2)\frac{t}{1-t}}\left(1-\frac{t}{\mu}\right)^x(1-t)^{-x+f_3}$
$\mathbf{G_1}$	$(1-t)^x\left(1-\frac{t}{\lambda_1}\right)^{g_1x+g_1'}\left(1-\frac{t}{\lambda_2}\right)^{g_2x+g_2'}$
$\mathbf{G_2}$	$\left(1-\frac{t}{\nu}\right)^{g_3x+g_3'}\left(\left(1+\delta t\right)^2+t^2\right)^{g_4x+g_4'}\ e^{x\arctan\left(\frac{t}{1+\delta t}\right)}$

After some affine transformations, the corresponding nine sequences are quoted in Table 2 (see page 236), among which we find 5 families that generalize, respectively, the Hermite polynomials, the Laguerre polynomials, the Charlier polynomials, the Meixner polynomials and the Meixner-Pollaczek polynomials.

4.4.3 Case $d = 3$

The case $d = 3$ was recently studied by Chaggara and Mbarki [16]. According to (1.5), any 3-OPS $\{P_n\}_{n\geq 0}$ satisfies a five-term recurrence relation as follows:

$$P_{n+1}(x) = (x+\alpha_{n+1})P_n(x) + \widehat{\beta}_1^{(n+1)}P_{n-1}(x) + \widehat{\beta}_2^{(n+1)}P_{n-2}(x) + \widehat{\beta}_3^{(n+1)}P_{n-3}(x), \tag{4.25}$$

where

$$\begin{cases} \alpha_{n+1} = -n\theta_0 + \sigma_0, \\ \widehat{\beta}_1^{(n+1)} = -n(n-1)\theta_1 + n\sigma_1, \\ \widehat{\beta}_2^{(n+1)} = -n(n-1)(n-2)\theta_2 + n(n-1)\sigma_2, \\ \widehat{\beta}_3^{(n+1)} = -n(n-1)(n-2)(n-3)\theta_3 + n(n-1)(n-2)\sigma_3. \end{cases}$$

with $\sigma_3 \neq 0$, $\frac{\theta_3}{\sigma_3} \neq \frac{1}{k}$, $k \in \mathbb{N}^*$. By Theorem 4.1, the relationships (4.1) and (4.2) are:

$$C'(t) = \frac{1}{1+\theta_0 t+\theta_1 t^2+\theta_2 t^3+\theta_3 t^4}, \quad \text{and} \quad \frac{A'(t)}{A(t)} = \frac{\sigma_0+\sigma_1 t+\sigma_2 t^2+\sigma_3 t^3}{1+\theta_0 t+\theta_1 t^2+\theta_2 t^3+\theta_3 t^4}. \tag{4.26}$$

By writing:

$$1+\theta_0 t+\theta_1 t^2+\theta_2 t^3+\theta_3 t^4=(1-\alpha t)(1-\beta t)(1-\gamma t)(1-\delta t);$$

with $\sigma_3 \neq 0$, $\frac{\theta_3}{\sigma_3}=\frac{\alpha\beta\gamma\delta}{\sigma_3}\neq\frac{1}{k}$, $k\in\mathbb{N}^*$, we get the following cases:

- **(A):** $\alpha=\beta=\gamma=\delta=0$; $(\theta_0=\theta_1=\theta_2=\theta_3=0)$.
- **(B):** $\alpha\neq 0, \beta=\gamma=\delta=0$; $(\theta_0\neq 0, \theta_1=\theta_2=\theta_3=0)$.
- **(C):** $\alpha=\beta\neq 0$; $(\gamma=\delta=\theta_2=\theta_3=0)$.
- **(D):** $\alpha\neq\beta$, $\alpha\cdot\beta\neq 0, \gamma=\delta=0$; $(\theta_0\neq 0, \theta_1\neq 0, \theta_2=\theta_3=0)$.
  - **(D$_1$):** α and β are two real roots.
  - **(D$_2$):** α and β are two conjugate roots.
- **(E):** $\alpha=\beta=\gamma\neq 0$; $(\delta=\theta_3=0)$
- **(F):** $\alpha=\beta\neq\gamma$; $(\theta_2\neq 0,\ \delta=\theta_3=0)$
- **(G):** $\alpha\neq\beta$, $\alpha\neq\gamma$, $\gamma\neq\beta$; $(\theta_2\neq 0,\ \delta=\theta_3=0)$.
  - **(G$_1$):** α, β, and γ are real roots.
  - **(G$_2$):** α, β are two conjugate roots and γ is real root.
- **(H):** $\alpha=\beta=\gamma=\delta\neq 0$; $(\theta_3\neq 0)$.
- **(I):** $\alpha=\beta=\gamma\neq\delta$; $(\theta_3\neq 0)$.
- **(J):** $\alpha=\beta\neq\gamma$, $\gamma\neq\delta$, $\beta\neq\delta$; $(\theta_3\neq 0)$.
  - **(J$_1$):** γ and δ are two real roots.
  - **(J$_2$):** γ and δ are two conjugate roots.
- **(K):** $\alpha=\beta\neq\gamma=\delta$; $(\theta_3\neq 0)$.
  - **(K$_1$):** α, β, γ and δ are real roots.
  - **(K$_2$):** α and γ are two conjugate roots.
- **(L):** α, β, γ, δ pairwise distinct, $(\theta_3\neq 0)$
  - **(L$_1$):** α, β, γ and δ are real roots.
  - **(L$_2$):** α, β are two real roots and γ, δ are two conjugate roots.
  - **(L$_3$):** α, β, γ and δ are conjugate roots in two.

We have exactly 18 Sheffer type 3-OPSs.

After affine transformations similar to the changes made in Tables 1 and 2, one obtains Table 3 (see page 238).

Note here that some slight modifications have been made in Tables 1, 2, and 3 in the expressions of the obtained generating functions in order to simplify the given forms. So, in some cases, new constants, depending of course on $\sigma_i, \theta_i, \ldots$, were introduced. In some cases, the considered families are not monic.

Table 3 3-OPS of Sheffer type (Chaggara and Mbarki [16])

Case	Generating function: $G(x,t)$
A	$e^{a_1t^2+a_2t^3+a_3t^4}\,e^{xt}$
B	$e^{b_1t+b_2t^2+b_3t^3}\,(1+t)^x$
C	$(1-t)^{c_1}\,e^{c_2t^2+c_3t}\,e^{\frac{xt}{1-t}}$
D₁	$e^{d_1't+d_1''t^2}\left(1-\frac{t}{\eta}\right)^x(1-t)^{-x+d_1}$
D₂	$\left(\left(1+\delta t\right)^2+t^2\right)^{d_2}\,e^{d_2't+d_2''t^2}\,e^{x\arctan\left(\frac{t}{1+\delta t}\right)}$
E	$e^{\varepsilon_0 t}\,(1-t)^{\varepsilon_1}\,e^{\varepsilon_2\frac{t}{(1-t)}}\,e^{x\frac{t(2-t)}{(1-t)^2}}$
F	$e^{f_0t}e^{\left(f_1x+f_2\right)\frac{t}{1-t}}\left(1-\frac{t}{\mu}\right)^x\left(1-t\right)^{-x+f_3}$
G₁	$e^{g_0t}\,(1-t)^x\left(1-\frac{t}{\lambda_1}\right)^{g_1x+g_1'}\left(1-\frac{t}{\lambda_2}\right)^{g_2x+g_2'}$
G₂	$e^{g_0't}\left(1-\frac{t}{\nu}\right)^{g_3x+g_3'}\left(\left(1+\delta t\right)^2+t^2\right)^{g_4x+g_4'}\,e^{x\arctan\left(\frac{t}{1+\delta t}\right)}$
H	$(1-t)^{h_1}\,e^{h_2\frac{t}{1-t}}\,e^{h_3\frac{t(2-t)}{(1-t)^2}}\,e^{x\frac{t\left(3-3t+t^2\right)}{(1-t)^3}}$
I	$e^{\left(i_1x+i_1'\right)\frac{t}{1-t}}\,e^{\left(i_2x+i_2'\right)\frac{t(2-t)}{(1-t)^2}}\,(1-\frac{t}{\upsilon})^x\left(1-t\right)^{x+i_3}$
J₁	$(1-t)^x\,(1-\frac{t}{\omega_1})^{j_1x+j_1'}\,(1-\frac{t}{\omega_2})^{j_2x+j_2'}\,e^{x\frac{t}{1-t}}$
J₂	$\left(1-\frac{t}{\varsigma}\right)^{j_3x+j_3'}\left(\left(1+\delta' t\right)^2+t^2\right)^{j_4x+j_4'}\,e^{\left(j_5x+j_5'\right)\frac{t}{1-\frac{t}{\varsigma}}}\,e^{x\arctan\left(\frac{t}{1+\delta' t}\right)}$
K₁	$(1-\frac{t}{\kappa})^x\left(1-t\right)^{-x+k_1}\,e^{\left(k_2x+k_2'\right)\frac{t}{1-t}}\,e^{\left(k_3x+k_3'\right)\frac{t}{1-\frac{t}{\kappa}}}$
K₂	$\left(\left(1+\delta t\right)^2+t^2\right)^{k_4}\,e^{\frac{k_5t+k_5'}{\left((1+\delta t)^2+t^2\right)}t}\,e^{\left(k_6x+k_6'\right)\left[\frac{k_7t+k_7'}{\left((1+\delta t)^2+t^2\right)}t\right]}\,e^{x\arctan\left(\frac{t}{1+\delta t}\right)}$
L₁	$\left(1-t\right)^x\left(1-\frac{t}{\eta_1}\right)^{l_1x+l_1'}\left(1-\frac{t}{\eta_2}\right)^{l_2x+l_2'}\left(1-\frac{t}{\eta_3}\right)^{l_3x+l_3'}$
L₂	$\left(1-\frac{t}{\tau_1}\right)^{l_4x+l_4'}\left(1-\frac{t}{\tau_2}\right)^{l_5x+l_5'}\left((1+\delta' t)^2+t^2\right)^{l_6x+l_6'}\,e^{x\arctan\left(\frac{t}{1+\delta' t}\right)}$
L₃	$\left(\left(1+\delta t\right)^2+t^2\right)^{l_1''x+l_2''}\,e^{x\arctan\left(\frac{t}{1+\delta t}\right)}\left(rt^2+st+1\right)^{l_3''x+l_4''}\,e^{\left(l_5''x+l_6''\right)\arctan\left(\frac{t}{r't+s'}\right)}$

Differential and Difference Equations By the expression of λ and ρ given in (3.1) and Theorem 4.1, it is obvious that if $\{P_n\}_{n\geq 0}$ is generated by (1.1), then we have:

$$\rho\lambda = \frac{A'}{A}(C^{-1}(D))C^{-1}(D) + xC'(C^{-1}(D))C^{-1}(D)\sum_{k=0}^{d}\sigma_k\frac{(C^{-1}(D))^{k+1}}{(C^{-1})'(D)} + x\frac{C^{-1}(D)}{(C^{-1})'(D)}. \tag{4.27}$$

According to Relation (2.3) and Equation (4.27), any Sheffer d-OPS satisfies the equation:

$$\Big(\sum_{k=0}^{d}\sigma_k(C^{-1}(D))^{k+1}+xC^{-1}(D)+\frac{C^{-1}(D)(C^{-1})''(D)}{(C^{-1})'(D)}\Big)P_n=n(C^{-1})'(D)P_n. \tag{4.28}$$

Next, we consider some particular cases among Table 3.

- Case [**A**]: For this case, we have:

$$C(t)=t \text{ and } \rho\lambda=\sigma_3D^4+\sigma_2D^3+\sigma_1D^2+\sigma_0D+xD.$$

  By virtue of (4.28), we obtain the forth-order differential equation:

$$\sigma_3P_n^{(4)}(x)+\sigma_2P_n^{(3)}(x)+\sigma_1P_n''(x)+(x+\sigma_0)P_n'(x)-nP_n(x)=0. \tag{4.29}$$

- Case [**B**]: In that event $C(t)=\ln(1+t)$ then $C^{-1}(D)=e^D-1=\Delta$. Hence,

$$\rho\lambda=\sum_{k=0}^{3}\sigma_k\frac{\Delta^{k+1}}{\Delta+1}+x\frac{\Delta}{\Delta+1}.$$

  We use (4.28) to get the forth-order finite difference equation,

$$\sigma_3\Delta^4P_n(x)+\sigma_2\Delta^3P_n(x)+\sigma_1\Delta^2P_n(x)+[x+\sigma_0-(n-1)]\Delta P_n(x)-nP_n(x)=0. \tag{4.30}$$

- Case [**C**]: In this case $C(t)=\frac{t}{1-t}$ and $\lambda=\frac{D}{1+D}$. The Eq. (4.28) reduces to

$$\rho\lambda=\sum_{k=0}^{3}\sigma_kD^{k+1}(1+D)^{1-k}+xD(1+D).$$

  Thus, we obtain a forth-order differential equation verified by the family (**C**).

$$\begin{aligned}&[\sigma_3+\sigma_2+\sigma_1+(x+\sigma_0)]P_n^{(4)}(x)+[\sigma_2+2\sigma_1+3(x+\sigma_0)+2]P_n^{(3)}(x)\\&+[\sigma_1+3(x+\sigma_0)+(4-n)]P_n''(x)+[x+\sigma_0+2(1-n)]P_n'(x)-nP_n(x)=0.\end{aligned} \tag{4.31}$$

- Case [$\mathbf{D_1}$]: In that event $C(t)=\ln\Big(\frac{1-\frac{t}{\eta}}{1-t}\Big)$ then $\lambda=\frac{\Delta}{\Delta+\delta}$, $\delta=\frac{\eta-1}{\eta}$, which gives, by virtue of (4.27),

$$\rho\lambda=\sum_{k=0}^{3}\sigma_k\frac{\Delta^{k+1}(\Delta+\delta)^{1-k}}{\delta(\Delta+1)}+x\frac{\Delta(\Delta+\delta)}{\delta(\Delta+1)}.$$

We find that any element of the sequence ($\mathbf{D}_1$) in Table 3 satisfies the forth-order finite difference equation:

$$\begin{aligned}
&\frac{1}{\delta^3}\Big[\sigma_3+\sigma_2+\sigma_1+(x+\sigma_0)+3\Big]\Delta^4 P_n(x)\\
&+\frac{1}{\delta^2}\Big[\sigma_2+2\sigma_1+3(x+\sigma_0)+\Big(7-n-\frac{2}{\delta}\Big)\Big]\Delta^3 P_n(x)\\
&+\frac{1}{\delta}\Big[\sigma_1+3(x+\sigma_0)+\Big(5-2n-\frac{4+n}{\delta}\Big)\Big]\Delta^2 P_n(x)\\
&+\Big[(x+\sigma_0)+\Big(1-n-\frac{1+2n}{\delta}\Big)\Big]\Delta P_n(x)-nP_n(x)=0. \qquad (4.32)
\end{aligned}$$

4.4.4 $(d+1)$-Fold Symmetric d-OPS of Sheffer Type

In the sequel, we will be interested by d-SOPSs families among the Sheffer type PSs. We discuss the corresponding form of the generating function according to the sign of θ_d given by (4.12).

From (4.13), we conclude that:

$$\begin{cases} A(t)=\exp\Big(\frac{\sigma_d}{d+1}t^{d+1}\Big) & \text{if } \theta_d=0,\\[2ex] A(t)=\Big(1+\theta_d t^{d+1}\Big)^{\frac{\sigma_d}{(d+1)\theta_d}} & \text{if } \theta_d\neq 0. \end{cases} \qquad (4.33)$$

Firstly, we consider the cases $d=1$ and $d=2$ and then we study the general case. Without loss of generality, we will discuss the cases $\theta_d=0,\quad \theta_d=1$ and $\theta_d=-1$.

Case $d=1$ We have in this case, $C'(t)=\frac{1}{1+\theta_1 t^2}$.

- $\theta_1=0$: In this case, $C(t)=t$, and then $\{P_n\}_{n\geq 0}$ reduces to the Hermite PS which corresponds to Case (**A**) in Table 1.
- $\theta_1=-1$: We have

$$C(t)=\frac{1}{2}\ln\Big(\frac{1+t}{1-t}\Big).$$

This solution corresponds to Meixner PS is Case ($\mathbf{D_1}$) in Table 1 (with $\eta=-1$).
- $\theta_1=1$: $C'(t)=\frac{1}{1+t^2}$, then

$$C(t)=\arctan(t).$$

The obtained sequence corresponding to Meixner PS in Case ($\mathbf{D_2}$) in Table 1.

Case $d = 2$ we have $C'(t) = \frac{1}{1+\theta_2 t^3}$.

- $\theta_2 = 0$: This is similar to the case $\theta_1 = 0$ when $d = 1$, $\{P_n\}_{n\geq 0}$ corresponding to 2-Hermite PS (case (**A**) in Table 2.
- $\theta_2 = \pm 1$: $1 \pm t^3$ admits one real root ± 1 and two conjugate complex roots $\pm j,\ \pm\overline{j}$. Thus $C(t)$ has the following form:

$$C(t) = \frac{\pm 1}{3}\left[\ln\Big(1 \pm t\Big) - \frac{1}{2}\ln\Big(t^2 \mp t + 1\Big) \pm \sqrt{3}\arctan\Big(\frac{\sqrt{3}t}{t \mp 2}\Big)\right],$$

  which corresponds to case ($\mathbf{G_2}$) in Table 2.

General Case This classification depends on the parity of d and the values of θ_d.

Case 1**:** d **is Odd** ($d \geq 3$)

- If $\theta_d = 0$, then $C(t) = t$, $\{P_n\}_{n\geq 0}$ is the Gould–Hopper PS [21] or Hermite-type d-SOPS already mentioned in Sect. 4.

$$\sum_{n=0}^{\infty} \frac{P_n(x)}{n!} t^n = \exp\Big(\frac{\sigma_d}{d+1} t^{d+1}\Big)\exp(xt).$$

- If $\theta_d = -1$, then $C'(t) = \frac{1}{1-t^{d+1}}$.
  The polynomial $1 - t^{d+1}$ admits two real roots: ± 1 and $(d-1)$ complex two by two conjugate roots $\xi_1, \ldots, \xi_{\frac{d-1}{2}}$ and $\overline{\xi}_1, \ldots, \overline{\xi}_{\frac{d-1}{2}}$, $\xi_k = e^{\frac{2ik\pi}{d+1}}$.
  It follows:

$$C'(t) = \frac{1}{d+1}\Big(\frac{1}{1+t} + \frac{1}{1-t}\Big) + \sum_{k=1}^{\frac{d-1}{2}} \frac{\alpha_k t + \beta_k}{t^2 - 2\cos\Big(\frac{2k\pi}{d+1}\Big)t + 1}.$$

  Then $C(t)$ takes the form:

$$C(t) = \frac{1}{d+1}\ln\Big(\frac{1+t}{1-t}\Big) + \sum_{k=1}^{\frac{d-1}{2}} r_k \ln\Big(t^2 - 2\cos\Big(\frac{2k\pi}{d+1}\Big)t + 1\Big) + \sum_{k=1}^{\frac{d-1}{2}} s_k \arctan\Big(\frac{a_k t}{t + b_k}\Big).$$

  So, the generating function of $\{P_n\}_{n\geq 0}$ becomes:

$$\Big(1 - t^{d+1}\Big)^{\frac{-\sigma_d}{d+1}}\Big(\frac{1+t}{1-t}\Big)^{\frac{x}{d+1}} \prod_{k=1}^{\frac{d-1}{2}}\Big(t^2 - 2\cos\Big(\frac{2k\pi}{d+1}\Big)t + 1\Big)^{r_k x} \prod_{k=1}^{\frac{d-1}{2}} e^{s_k x \arctan\Big(\frac{a_k t}{t+b_k}\Big)}.$$

- If $\theta_d = 1$: Similarly to the above case,

$$1 + t^{d+1} = \prod_{k=0}^{d} \Big(t - \zeta_k\Big); \quad \zeta_k = e^{\frac{(2k+1)i\pi}{d+1}}.$$

  Then $1 + t^{d+1}$ admits $(d+1)$ two by two conjugate roots: $\zeta_0, \dots, \zeta_{\frac{d-1}{2}}, \overline{\zeta}_0, \dots, \overline{\zeta}_{\frac{d-1}{2}}$.
  Thus,

$$C(t) = \sum_{k=0}^{\frac{d-1}{2}} r_k \ln\Big(t^2 - 2\cos\Big(\frac{(2k+1)\pi}{d+1}\Big)t + 1\Big) + \sum_{k=0}^{\frac{d-1}{2}} s_k \arctan\Big(\frac{a_k t}{t + b_k}\Big).$$

  Hence

$$\sum_{n=0}^{\infty} \frac{P_n(x)}{n!} t^n = \Big(1 + t^{d+1}\Big)^{\frac{\sigma_d}{d+1}} \prod_{k=0}^{\frac{d-1}{2}} \Big(t^2 - 2\cos\Big(\frac{(2k+1)\pi}{d+1}\Big)t + 1\Big)^{r_k x} \prod_{k=0}^{\frac{d-1}{2}} e^{s_k x \arctan\Big(\frac{a_k t}{t + b_k}\Big)}.$$

Case 2: d Is Even

- If $\theta_d = 0$, similar to the Case I (d is odd).
- If $\theta_d \neq 0$ for both cases ($\theta_d = 1$ and $\theta_d = -1$), $1 \pm t^{d+1}$ has a single real root ± 1 and d two by two conjugate complex roots. Then, for $\theta_d = 1$, $C(t)$ is given by:

$$C(t) = \frac{1}{d+1} \ln(1+t) + \sum_{k=1}^{\frac{d}{2}} r_k \ln\Big(t^2 - 2\cos\Big(\frac{(2k-1)\pi}{d+1}\Big)t + 1\Big) + \sum_{k=1}^{\frac{d}{2}} s_k \arctan\Big(\frac{a_k t}{t + b_k}\Big).$$

  Thus $\{P_n\}_{n \geq 0}$ is generated by:

$$\Big(1 + t^{d+1}\Big)^{\frac{\sigma_d}{d+1}} (1+t)^{\frac{x}{d+1}} \prod_{k=1}^{\frac{d}{2}} \Big(t^2 - 2\cos\Big(\frac{(2k-1)\pi}{d+1}\Big)t + 1\Big)^{r_k x} \prod_{k=1}^{\frac{d}{2}} e^{s_k x \arctan\Big(\frac{a_k t}{t + b_k}\Big)}.$$

From the above inquiry, we conclude that there are exactly three families of d-SOPSs of Sheffer type if d is odd and exactly two in the even case (Table 4).

Table 4 d-SOPS of Sheffer type (Chaggara and Mbarki [16])

d	θ_d	C(t)	A(t)
Odd	$\theta_d = 0$	t	$\exp\left(\frac{\sigma_d}{d+1}t^{d+1}\right);\ \sigma_d \neq 0$
	$\theta_d = -1$	$\frac{1}{d+1}\ln\left(\frac{1+t}{1-t}\right)$ $+\sum_{k=1}^{\frac{d-1}{2}} r_k \ln\left(t^2 - 2\cos\left(\frac{2k\pi}{d+1}\right)t + 1\right)$ $+\sum_{k=1}^{\frac{d-1}{2}} s_k \arctan\left(\frac{a_k t}{t+b_k}\right)$	$\left(1 - t^{d+1}\right)^{\frac{-\sigma_d}{d+1}};\ -\sigma_d \notin \mathbb{N}$
	$\theta_d = 1$	$\sum_{k=0}^{\frac{d-1}{2}} r_k \ln\left(t^2 - 2\cos\left(\frac{(2k+1)\pi}{d+1}\right)t + 1\right)$ $+\sum_{k=0}^{\frac{d-1}{2}} s_k \arctan\left(\frac{a_k t}{t+b_k}\right)$	$\left(1 + t^{d+1}\right)^{\frac{\sigma_d}{d+1}};\ \sigma_d \notin \mathbb{N}$
Even	$\theta_d = 0$	t	$\exp\left(\frac{\sigma_d}{d+1}t^{d+1}\right);\ \sigma_d \neq 0$
	$\theta_d = \pm 1$	$\frac{\pm 1}{d+1}\ln(1+t)$ $+\sum_{k=1}^{\frac{d}{2}} r_k \ln\left(t^2 - 2\cos\left(\frac{\left(2k-\left(\frac{1\pm 1}{2}\right)\right)\pi}{d+1}\right)t + 1\right)$ $+\sum_{k=1}^{\frac{d}{2}} s_k \arctan\left(\frac{a_k t}{t+b_k}\right)$	$\left(1 + t^{d+1}\right)^{\frac{\sigma_d}{(d+1)}};\ \sigma_d \notin \mathbb{N}$

Acknowledgements We are much indebted to the organisers, Wolfram Koepf and Mama Foupouagnigni, of the *AIMS-Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications*, Douala, Cameroon, October 5–12, 2018 for the invitation, the interesting program and the kind hospitality.

The first author would like to extend his appreciation to the Deanship of Scientific Research at King Khalid University for funding his work through research groups program under grant (2019).

References

1. A.I. Aptekarev, Multiple polynomials. J. Comput. Appl. Math. **99**, 423–447 (1998)
2. A.I. Aptekarev, V.A. Kaliaguine, W. Van Assche, Criterion for the resolvent set of non-symmetric tridiagonal operators. Proc. Amer. Math. Soc. **123**, 2423–2430 (1995)
3. Y. Ben Cheikh, On obtaining dual sequences via quasi-monomiality. Georgian Math. J. **9**, 413–422 (2002)
4. Y. Ben Cheikh, Some results on quasi-monomiality. Appl. Math. Comput. **141**, 63–76 (2003)
5. Y. Ben Cheikh, N. Ben Romdhane, d-symmetric d-orthogonal polynomials of Brenke type. J. Math. Anal. Appl. **416**, 735–747 (2014)
6. Y. Ben Cheikh, M. Gaied, Dunkl-Appell d-orthogonal polynomials. Integral Transform. Spec. Funct. **28**, 581–597 (2007)
7. Y. Ben Cheikh, I. Gam, On some operators varying the dimensional parameters of d-orthogonality. Integral Transform. Spec. Funct. **27**, 731–746 (2016)
8. Y. Ben Cheikh, A. Ouni, Some generalized hypergeometric d-orthogonal polynomial sets. J. Math. Anal. Appl. **343**, 464–478 (2008)
9. Y. Ben Cheikh, A. Zaghouani, Some discrete d-orthogonal polynomial sets. J. Comput. Appl. Math. **156**, 253–263 (2003)

10. Y. Ben Cheikh, A. Zaghouani, d-orthogonality via generating functions. J. Comput. Appl. Math. **199**, 2–22 (2007)
11. A. Boukhemis, On the classical 2-orthogonal polynomials sequences of Sheffer-Meixner type. CUBO A Math. J. Univ. Frontera **7**, 39–55 (2005)
12. A. Boukhemis, P. Maroni, Une caractérisation des polynômes strictement $1/p$ orthogonaux de type Sheffer. Étude du cas p = 2. J. Math. Approx. Theory. **54**, 67–91 (1998)
13. E. Bourreau, Polynômes orthogonaux simultanés et systèmes dynamiques infinis. Ph.D. Thesis Université des Sciences et Technologie de Lille-Lille I. France (2002)
14. M.G. Bruin, Simultaneous Padé approximation and orthogonality, in *Polynômes Orthogonaux et Applications*, ed. by C. Brezinski, A. Draux, A.P. Magnus, P. Maroni, A. Ronveaux. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 74–83
15. H. Chaggara, Quasi monomiality and linearization cofficients for Sheffer polynomial sets, in Proceedings of the International Conference: Difference Equations, Special Functions and Orthogonal Polynomials Munich, Germany, 2005. ed. by S. Elaydi, J. Cushing, R. Lasser, V. Papageorgiou, A. Ruffing, W. Van Assche. (Special edition of World Scientific, Singapore, 2007), pp. 90–99
16. H. Chaggara, R. Mbarki, On d-orthogonal polynomials of Sheffer type. J. Diff. Equ. Appl. **24**, 1808–1829 (2018)
17. E. Coussement, W. Van Assche, Multiple orthogonal polynomials associated with the modified Bessel functions of the first kind. Constr. Approx. **19**, 237–263 (2003)
18. K. Douak, The relation of the d-orthogonal polynomials to the Appell polynomials. J. Comput. Appl. Math. **70**, 279–295 (1996)
19. K. Douak, On 2-orthogonal polynomials of Laguerre type. Inter. J. Math. Math. Sci. **22**, 29–48 (1999)
20. K. Douak, P. Maroni, Les polynômes orthogonaux classiques de dimension deux. Analysis **12**, 71–107 (1992)
21. H.W. Gould, A.T. Hopper, Operational formulas connected with two generalizations of Hermite polynomials. Duke Math. J. **29**, 51–63 (1962)
22. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable* (Cambridge University Press, Cambridge, 2005)
23. I. Kubo, Generating functions of exponential type for orthogonal polynomials. Infin. Dimens. Anal. Quantum Probab. Relat. Top. **7**, 155–159 (2004)
24. P. Maroni, L'orthogonalité et les récurrences de polynômes d'ordre supérieur à deux. Ann. Fac. Sci. Toulouse Math. **10**, 105–139 (1998)
25. J. Meixner, Orthogonale polynomsysteme mit einer besonderen Gestalt der erzeugenden Funktion. J. Lond. Math. Soc. **1**, 6–13 (1934)
26. E.D. Rainville, *Special Functions* (The Macmillan Company, New York, 1960)
27. S. Roman, G.C. Rota, The Umbral Calculus. Adv. Math. **27**, 95–188 (1978)
28. A. Saib, On semi-classical d-orthogonal polynomials. Math. Nachr. **286**, 1863–1885 (2013)
29. I.M. Sheffer, Some properties of polynomial sets of type zero. Duke Math. J. **5**, 590–622 (1939)
30. J. Van Iseghem, Approximants de Padé vectoriels. Ph.D. Thesis, Université des Sciences et Techniques de Lille-Flandre-Artois (1987)
31. J. Van Iseghem, Vector orthogonal relations. Vector QD-algorithm. J. Comput. Appl. Math. **19**, 141–150 (1987)
32. S. Varma, Some new d-orthogonal polynomial sets of Sheffer type. Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat. **68**, 1–10 (2018)

From Standard Orthogonal Polynomials to Sobolev Orthogonal Polynomials: The Role of Semiclassical Linear Functionals

Juan C. García-Ardila, Francisco Marcellán, and Misael E. Marriaga

Abstract In this contribution, we present an overview of standard orthogonal polynomials by using an algebraic approach. Discrete Darboux transformations of Jacobi matrices are studied. Next, we emphasize the role of semiclassical orthogonal polynomials as a basic background to analyze sequences of polynomials orthogonal with respect to a Sobolev inner product. Some illustrative examples are discussed. Finally, we summarize some results in multivariate Sobolev orthogonal polynomials.

Keywords Orthogonal polynomials · Discrete Darboux transformations · Semi-classical functionals · Sobolev orthogonal polynomials

Mathematics Subject Classification (2000) Primary 42C05; Secondary 33C45, 33D50

J. C. García-Ardila
Departamento de Matemática Aplicada a la Ingeniería Industrial, Universidad Politécnica de Madrid, Madrid, Spain
e-mail: juancarlos.garciaa@upm.es

F. Marcellán (✉)
Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain

Instituto de Ciencias Matemáticas (ICMAT), Madrid, Spain
e-mail: pacomarc@ing.uc3m.es

M. E. Marriaga
Departamento de Matemática Aplicada, Ciencia e Ingeniería de Materiales y Tecnología Electrónica, Universidad Rey Juan Carlos, Móstoles, Spain
e-mail: misael.marriaga@urjc.es

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_13

1 Introduction

The aim of this contribution is to provide a self-contained presentation of the so called Sobolev orthogonal polynomials, i.e., polynomials which are orthogonal with respect to a bilinear form involving derivatives of its inputs, defined in the linear space of polynomials with real coefficients. We start by focusing our attention on an algebraic approach to the so called standard orthogonal polynomials, which are polynomials orthogonal with respect to a linear functional, taking into account that we can associate with such polynomials a structured matrix for their moments (a Hankel matrix), a tridiagonal matrix (a Jacobi matrix reflecting the fact that the multiplication operator is symmetric with respect to the above linear functional), as well as an analytic function around infinity (the so called Stieltjes function, that is the z-transform of the sequence of moments of the linear functional), such that the denominators of the diagonal Padé approximants to such a function are the corresponding orthogonal polynomials (we refer the reader to [17] and [70] for an introduction to these topics). These three basic ingredients allow us to deal with a theory that knows an increasing interest in the last decades (see [64], [37]).

The most useful standard orthogonal polynomials appear as polynomial eigenfunctions of second-order differential operators and constitute the so called classical families—Hermite, Laguerre, Jacobi and Bessel—see [42] as well as [11]. All of them can be written in terms of hypergeometric functions and they can be characterized in several ways taking into account their hypergeometric character. Beyond the above classical orthogonal polynomials, the so-called semiclassical orthogonal polynomials constitute a wide class with an increasing interest for researchers, taking into account their connections with Painlevé equations and integrable systems [73]. They have been introduced in [69] from the point of view of holonomic equations satisfied by orthogonal polynomials associated with weight functions $w(x)$ satisfying a Pearson differential equation $(A(x)w(x))' = B(x)w(x)$, where A and B are polynomials. In the 80s, they have been intensively studied by P. Maroni and co-workers (see [59] as an excellent and stimulating survey paper). The role of semiclassical orthogonal polynomials in the study of orthogonal polynomials with respect to univariate Sobolev inner products has been emphasized when the so called coherent pairs of measures are introduced (see [63]) as well as some of their generalizations (see [22]).

The structure of the paper is the following. In Sect. 2, a basic background concerning linear functionals and the algebraic structure of the topological dual space corresponding to the linear space of polynomials with real coefficients is presented. Orthogonal polynomials with respect to linear functionals are defined and the three-term recurrence relation they satisfy constitute a key point in the analysis of their zeros. Discrete Darboux transformations for linear functionals are studied in Sect. 3 in the framework of LU and UL factorizations of Jacobi matrices (see [14]). The connection formulas between the corresponding sequences of orthogonal polynomials are studied in the framework of the linear spectral transformations of the Stieltjes functions associated with linear functionals. In Sect. 4, following [59], semiclassical linear functionals are introduced and some of their characterizations are provided. The definition of class plays a key role in order to give a classification of semiclassical orthogonal polynomials, mainly those of class zero (the classical ones) and of class one (see [10]), which will play a central role in the sequel. Thus, a constructive approach to describe families of semiclassical linear functionals is presented. In particular, every linear spectral transformation of a semiclassical linear functional is also semiclassical. On the other hand, the symmetrization process of linear functionals is also studied and the invariance of the semiclassical character of a linear functional by symmetrization is pointed out. This constitutes the core of Sect. 5. In Sect. 6, orthogonal polynomials with respect to Sobolev inner products associated with a vector of measures supported on the real line are introduced. We emphasize the case where this vector of measures is coherent, i.e., their corresponding sequences of orthogonal polynomials satisfy a simple algebraic relation. This fact allows to deal with an algorithm to generate Sobolev orthogonal polynomials associated with coherent pairs of measures. Some analytic properties of these polynomials are shown. Notice that the three-term recurrence relation that constitutes a basic tool for the standard orthogonality is lost and, as a consequence, new techniques for studying asymptotic properties of such orthogonal polynomials are needed. In Sect. 7, multivariate Sobolev orthogonal polynomials are studied and their representations in terms of semiclassical orthogonal polynomials and spherical harmonics are given. A recent survey on Sobolev orthogonal polynomials can be found in [56], both in the univariate and multivariate case. Finally, an updated list of references provides a good guideline for the readers interested in these topics.

2 Background

Recall that a linear functional $\mathbf{u}$ defined on the linear space $\mathbb{P}$ of polynomials with real coefficients is a mapping

$$\begin{aligned} \mathbf{u} :& \mathbb{P} \to \mathbb{R} \\ & p \to \langle \mathbf{u}, p \rangle \end{aligned}$$

such that for every polynomials p, q, and every real numbers α, β,

$$\langle \mathbf{u}, \alpha p + \beta q \rangle = \alpha \langle \mathbf{u}, p \rangle + \beta \langle \mathbf{u}, q \rangle.$$

In general, given a basis of polynomials $\{p_n(x)\}_{n\geq 0}$, and a sequence of real numbers $\{\mu_n\}_{n\geq 0}$, a linear functional $\mathbf{u}$ is defined by means of its action on the basis

$$\langle \mathbf{u}, p_n \rangle = \mu_n, \quad n \geq 0,$$

and extended by linearity to all polynomials. If $p_n(x) = x^n, n \geq 0$, then the real numbers $\mu_n = \langle \mathbf{u}, x^n \rangle, n \geq 0$, are called *moments* with respect to the canonical basis and we usually say that $\mathbf{u}$ is a *moment functional*. If $p_n(x) = a_n x^n +$ lower degree terms, $n \geq 0,, a_n \neq 0$, the real numbers $\tilde{\mu}_n = \langle \mathbf{u}, p_n \rangle, n \geq 0$, are called the *modified moments* associated with the linear functional $\mathbf{u}$.

For a linear functional $\mathbf{u}$, we define its moment matrix as the semi-infinite Hankel matrix $M = (\mu_{i+j})_{i,j=0}^{\infty}$. If we denote

$$\Delta_n = \det[(\mu_j)_{j=0}^n] = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix},$$

then $\mathbf{u}$ is said to be quasi-definite if $\Delta_n \neq 0$ for $n \geq 0$, and $\mathbf{u}$ is said to be positive-definite if $\Delta_n > 0$ for $n \geq 0$.

Definition 2.1 Given a linear functional $\mathbf{u}$ and a polynomial $q(x)$ we define a new linear functional $q(x)\mathbf{u}$ as

$$\langle q(x)\mathbf{u}, p \rangle = \langle \mathbf{u}, q(x)\, p(x) \rangle,$$

for every polynomial $p \in \mathbb{P}$.

Definition 2.2 Given a linear functional $\mathbf{u}$ and a polynomial $p(x) = \sum_{k=0}^{n} a_k\, x^k$, we define the polynomial $(\mathbf{u} * p)(x)$ as

$$(\mathbf{u} * p)(x) := \left\langle \mathbf{u}_y, \frac{xp(x) - yp(y)}{x - y} \right\rangle = \sum_{k=0}^{n} \left(\sum_{m=k}^{n} a_m\, \mu_{m-k} \right) x^k$$
$$= (1, x, \ldots, x^n) \begin{pmatrix} \mu_0 & \ldots & \mu_n \\ & \ddots & \vdots \\ & & \mu_0 \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}.$$

Definition 2.3 A sequence of polynomials $\{P_n(x)\}_{n \geq 0}$ is said to be a sequence of orthogonal polynomials with respect to $\mathbf{u}$ if

(i) $\deg(P_n) = n$, and
(ii) $\langle \mathbf{u}, P_n\, P_m \rangle = \delta_{n,m}\, K_n$ with $K_n \neq 0$,

where, as usual, $\delta_{n,m}$ denotes the Kronecker delta.

Theorem 2.4 (Existence and Uniqueness of Orthogonal Polynomials)

1. *If* $\mathbf{u}$ *is a quasi-definite functional, then there exists a sequence of orthogonal polynomials* $\{P_n(x)\}_{n \geq 0}$ *associated with* $\mathbf{u}$.

2. *If* $\{Q_n(x)\}_{n\geq 0}$ *is another sequence of orthogonal polynomials associated with* $\mathbf{u}$*, then*

$$Q_n(x) = c_n\, P_n(x), \quad n \geq 0,$$

where c_n *are non zero real numbers. That is,* $\{P_n(x)\}_{n\geq 0}$ *is unique up to multiplicative scalar factors.*

Let $\mathbf{u}$ be a quasi-definite linear functional and $\{P_n(x)\}_{n\geq 0}$ a sequence of orthogonal polynomials associated with $\mathbf{u}$. For each $n \geq 0$, let k_n denote the leading coefficient of the polynomial $P_n(x)$. The sequence of polynomials $\{\hat{P}_n(x)\}_{n\geq 0}$ with

$$\hat{P}_n(x) := k_n^{-1}\, P_n(x), \qquad n \geq 0,$$

is called a sequence of monic orthogonal polynomials associated with $\mathbf{u}$. In particular, if $\mathbf{u}$ is positive-definite, then we can define a norm on $\mathbb{P}$ by

$$||p||_{\mathbf{u}} = \sqrt{\langle \mathbf{u}, p^2\rangle}.$$

The sequence of polynomials $\{Q_n(x)\}_{n\geq 0}$ with

$$Q_n(x) := \frac{P_n(x)}{||P_n(x)||_{\mathbf{u}}}, \qquad n \geq 0,$$

is called the sequence of orthonormal polynomials with respect to $\mathbf{u}$.

Using a matrix approach, we can rewrite the orthogonality as follows. If M is the Hankel moment matrix associated with a quasi-definite linear functional, then M has a unique Gauss-Borel factorization [35, p. 441] with

$$M = S^{-1} D\, S^{-t}, \tag{2.1}$$

where, as usual, the superscript t denotes the transpose, S is a non-singular lower triangular matrix with 1's in the main diagonal, $S^{-t} := (S^{-1})^t$, and D is a diagonal matrix. With this in mind, if $\chi(x)$ denotes the semi-infinite column vector $\chi(x) := (1, x, x^2, \cdots)^t$, then the sequence of monic orthogonal polynomials arranged in a column vector as $\mathbf{P} := (P_0(x), P_1(x), P_2(x), \cdots)^t$ can be written as $\mathbf{P} = S\chi(x)$. In other words, S is the matrix of change of basis from the canonical basis to the basis of monic orthogonal polynomials.

Notice also that if $\mathbf{u}$ is a positive-definite linear functional, then the entries of D in (2.1) are positive and thus the factorization of the moment matrix M is the standard Cholesky factorization. Moreover, if $\mathbf{Q} := (Q_0(x), Q_1(x), Q_2(x), \ldots)^t$ is the vector of orthonormal polynomials, then $\mathbf{Q} := \tilde{S}\chi(x)$, where $\tilde{S} = D^{-1/2}S$.

Definition 2.5 The shift matrix is the semi-infinite matrix

$$\Lambda := \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}.$$

The shift matrix satisfies the spectral property $\Lambda \chi(x) = x\, \chi(x)$. Notice also that from the symmetry of the Hankel moment matrix M, we have that $\Lambda M = M \Lambda^t$.

Theorem 2.6 (Three-Term Recurrence Relation) *Let* $\mathbf{u}$ *be a quasi-definite linear functional and let* $\{P_n(x)\}_{n\geq 0}$ *be the sequence of monic orthogonal polynomials with respect to* $\mathbf{u}$*. Then there exist two sequences of real numbers* $\{a_n\}_{n\geq 1}$ *and* $\{b_n\}_{n\geq 0}$*, with* $a_n \neq 0$ *for* $n \geq 1$*, such that*

$$\begin{aligned} &x\, P_n(x) = P_{n+1}(x) + b_n\, P_n(x) + a_n\, P_{n-1}(x), \quad n \geq 0, \\ &P_{-1}(x) = 0, \quad P_0(x) = 1. \end{aligned} \tag{2.2}$$

Moreover,

$$b_n = \frac{\langle \mathbf{u}, x\, P_n^2 \rangle}{\langle \mathbf{u}, P_n^2 \rangle}, \quad n \geq 0, \qquad a_n = \frac{\langle \mathbf{u}, P_n^2 \rangle}{\langle \mathbf{u}, P_{n-1}^2 \rangle}, \quad n \geq 1.$$

The above three-term recurrence relation can be written in matrix form as follows

$$x\, \mathbf{P} = J_{\text{mon}} \mathbf{P}, \quad \text{where} \quad J_{\text{mon}} = \begin{pmatrix} b_0 & 1 & & \\ a_1 & b_1 & 1 & \\ & a_2 & b_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

The matrix J_{mon} is called a monic Jacobi tridiagonal matrix associated with the sequence of monic orthogonal polynomials $\{P_n(x)\}_{n\geq 0}$.

Similarly, if $\mathbf{u}$ is positive-definite, then the sequence of orthonormal polynomials satisfies a three-term recurrence relation $x\, \mathbf{Q} = J\, \mathbf{Q}$, where J is a tridiagonal semi-infinite symmetric matrix called a Jacobi matrix.

Theorem 2.7 *If* S *is the semi-infinite upper triangular matrix obtained in the Gauss-Borel factorization* (2.1)*, then*

$$J_{\text{mon}} = S\, \Lambda\, S^{-1}.$$

Similarly, if $\mathbf{u}$ *is a positive-definite linear functional and* $\tilde{S}$ *is the upper triangular matrix obtained from the Cholesky factorization of the moment matrix, then* $J = \tilde{S}\, \Lambda\, \tilde{S}^{-1}$.

Proof Let $\mathbf{P} = (P_0(x), P_1(x) \cdots)^t$ be the vector of monic orthogonal polynomials. Using the shift matrix properties, we have

$$x\,\mathbf{P} = x\,S\chi(x) = S\,\Lambda\,\chi(x) = S\,\Lambda\,S^{-1}\,(S\,\chi(x)) = S\,\Lambda\,S^{-1}\,\mathbf{P},$$

and the result follows.

If $\mathbf{u}$ is positive-definite, let $\mathbf{Q} = (Q_0(x), Q_1(x) \cdots)^t$ be the vector of orthonormal polynomials.

$$x\,\mathbf{Q} = x\,\tilde{S}\chi(x) = \tilde{S}\,\Lambda\,\chi(x) = \tilde{S}\,\Lambda\,\tilde{S}^{-1}\,(\tilde{S}\,\chi(x)) = \tilde{S}\,\Lambda\,\tilde{S}^{-1}\,\mathbf{Q},$$

and we obtain the result □

Theorem 2.8 (Favard's Theorem) *Let $\{b_n\}_{n\geq 0}$ and $\{a_n\}_{n\geq 1}$ be arbitrary sequences of real numbers with $a_n \neq 0$ for $n \geq 1$, and let $\{P_n(x)\}_{n\geq 0}$ be a sequence of monic polynomials defined by the recurrence formula*

$$\begin{aligned} &x\,P_n(x) = P_{n+1}(x) + b_n\,P_n(x) + a_n\,P_{n-1}(x), \quad n \geq 0, \\ &P_{-1}(x) = 0, \qquad P_0(x) = 1, \end{aligned}$$

then there exists a quasi-definite linear functional $\mathbf{u}$ such that $\{P_n(x)\}_{\geq 0}$ is the sequence of monic orthogonal polynomials with respect to $\mathbf{u}$. Furthermore, if $\{a_n\}_{n\geq 1}$ is a sequence such that $a_n > 0$ for $n \geq 1$, then $\mathbf{u}$ is positive-definite.

If $\{P_n(x)\}_{n\geq 0}$ is a sequence of monic orthogonal polynomials satisfying a three-term recurrence (2.2), we define the sequence of associated polynomials of the first kind $\{P_n^{(1)}(x)\}_{n\geq 0}$ as the sequence of polynomials that satisfy the three-term recurrence relation

$$\begin{aligned} &x\,P_n^{(1)}(x) = P_{n+1}^{(1)}(x) + b_{n+1}\,P_n^{(1)}(x) + a_{n+1}\,P_{n-1}^{(1)}(x), \quad n \geq 0, \\ &P_0^{(1)}(x) = 1, \quad P_{-1}^{(1)}(x) = 0. \end{aligned} \tag{2.3}$$

Proposition 2.9 *Let $\mathbf{u}$ be a quasi-definite linear functional and $\{P_n(x)\}_{n\geq 0}$ its corresponding sequence of monic orthogonal polynomials. The sequence of associated polynomials of the first kind is given by*

$$P_{n-1}^{(1)}(x) = \frac{1}{\mu_0}\left\langle \mathbf{u}_y, \frac{P_n(x) - P_n(y)}{x - y} \right\rangle, \quad n \geq 1.$$

Notice that the families of polynomials $\{P_n(x)\}_{n\geq 0}$ and $\{P_{n-1}^{(1)}(x)\}_{n\geq 0}$ are linearly independent solutions of (2.2). Thus, any other solution can be written as a linear combination of $\{P_n(x)\}_{n\geq 0}$ and $\{P_{n-1}^{(1)}(x)\}_{n\geq 0}$ with polynomial coefficients.

Definition 2.10 For each $n \geq 0$, the nth kernel polynomial is defined by

$$\mathcal{K}_n(x, y) = \sum_{j=0}^{n} \frac{P_j(x)\, P_j(y)}{\| P_j(x) \|_{\mathbf{u}}^2}. \tag{2.4}$$

Definition 2.11 Let $\mathbf{u}$ be a quasi-definite functional with moments $\{\mu_n\}_{n\geq 0}$. We define the Stieltjes function associated with $\mathbf{u}$ as the formal power series

$$\mathcal{S}(z) = \sum_{n=0}^{\infty} \frac{\mu_n}{z^{n+1}}.$$

By a linear spectral transformation of $\mathcal{S}(z)$ we mean the following transformation

$$\tilde{\mathcal{S}}(z) = \frac{A(z)\mathcal{S}(z) + B(z)}{C(z)}$$

where $A(z)$, $B(z)$, $C(z)$ are polynomials in the variable z such that

$$\tilde{\mathcal{S}}(z) = \sum_{n=0}^{\infty} \frac{\tilde{\mu}_n}{z^{n+1}}.$$

Definition 2.12 Let $\mathbf{u}$ be a linear functional and let $\{P_n(x)\}_{n\geq 0}$ be a sequence of polynomials with $\deg(P_n) = n$. We say that $\{P_n(x)\}_{n\geq 0}$ is quasi-orthogonal of order m with respect to $\mathbf{u}$ if

$$\begin{aligned} \langle \mathbf{u}, P_k\, P_n \rangle \quad &= 0, \; m + 1 \leq |n - k|, \\ \langle \mathbf{u}, P_{n-m}\, P_n \rangle &\neq 0, \; \text{for some } n \geq m. \end{aligned}$$

The sequence of polynomials $\{P_n(x)\}_{n\geq 0}$ is said to be strictly quasi-orthogonal of order m with respect to $\mathbf{u}$ if

$$\begin{aligned} \langle \mathbf{u}, P_k\, P_n \rangle \quad &= 0, \; m + 1 \leq |n - k|, \\ \langle \mathbf{u}, P_{n-m}\, P_n \rangle &\neq 0, \; \text{for every } n \geq m. \end{aligned}$$

3 Discrete Darboux Transformations

Several examples of perturbations of a quasi-definite linear functional $\mathbf{u}$ have been studied (see for example [14, 17, 18, 23, 31, 32, 71, 72, 76, 77]). In particular, the following three canonical cases (see [14, 76]) have attracted the interest of researchers. These transformations are known in the literature as discrete Darboux transformations.

3.1 Christoffel Transformation

Let $\mathbf{u}$ be a quasi-definite linear functional and $\{P_n(x)\}_{n\geq 0}$ a sequence of monic orthogonal polynomials associated with $\mathbf{u}$. Suppose that the linear functional $\tilde{\mathbf{u}}$ satisfies

$$\tilde{\mathbf{u}} = (x-a)\mathbf{u}, \tag{3.1}$$

with $a \in \mathbb{R}$. Then $\tilde{\mathbf{u}}$ is called a canonical Christoffel transformation of $\mathbf{u}$ (see [14]). Necessary and sufficient conditions for the functional $\tilde{\mathbf{u}}$ to be quasi-definite are given in [16, 76]. If $\tilde{\mathbf{u}}$ is also a quasi-definite functional, then its sequence of monic orthogonal polynomials $\{\tilde{P}_n(x)\}_{n\geq 0}$ satisfies the following connection formulas.

Proposition 3.1 *The sequences of monic orthogonal polynomial $\{P_n(x)\}_{n\geq 0}$ and $\{\tilde{P}_n(x)\}_{n\geq 0}$ are related by*

$$\begin{aligned} (x-a)\,\tilde{P}_n(x) &= P_{n+1}(x) + \lambda_n\, P_n(x), \quad n \geq 0,\\ P_n(x) &= \tilde{P}_n(x) + \nu_n\, \tilde{P}_{n-1}(x), \quad n \geq 1, \end{aligned} \tag{3.2}$$

with

$$\lambda_n = -\frac{P_{n+1}(a)}{P_n(a)}, \quad n \geq 0, \qquad \nu_n = \frac{\langle \mathbf{u}, P_n^2\rangle}{\lambda_{n-1}\langle \mathbf{u}, P_{n-1}^2\rangle}, \quad n \geq 1.$$

Notice that (3.2) can be written in matrix form

$$\begin{pmatrix} P_0(x)\\ P_1(x)\\ P_2(x)\\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & & & \\ \nu_1 & 1 & & \\ & \nu_2 & 1 & \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \tilde{P}_0(x)\\ \tilde{P}_1(x)\\ \tilde{P}_2(x)\\ \vdots \end{pmatrix},$$

$$(x-a)\begin{pmatrix} \tilde{P}_0(x)\\ \tilde{P}_1(x)\\ \tilde{P}_2(x)\\ \vdots \end{pmatrix} = \begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_1 & 1 & \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} P_0(x)\\ P_1(x)\\ P_2(x)\\ \vdots \end{pmatrix}.$$

Theorem 3.2 ([14, 76]) *Let J_{mon} and $\tilde{J}_{\text{mon}}$ be the Jacobi matrices associated with $\mathbf{u}$ and $\tilde{\mathbf{u}} = (x-a)\,\mathbf{u}$, respectively. If $J_{\text{mon}} - aI$ can be written as*

$$J_{\text{mon}} - aI = L\,U,$$

where L is a lower bidiagonal matrix with 1's in the main diagonal and U is an upper bidiagonal matrix, then

$$\tilde{J}_{\text{mon}} - aI = UL.$$

Proof Recall that from (3.2),

$$(x-a)\,\tilde{\mathbf{P}} = U\,\mathbf{P} \quad \text{and} \quad \mathbf{P} = L\,\tilde{\mathbf{P}},$$

where $\mathbf{P} = (P_0(x), P_1(x)\cdots)^t$, $\tilde{\mathbf{P}} = (\tilde{P}_0(x), \tilde{P}_1(x)\cdots)^t$, L is a lower bidiagonal matrix with 1's in the main diagonal, and U is an upper bidiagonal matrix. Thus,

$$(x-a)\,\mathbf{P} = (x-a)\,L\,\tilde{\mathbf{P}} = L\,(x-a)\,\tilde{\mathbf{P}} = (L\,U)\,\mathbf{P},$$

and since $(x-a)\,\mathbf{P} = (J_{\text{mon}} - aI)\mathbf{P}$, it follows that

$$(J_{\text{mon}} - aI)\mathbf{P} = (L\,U)\mathbf{P}.$$

Since $\{P_n(x)\}_{n\geq 0}$ constitutes a basis of the linear space of polynomials, then $J_{\text{mon}} - aI = L\,U$. On the other hand,

$$(x-a)\,\tilde{\mathbf{P}} = U\,\mathbf{P} = (UL)\,\tilde{\mathbf{P}},$$

but, as above, this implies that $\tilde{J}_{\text{mon}} - aI = UL$. □

3.2 Geronimus Transformation

Let $\mathbf{u}$ be a quasi-definite linear functional, and introduce the linear functional $\hat{\mathbf{u}}$

$$\hat{\mathbf{u}} = (x-a)^{-1}\mathbf{u} + M\delta(x-a), \tag{3.3}$$

i.e., for every polynomial $p(x)$,

$$\langle \hat{\mathbf{u}}, p(x)\rangle = \left\langle \mathbf{u}, \frac{p(x)-p(a)}{x-a} \right\rangle + Mp(a). \tag{3.4}$$

We say that $\hat{\mathbf{u}}$ is a canonical Geronimus transformation of $\mathbf{u}$ (see [31]). Necessary and sufficient conditions for the functional $\hat{\mathbf{u}}$ to be quasi-definite are given in [23, 32, 76]. If $\hat{\mathbf{u}}$ is also a quasi-definite linear functional, then we denote by $\{\hat{P}_n(x)\}_{n\geq 0}$ its sequence of monic orthogonal polynomials.

Proposition 3.3 *The sequences of monic orthogonal polynomials* $\{P_n(x)\}_{n\geq 0}$ *and* $\{\hat{P}_n(x)\}_{n\geq 0}$ *are related by*

$$\begin{aligned} \hat{P}_n(x) = &\quad P_n(x) + \varsigma_n\, P_{n-1}(x), \quad n \geq 1, \\ (x-a)\, P_n(x) = &\quad \hat{P}_{n+1}(x) + \rho_n\, \hat{P}_n(x), \quad n \geq 0, \end{aligned} \tag{3.5}$$

where

$$\varsigma_n = -\frac{\mu_0 P_{n-1}^{(1)}(a) + M P_n(a)}{\mu_0 P_{n-2}^{(1)}(a) + M P_{n-1}(a)}, \quad n \geq 1,$$

$$\rho_n = \frac{\left(\mu_0 P_{n-2}^{(1)}(a) + M P_{n-1}(a)\right) \langle \mathbf{u}, P_n^2(x)\rangle}{\left(\mu_0 P_{n-1}^{(1)}(a) + M P_n(a)\right) \langle \mathbf{u}, P_{n-1}^2(x)\rangle}, \quad n \geq 1,$$

$$\rho_0 = \frac{\mu_0}{\hat{\mu}_0},$$

and $\{P_n^{(1)}(x)\}_{n\geq 0}$ *is the sequence of polynomials of the first kind* (2.3).

Notice that (3.5) can be written in matrix form as

$$\begin{pmatrix} \hat{P}_0(x) \\ \hat{P}_1(x) \\ \hat{P}_2(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & & & \\ \varsigma_1 & 1 & & \\ & \varsigma_2 & 1 & \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} P_0(x) \\ P_1(x) \\ P_2(x) \\ \vdots \end{pmatrix},$$

$$(x-a) \begin{pmatrix} P_0(x) \\ P_1(x) \\ P_2(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} \rho_0 & 1 & & \\ & \rho_1 & 1 & \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \hat{P}_0(x) \\ \hat{P}_1(x) \\ \hat{P}_2(x) \\ \vdots \end{pmatrix}.$$

Theorem 3.4 ([14, 76]) *Let* J_{mon} *and* $\hat{J}_{\text{mon}}$ *be the Jacobi matrices associated with* $\mathbf{u}$ *and* $\hat{\mathbf{u}}$, *respectively. If the semi-infinite matrix* $J_{\text{mon}} - aI$ *can be written as*

$$J_{\text{mon}} - aI = UL,$$

where L *is a lower bidiagonal matrix and* U *is a upper bidiagonal matrix, then*

$$\hat{J}_{\text{mon}} - aI = L\,U.$$

Proof From (3.5),

$$(x-a)\,\mathbf{P} = U\,\hat{\mathbf{P}}, \quad \text{and} \quad \hat{\mathbf{P}} = L\,\mathbf{P},$$

where, $\mathbf{P} = (P_0(x), P_1(x)\cdots)^t$, $\hat{\mathbf{P}} = (\hat{P}_0(x), \hat{P}_1(x)\cdots)^t$, L is a lower bidiagonal matrix with 1's in the main diagonal, and U is an upper bidiagonal matrix. Then,

$$(x-a)\,\mathbf{P} = U\,\hat{\mathbf{P}} = (UL)\,\mathbf{P}.$$

But $\{P_n(x)\}_{n\geq 0}$ is a basis for $\mathbb{P}$, and $(x-a)\,\mathbf{P} = (J_{\text{mon}} - aI)\,\mathbf{P}$, we get

$$J_{\text{mon}} - aI = UL.$$

Notice that this factorization depends on the choice of the free parameter $\hat{\mu}_0 \neq 0$. For a fixed $\hat{\mu}_0$,

$$(x-a)\,\hat{\mathbf{P}} = (x-a)\,L\,\mathbf{P} = L\,(x-a)\,\mathbf{P} = (L\,U)\,\hat{\mathbf{P}}.$$

As above, $\hat{J}_{\text{mon}} - aI = LU$. □

3.3 Uvarov Transformation

Let $\mathbf{u}$ be a quasi-definite linear functional and suppose that the linear functional $\check{\mathbf{u}}$ is defined by

$$\check{\mathbf{u}} = \mathbf{u} + M\delta(x-a). \tag{3.6}$$

The linear functional $\check{\mathbf{u}}$ is said to be a canonical Uvarov transformation of $\mathbf{u}$ (see [71, 72]). Necessary and sufficient conditions for the quasi-definiteness of the linear functional $\check{\mathbf{u}}$ are given in [49].

Proposition 3.5 *Suppose that $\check{\mathbf{u}}$ is quasi-definite, and let $\{\check{P}_n(x)\}_{n\geq 0}$ denote the sequence of monic orthogonal polynomials associated with $\check{\mathbf{u}}$. The sequences of polynomial $\{P_n(x)\}_{n\geq 0}$, and $\{\check{P}_n\}_{n\geq 0}$ are related by*

$$\check{P}_n(x) = P_n(x) - \frac{M P_n(a)}{1 + M\mathcal{K}_{n-1}(a,a)}\mathcal{K}_{n-1}(x,a), \quad n \geq 1,$$

where $\mathcal{K}_n(x,y)$ denotes the nth kernel polynomial defined in (2.4).

For any linear functional $\mathbf{u}$, it is straightforward to verify that if a canonical Christoffel transformation is applied to $\hat{\mathbf{u}}$ in (3.3) with the same parameter a, then we recover the original linear functional $\mathbf{u}$, that is, the canonical Christoffel transformation is the left inverse of the canonical Geronimus transformation.

However, a canonical Geronimus transformation applied to the linear functional $\tilde{\mathbf{u}}$ in (3.1) with the same parameter a, transforms $\tilde{\mathbf{u}}$ into a linear functional $\check{\mathbf{u}}$ as in (3.6), that is, a canonical Uvarov transformation. It is important to notice that the following result holds.

Theorem 3.6 ([77]) *Every linear spectral transform is a finite composition of Christoffel and Geronimus transformations.*

4 Semiclassical Linear Functionals

Let D denote the derivative operator. Given a linear functional $\mathbf{u}$, we define $D\mathbf{u}$ as

$$\langle D\mathbf{u}, p\rangle = -\langle \mathbf{u}, p'\rangle,$$

for every polynomial $p \in \mathbb{P}$. Inductively, we define

$$\langle D^n\mathbf{u}, p\rangle = (-1)^n\langle \mathbf{u}, p^{(n)}\rangle.$$

Notice that, for any polynomial $q(x)$,

$$D(q(x)\,\mathbf{u}) = q'(x)\,\mathbf{u} + q(x)\,D\mathbf{u}.$$

Definition 4.1 A quasi-definite linear functional $\mathbf{u}$ is said to be semiclassical if there exist non-zero polynomials ϕ and ψ with $\deg(\phi) =: r \geq 0$ and $\deg(\psi) =: t \geq 1$, such that $\mathbf{u}$ satisfies the Pearson equation

$$D(\phi(x)\mathbf{u}) + \psi(x)\mathbf{u} = 0. \tag{4.1}$$

In general, if $\mathbf{u}$ satisfies (4.1), then it satisfies an infinite number of Pearson equations. Indeed, for any non-zero polynomial $q(x)$, $\mathbf{u}$ satisfies

$$D(\tilde{\phi}\,\mathbf{u}) + \tilde{\psi}\,\mathbf{u} = 0,$$

where $\tilde{\phi}(x) = q(x)\,\phi(x)$ and $\tilde{\psi}(x) = q'(x)\,\phi(x) + q(x)\,\psi(x)$.

Remark 4.2 In order to avoid any incompatibility with the quasi-definite character of the semiclassical functional $\mathbf{u}$, it will be required from now on that, if

$$\phi(x) = a_r\,x^r + \cdots \quad \text{and} \quad \psi(x) = b_t\,x^t + \cdots,$$

then, for any $n = 0, 1, 2, \ldots$, if $t = r - 1$, then $n\,a_r - b_t \neq 0$. In such a case, every moment of the linear functional $\mathbf{u}$ is well defined.

This motivates the following definition.

Definition 4.3 The *class* of a semiclassical linear functional $\mathbf{u}$ is defined as

$$\mathfrak{s}(\mathbf{u}) := \min \max\{\deg(\phi) - 2,\ \deg(\psi) - 1\},$$

where the minimum is taken among all pairs of polynomials ϕ and ψ such that $\mathbf{u}$ satisfies (4.1).

Lemma 4.4 *Let* $\mathbf{u}$ *be a semiclassical functional such that*

$$D(\phi_1\mathbf{u}) + \psi_1\mathbf{u} = 0, \qquad s_1 := \max\{\deg(\phi_1) - 2, \deg(\psi_1) - 1\}, \tag{4.2}$$

$$D(\phi_2\mathbf{u}) + \psi_2\mathbf{u} = 0, \qquad s_2 := \max\{\deg(\phi_2) - 2, \deg(\psi)_2 - 1\}, \tag{4.3}$$

where $\phi_i(x)$ *and* $\psi_i(x)$, $i = 1, 2$, *are non-zero polynomials with* $\deg(\phi_i) \geq 0$ *and* $\deg(\psi_i) \geq 1$. *Let* $\phi(x)$ *be the greatest common divisor of* $\phi_1(x)$ *and* $\phi_2(x)$.

Then there exists a polynomial $\psi(x)$ *such that*

$$D(\phi\, \mathbf{u}) + \psi\, \mathbf{u} = 0, \qquad s := \max\{\deg(\phi) - 2, \deg(\psi) - 1\}.$$

Moreover, $s - \deg(\phi) = s_1 - \deg(\phi_1) = s_2 - \deg(\phi_2)$.

Proof From the hypothesis, there exist polynomials $\tilde{\phi}_1$ and $\tilde{\phi}_2$ such that $\phi_1 = \phi\, \tilde{\phi}_1$ and $\phi_2 = \phi\, \tilde{\phi}_2$. If ϕ_1 and ϕ_2 are coprime, then set $\phi = 1$. From (4.2) and (4.3), we obtain

$$\tilde{\phi}_2 D(\phi_1\mathbf{u}) - \tilde{\phi}_1 D(\phi_2\mathbf{u}) + (\tilde{\phi}_2\, \psi_1 - \tilde{\phi}_1\, \psi_2)\mathbf{u} = 0. \tag{4.4}$$

Observe that, for any polynomial $p \in \mathbb{P}$,

$$\begin{aligned}
\langle \tilde{\phi}_2 D(\phi_1\mathbf{u}) - \tilde{\phi}_1 D(\phi_2\mathbf{u}), p\rangle &= -\langle \mathbf{u}, \phi_1(\tilde{\phi}_2\, p)'\rangle + \langle \mathbf{u}, \phi_2\, (\tilde{\phi}_1\, p)'\rangle \\
&= \langle \mathbf{u}, (\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2')\, p + (\tilde{\phi}_1\, \phi_2 - \tilde{\phi}_2\, \phi_1)\, p'\rangle \\
&= \langle \mathbf{u}, (\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2')\, p + \phi\, (\tilde{\phi}_1\, \tilde{\phi}_2 - \tilde{\phi}_2\, \tilde{\phi}_1)\, p'\rangle \\
&= \langle \mathbf{u}, (\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2')\, p\rangle \\
&= \langle (\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2')\, \mathbf{u}, p\rangle.
\end{aligned}$$

Therefore, (4.4) becomes $(\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2' + \tilde{\phi}_2\, \psi_1 - \tilde{\phi}_1\, \psi_2)\, \mathbf{u} = 0$. Since $\mathbf{u}$ is quasi-definite, then

$$\phi_2\, \tilde{\phi}_1' - \phi_1\, \tilde{\phi}_2' + \tilde{\phi}_2\, \psi_1 - \tilde{\phi}_1\, \psi_2 = 0,$$

or, equivalently, $(\tilde{\phi}_1'\, \phi + \psi_1)\tilde{\phi}_2 = (\tilde{\phi}_2'\, \phi + \psi_2)\tilde{\phi}_1$.

But $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are coprime polynomials. Hence, there exists a polynomial ψ such that

$$\tilde{\phi}_1' \, \phi + \psi_1 = \psi \, \tilde{\phi}_1, \quad \tilde{\phi}_2' \, \phi + \psi_2 = \psi \, \tilde{\phi}_2. \tag{4.5}$$

Since $\phi_1 = \phi \, \tilde{\phi}_1$ and $\phi_2 = \phi \, \tilde{\phi}_2$, (4.2) and (4.3) can be written as

$$\tilde{\phi}_1 \, D(\phi \, \mathbf{u}) + (\tilde{\phi}_1' \, \phi + \psi_1) \, \mathbf{u} = 0, \quad \tilde{\phi}_2 \, D(\phi \, \mathbf{u}) + (\tilde{\phi}_2 \, \phi + \psi_2) \, \mathbf{u} = 0.$$

Using (4.5), we write

$$\tilde{\phi}_1 \, (D(\phi \mathbf{u}) + \psi \mathbf{u}) = 0, \quad \tilde{\phi}_2 \, (D(\phi \mathbf{u}) + \psi \mathbf{u}) = 0,$$

and the result follows from the Bézout identity for coprime polynomials. □

Theorem 4.5 ([59]) *For any semiclassical linear functional* $\mathbf{u}$, *the polynomials* ϕ *and* ψ *in* (4.1) *such that*

$$\mathfrak{s}(\mathbf{u}) = \max\{\deg(\phi) - 2, \deg(\psi) - 1\}$$

are unique up to a constant factor.

Proof Suppose that $\mathbf{u}$ satisfies (4.1) with ϕ_i and ψ_i, $i = 1, 2$, and suppose that $\mathfrak{s}(\mathbf{u}) = \max\{\deg(\phi_i) - 2, \deg(\psi_i) - 1\}$, $i = 1, 2$. If in Lemma 4.4 we take $s_1 = s_2$, then $s = s_1 = s_2$. But this implies that $\deg(\phi) = \deg(\phi_1) = \deg(\phi_2)$, or, equivalently, $\phi = \phi_1 = \phi_2$. Notice also that ψ is unique up to a constant factor. □

The polynomials ϕ and ψ such that $\mathfrak{s}(\mathbf{u}) = \max\{\deg(\phi) - 2, \deg(\psi) - 1\}$ are characterized in the following result.

Proposition 4.6 ([57]) *Let* $\mathbf{u}$ *be a semi-classical linear functional and let* $\phi(x)$ *and* $\psi(x)$ *be non-zero polynomials with* $\deg(\phi) =: r$ *and* $\deg(\psi) =: t$, *such that* (4.1) *holds. Let* $s := \max(r - 2, t - 1)$. *Then* $s = \mathfrak{s}(\mathbf{u})$ *if and only if*

$$\prod_{c:\, \phi(c)=0} \Big(|\psi(c) + \phi'(c)| + |\langle \mathbf{u}, \theta_c \psi + \theta_c^2 \phi \rangle| \Big) > 0. \tag{4.6}$$

Here, $(\theta_c f)(x) = \dfrac{f(x) - f(c)}{x - c}$.

Proof Let c be a zero of ϕ, then there exists a polynomial $\phi_c(x)$ of degree $r - 1$ such that $\phi(x) = (x - c)\phi_c(x)$. On the other hand, since

$$\theta_c^2 \phi(x) = \frac{\phi(x) - \phi(c)}{(x - c)^2} - \frac{\phi'(c)}{x - c},$$

then

$$\psi(x) + \phi_c(x) = (x - c)\psi_c(x) + r_c,$$

where

$$\psi_c(x) = \theta_c\psi(x) + \theta_c^2\phi(x), \quad r_c = \psi(c) + \phi'(c).$$

With this in mind, (4.1) can be written as $(x-c)(D(\phi_c\,\mathbf{u}) + \psi_c\,\mathbf{u}) + r_c\mathbf{u} = 0$. From here, we obtain

$$\begin{aligned} D(\phi_c\mathbf{u}) + \psi_c\mathbf{u} &= -\frac{r_c}{(x-c)}\mathbf{u} + \langle \mathbf{u}, \psi_c\rangle\, \delta(x-c) \\ &= -\frac{\psi(c) + \phi'(c)}{(x-c)}\mathbf{u} + \langle \mathbf{u}, \theta_c\psi + \theta_c^2\phi\rangle\, \delta(x-c). \end{aligned}$$

Next, we proceed to the proof of the proposition.

Suppose that $\mathfrak{s}(\mathbf{u}) = s$, $r_c = 0$ and $\langle \mathbf{u}, \psi_c\rangle = 0$ for some c such that $\phi(c) = 0$. Then $D(\phi_c\mathbf{u}) + \psi_c\mathbf{u} = 0$. But $\deg(\phi_c) = r - 1$ and $\deg(\psi_c) = t - 1$. This means that $\mathfrak{s}(\mathbf{u}) = s - 1$, which is a contradiction.

Now, suppose that (4.6) holds and that $\mathbf{u}$ is of class $\tilde{s} \leq s$, with $D(\tilde{\phi}\,\mathbf{u}) + \tilde{\psi}\,\mathbf{u} = 0$. From Lemma 4.4, there exists a polynomial $\rho(x)$ such that

$$\phi(x) = \rho(x)\tilde{\phi}(x), \qquad \psi(x) = \rho(x)\tilde{\psi}(x) - \rho'(x)\tilde{\phi}(x).$$

If $\tilde{s} < s$, then necessarily $\deg(\rho) \geq 1$. Let c be a zero of $\rho(x)$ and let $\rho_c(x)$ be the polynomial such that $\rho(x) = (x-c)\,\rho_c(x)$. Then,

$$\psi(x) + \phi_c(x) = (x - c)\,\left(\rho_c(x)\,\tilde{\psi}(x) - \rho_c'(x)\,\tilde{\phi}(x)\right).$$

It follows that

$$r_c = 0, \qquad \psi_c(x) = \rho_c(x)\,\tilde{\psi}(x) - \rho_c'(x)\,\tilde{\phi}(x).$$

Hence,

$$\langle \mathbf{u}, \psi_c\rangle = \langle \mathbf{u}, \tilde{\psi}\,\rho_c\rangle - \langle \mathbf{u}, \rho_c'\,\tilde{\phi}\rangle = \langle D(\tilde{\phi}\,\mathbf{u}) + \tilde{\psi}\,\mathbf{u}, \rho_c\rangle = 0.$$

But this means that $\phi(c) = 0$ and

$$|\psi(c) + \phi'(c)| + |\langle \mathbf{u}, \theta_c\psi + \theta_c^2\phi\rangle| = 0,$$

which contradicts (4.6). Thus, $s = \tilde{s}$ and, by Theorem 4.5, $\tilde{\phi}$ and $\tilde{\psi}$ are multiple of ϕ and ψ, respectively, up to a constant factor. $\square$

Proposition 4.7 ([34, 58]) *Let* $\mathbf{u}$ *be a linear functional. The following statements are equivalent.*

(1) $\mathbf{u}$ *is semiclassical.*
(2) *There exist two non-zero polynomials* ϕ *and* ψ *with* $\deg(\phi) =: r \geq 0$ *and* $\deg(\psi) =: t \geq 1$*, such that the Stieltjes function* $\mathcal{S}(z)$ *associated with* $\mathbf{u}$ *satisfies*

$$\phi(z)\,\mathcal{S}'(z) + (\psi(z) + \phi'(z))\,\mathcal{S}(z) = C(z), \tag{4.7}$$

where

$$C(z) = (\mathbf{u} * \theta_0 \left(\psi + \phi'\right))(z) - (D\mathbf{u} * \theta_0\phi)(z).$$

Proof (1)⇒(2) Let $\mathbf{u}$ be a semiclassical functional of class s satisfying (4.1).

$$\phi(z) = \sum_{k=0}^{r} \frac{\phi^{(k)}(0)}{k!} z^k, \qquad \psi(z) = \sum_{m=0}^{t} \frac{\psi^{(m)}(0)}{m!} z^m,$$

we have

$$\begin{aligned} 0 = \langle D(\phi\mathbf{u}) + \psi\mathbf{u}, x^n \rangle &= -\langle \mathbf{u}, n x^{n-1}\phi(x)\rangle + \langle \mathbf{u}, x^n \psi(x)\rangle \\ &= -n \sum_{k=0}^{r} \frac{\phi^{(k)}(0)}{k!} \mu_{n+k-1} + \sum_{m=0}^{t} \frac{\psi^{(m)}(0)}{m!} \mu_{n+m}. \end{aligned}$$

Multiplying the above relation by $1/z^{n+1}$ and taking the infinite sum over n, we obtain

$$0 = -\sum_{n=0}^{\infty} n \sum_{k=0}^{r} \frac{\phi^{(k)}(0)}{k!} \frac{\mu_{n+k-1}}{z^{n+1}} + \sum_{n=0}^{\infty}\sum_{m=0}^{t} \frac{\psi^{(m)}(0)}{m!} \frac{\mu_{n+m}}{z^{n+1}}. \tag{4.8}$$

It is straightforward to verify that

$$\begin{aligned} \sum_{n=0}^{\infty}\sum_{m=0}^{t} \frac{\psi^{(m)}(0)}{m!} \frac{\mu_{n+m}}{z^{n+1}} &= \psi(z)\mathcal{S}(z) - \sum_{m=1}^{t}\sum_{n=0}^{m-1} \frac{\psi^{(m)}(0)}{m!} \mu_n z^{m-1-n} \\ &= \psi(z)\mathcal{S}(z) - (\mathbf{u} * \theta_0\psi)(z). \end{aligned}$$

On the other hand,

$$\mathcal{S}'(z) = -\sum_{n=0}^{\infty} (n+1) \frac{\mu_n}{z^{n+2}}.$$

Thus,

$$\sum_{n=0}^{\infty}\sum_{k=0}^{r} n\frac{\phi^{(k)}(0)}{k!}\frac{\mu_{n+k-1}}{z^{n+1}} = -\phi(z)\mathcal{S}'(z) - \phi'(z)\mathcal{S}(z) + \sum_{k=2}^{r}\sum_{n=0}^{k-2}\frac{\phi^{(k)}(0)}{(k-1)!}\frac{\mu_n}{z^{n-k+2}}$$
$$- \sum_{k=2}^{r}\sum_{n=0}^{k-2}(n+1)\frac{\phi^{(k)}(0)}{k!}\frac{\mu_n}{z^{n-k+2}}$$
$$= -\phi(z)\mathcal{S}'(z) - \phi'(z)\mathcal{S}(z) + (\mathbf{u} * \theta_0\phi')(z) - (D\mathbf{u} * \theta_0\phi)(z).$$

Hence, (4.7) follows from (4.8).
(2)⇒(1) Suppose that (4.7) holds for some non-zero polynomials ϕ and ψ. Since each step above is also true in the reverse direction, then (4.7) is equivalently to (4.8). But this implies that, for every $n \geq 0$,

$$0 = -\langle \mathbf{u}, n\, x^{n-1}\, \phi\rangle + \langle \mathbf{u}, x^n\, \psi\rangle = \langle D(\phi\, \mathbf{u}) + \psi\, \mathbf{u}, x^n\rangle.$$

Therefore, $\mathbf{u}$ is semiclassical. □

Proposition 4.8 ([59]) *Let* $\mathbf{u}$ *be a linear functional, and let* $\{P_n(x)\}_{n\geq 0}$ *be its sequence of monic orthogonal polynomials. The following statements are equivalent.*

(1) *The linear functional* $\mathbf{u}$ *is semiclassical of class* s.
(2) *For* $n \geq 0$*, let* $R_n(x) = \dfrac{P'_{n+1}(x)}{n+1}$*. There exists a non-zero polynomial* $\phi(x)$ *with* $\deg(\phi) = r$*, such that the sequence of monic polynomials* $\{R_n(x)\}_{n\geq 0}$ *is quasi-orthogonal of order* s *with respect to the linear functional* $\phi(x)\,\mathbf{u}$.

Proof (1)⇒(2) Let ϕ and ψ be non-zero polynomials with $\deg(\phi) =: r \geq 0$ and $\deg(\psi) =: t \geq 1$ such that $\mathbf{u}$ satisfies $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ and $s := \max\{r-2, t-1\}$ is the class of $\mathbf{u}$. Note that

$$\langle \phi\,\mathbf{u}, x^m P'_{n+1}\rangle = \langle \phi\,\mathbf{u}, (x^m P_{n+1})'\rangle - \langle \phi\,\mathbf{u}, m\, x^{m-1} P_{n+1}\rangle$$
$$= \langle \mathbf{u}, \left(x^m\psi - m\, x^{m-1}\phi\right)\, P_{n+1}\rangle.$$

The above implies that $\langle \phi\,\mathbf{u}, x^m P'_{n+1}\rangle = 0$ for $0 \leq m \leq n-s-1$. Moreover, from Remark 4.2, $x^m\psi(x) - m\, x^{m-1}\phi(x)$ has degree $s+m+1$ and, thus, $\langle \phi\,\mathbf{u}, x^{n-s} P'_{n+1}\rangle \neq 0$. Hence, $R_n(x)$ is quasi-orthogonal of order s.

(2)⇒(1) Suppose that there exists some non-zero polynomial ϕ with $\deg(\phi) =: r \geq 0$, such that the sequence of polynomials $\{R_n(x)\}_{n\geq 0}$ is quasi-orthogonal of order s with respect to the linear functional $\phi(x)\mathbf{u}$. Since $\left\{\frac{P_n(x)}{\|P_n\|^2}\mathbf{u}\right\}_{n\geq 0}$ is a basis of the dual space of $\mathbb{P}$, then

$$D(\phi\,\mathbf{u}) = \sum_{n=0}^{\infty} \alpha_n \frac{P_n(x)}{\|P_n\|^2}\mathbf{u},$$

where $\alpha_n = \langle D(\phi\,\mathbf{u}), P_n\rangle = -\langle \phi\,\mathbf{u}, P_n'\rangle = -n\langle \mathbf{u}, \phi\, R_{n-1}\rangle$, $n \geq 1$, $\alpha_0 = 0$.

From the quasi-orthogonality of $\{R_n(x)\}_{n\geq 0}$, $\alpha_n = 0$ when $s+2 \leq n$. Thus,

$$D(\phi\mathbf{u}) + \psi\mathbf{u} = 0, \quad \text{where} \quad \psi(x) = -\sum_{n=1}^{s+1} \alpha_n \frac{P_n(x)}{\|P_n\|^2}.$$

□

Corollary 4.9 *A linear functional* $\mathbf{u}$ *with associated sequence of monic orthogonal polynomials* $\{P_n(x)\}_{n\geq 0}$ *is semiclassical of class* s *if and only if there is a non-zero polynomial* ϕ *such that sequence of monic polynomials* $\{F_n(x)\}_{n\geq 0}$*, where* $F_n(x) = \dfrac{P_{n+m}^{(m)}(x)}{(n+1)_m}$*, is quasi-orthogonal of order* s *with respect to the linear functional* $\phi^m\,\mathbf{u}$.

Proposition 4.10 ([59]) *Let* $\mathbf{u}$ *be a linear functional and* $\{P_n(x)\}_{n\geq 0}$ *its sequence of monic orthogonal polynomials. The following statements are equivalent.*

(1) $\mathbf{u}$ *is semiclassical of class* s.
(2) *There exist a nonnegative integer number* s *and a monic polynomial* $\phi(x)$ *of degree* r *with* $0 \leq r \leq s+2$*, such that*

$$\phi(x)\, P_{n+1}'(x) = \sum_{k=n-s}^{n+r} \lambda_{n,k}\, P_k(x), \quad n \geq s, \quad \lambda_{n,n-s} \neq 0. \tag{4.9}$$

If $s \geq 1$, $r \geq 1$ *and* $\lambda_{s,0} \neq 0$*, then* s *is the class of* $\mathbf{u}$.

Proof (1)⇒(2) Suppose that $\mathbf{u}$ is of class s satisfying $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ with $\deg(\phi) = r$. Since $\{P_n(x)\}_{n\geq 0}$ is a basis of $\mathbb{P}$, for each $n \geq 0$, there exists a set of real numbers $(\lambda_{n,k})_{k=0}^{n+r}$ such that

$$\phi(x)\, P_{n+1}'(x) = \sum_{k=0}^{r+n} \lambda_{n,k}\, P_k(x).$$

Using orthogonality,

$$\lambda_{n,k} = \frac{\langle \phi\,\mathbf{u}, P_{n+1}'\, P_k\rangle}{\langle \mathbf{u}, P_k^2\rangle} = \frac{(n+1)\langle \phi\,\mathbf{u}, R_n\, P_k\rangle}{\langle \mathbf{u}, P_k^2\rangle},$$

where, for each $n \geq 0$, $R_n(x) = \frac{P_{n+1}'(x)}{n+1}$. But $\mathbf{u}$ is semiclassical of class s, then, from Proposition 4.8, $R_n(x)$ is quasi-orthogonal of order s with respect to $\phi\,\mathbf{u}$. Therefore, $\lambda_{n,k} = 0$, when $s+1 < n-k$, and $\lambda_{n,n-s} \neq 0$.

(2)⇒(1) Assume that $\{P_n(x)\}_{n\geq 0}$ satisfies (4.9). Since $\left\{\frac{P_n(x)}{\|P_n(x)\|^2}\mathbf{u}\right\}_{n\geq 0}$ is a basis of the dual space of $\mathbb{P}$, then

$$D(\phi\mathbf{u}) = \sum_{n=0}^{\infty} \alpha_n \frac{P_n(x)}{\|P_n(x)\|^2}\mathbf{u}.$$

Using (4.9),

$$\alpha_n = \langle \mathbf{u}, \phi P_n'(x)\rangle = \sum_{k=n-s}^{n+r} \lambda_{n,k}\langle \mathbf{u}, P_k(x)\rangle = \begin{cases} 0, & n > s, \\ \lambda_{n,0}\langle \mathbf{u}, P_0\rangle, & n \leq s. \end{cases}$$

Therefore, $\mathbf{u}$ satisfies $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ with

$$\psi(x) = -\sum_{n=0}^{s+1} \alpha_n \frac{P_n(x)}{\|P_n\|^2},$$

hence, $\mathbf{u}$ is semiclassical. Observe that if in particular $\lambda_{s,0} \neq 0$, $\mathbf{u}$ is of class s. □

Using the three-term recurrence relation (2.2), (4.9) can be written in a compact form as shown in the following result.

Theorem 4.11 ([59]) *Let* $\mathbf{u}$ *be a semiclassical functional of class s, and* $\{P_n(x)\}_{n\geq 0}$ *its associated sequence of monic orthogonal polynomials. Then there exists a non-zero polynomial* ϕ *with* $\deg(\phi) =: r \geq 0$*, such that*

$$\phi(x)P_{n+1}'(x) = \frac{C_{n+1}(x) - C_0(x)}{2}P_{n+1}(x) - D_{n+1}(x)P_n(x), \quad n \geq 0, \tag{4.10}$$

where $\{C_n(x)\}_{n\geq 0}$ *and* $\{D_n(x)\}_{n\geq 0}$ *are polynomials satisfying*

$$\begin{aligned} C_{n+1}(x) &= -C_n(x) + \frac{2D_n(x)}{a_n}(x - b_n), \quad n \geq 0, \\ C_0(x) &= -\psi(x) - \phi'(x) \end{aligned} \tag{4.11}$$

and

$$D_{n+1}(x) = -\phi(x) + \frac{a_n}{a_{n-1}}D_{n-1}(x) + \frac{D_n(x)}{a_n}(x-b_n)^2 - C_n(x)(x-b_n),\ n \geq 0,$$
$$D_0(x) = -(\mathbf{u} * \theta_0\phi)'(x) - (\mathbf{u} * \theta_0\psi)(x), \quad D_{-1}(x) = 0.$$

The above expression leads to the so-called ladder operators associated with the linear functional $\mathbf{u}$. Using (4.11) and the three-term recurrence relation (2.2), we can deduce from (4.10) that, for $n \geq 0$,

$$\phi(x)P'_{n+1}(x) = -\left(\frac{C_{n+2}(x) + C_0(x)}{2}\right) P_{n+1}(x) + \frac{D_{n+1}(x)}{a_{n+1}} P_{n+2}(x). \qquad (4.12)$$

The relations (4.10) and (4.12) are essential to deduce a second-order linear differential equation satisfied by the polynomials $\{P_n(x)\}_{n\geq 0}$ (see [34, 37, 59]), which reads

$$J(x,n)P''_{n+1}(x) + K(x,n)P'_{n+1}(x) + L(x,n)P_{n+1}(x) = 0, \quad n \geq 0,$$

where, for $n \geq 0$,

$$J(x,n) \;= \phi(x)D_{n+1}(x),$$

$$K(x;n) = (\phi'(x) + C_0(x))D'_{n+1}(x) - \phi(x)\, D'_{n+1}(x),$$

and

$$L(x,n) = \left(\frac{C_{n+1}(x) - C_0(x)}{2}\right) D'_{n+1}(x) - \left(\frac{C'_{n+1}(x) - C'_0(x)}{2}\right) D_{n+1}(x) - D_{n+1}(x) \sum_{k=0}^{n} \frac{D_k(x)}{a_k}.$$

Notice that the degrees of the polynomials J, K, L are at most $2s+2$, $2s+1$, and $2s$, respectively.

Theorem 4.12 ([9, 10]) *Let* **u** *be a quasi-definite linear functional and* $\{P_n(x)\}_{n\geq 0}$ *the sequence of monic orthogonal polynomials associated with* **u**. *The following statements are equivalent.*

(1) **u** *is semiclassical.*
(2) $\{P_n(x)\}_{n\geq 0}$ *satisfies the following nonlinear differential equation.*

$$\phi(x)[P_{n+1}(x)\, P_n(x)]' = \frac{D_n(x)}{a_n} P^2_{n+1}(x) - C_0(x)P_{n+1}(x)P_n(x) - D_{n+1}(x)P^2_n(x), \qquad (4.13)$$

where $D_n(x)$, $C_0(x)$ *and* a_n *are the same as in* (4.10).

Proof (1)⇒(2) Suppose that $\mathbf{u}$ is semiclassical. From (4.10), we have

$$\begin{aligned}\phi(x)[P_{n+1}(x)P_n(x)]' &= P_n(x)\left(\frac{C_{n+1}(x)-C_0(x)}{2}P_{n+1}(x)-D_{n+1}(x)P_n(x)\right)\\ &\quad + P_{n+1}(x)\left(\frac{C_n(x)-C_0(x)}{2}P_n(x)-D_n(x)P_{n-1}(x)\right)\\ &= -D_{n+1}(x)P_n^2(x)+\left(\frac{C_{n+1}(x)+C_n(x)-2C_0(x)}{2}\right)P_{n+1}(x)P_n(x)\\ &\quad - D_n(x)P_{n+1}(x)P_{n-1}(x).\end{aligned}$$

Now, taking into account that $P_{n-1}(x) = \frac{(x-b_n)}{a_n}P_n(x) - \frac{1}{a_n}P_{n+1}(x)$ the above relation becomes

$$\begin{aligned}\phi[P_{n+1}(x)P_n(x)]' &= -D_{n+1}P_n^2(x)+\frac{D_n}{a_n}P_{n+1}^2(x)\\ &\quad + \left(\frac{C_{n+1}(x)+C_n(x)-2C_0(x)}{2}-\frac{(x-b_n)}{a_n}D_n\right)P_{n+1}(x)P_n(x).\end{aligned}$$

Using the relation (4.11), we get the result.

(2)⇒(1) Let $\mathbf{u}$ be a quasi-definite linear functional, and let $\{P_n(x)\}_{n\geq 0}$ be the sequence of monic orthogonal polynomials associated with $\mathbf{u}$.

Suppose that $\{P_n(x)\}_{n\geq 0}$ satisfies (4.13). Using the three-term recurrence relation $P_{n+1}(x) = (x-b_n)\,P_n(x) - a_n\,P_{n-1}(x)$ and (4.11), we can write (4.13) as

$$\begin{aligned}\phi(x)\,P'_{n+1}(x)P_n(x) &= \left(\frac{C_{n+1}(x)+C_n(x)-2C_0(x)}{2}\right)P_{n+1}(x)P_n(x)\\ &\quad - D_{n+1}(x)P_n^2(x)-D_n(x)P_{n+1}(x)P_{n-1}(x)-\phi(x)P_{n+1}(x)P_n'(x). \qquad (4.14)\end{aligned}$$

Multiplying the above relation by $P_{n-1}(x)$, and replacing $\phi(x)\,P_n'(x)\,P_{n-1}(x)$ with (4.14) for $n-1$, we obtain

$$\begin{aligned}\phi(x)\,P'_{n+1}(x)\,P_{n-1}(x) &= \left(\frac{C_{n+1}(x)-C_{n-1}(x)}{2}\right)P_{n+1}(x)\,P_{n-1}(x)\\ &\quad - D_{n+1}(x)\,P_n(x)\,P_{n-1}(x)+P_{n+1}(x)\,(D_{n-1}(x)\,P_{n-2}(x)+\phi(x)\,P'_{n-1}(x)).\end{aligned}$$

Similarly, multiplying the above relation by $P_{n-2}(x)$, and then replacing $\phi(x)\, P'_{n-1}(x)\, P_{n-2}(x)$ by (4.14) for $n-2$, we get

$$\phi(x)\, P'_{n+1}(x)\, P_{n-2}(x) = \left(\frac{C_{n+1}(x) + C_{n-2}(x) - 2C_0(x)}{2}\right) P_{n+1}(x)\, P_{n-2}(x)$$
$$- D_{n+1}(x)\, P_n(x)\, P_{n-2}(x) - P_{n+1}(x)\left(D_{n-2}(x)\, P_{n-3}(x) + \phi(x)\, P'_{n-2}(x)\right).$$

Iterating this process, we obtain that for odd $k \le n$,

$$\phi(x)\, P'_{n+1}(x)\, P_{n-k}(x) = \left(\frac{C_{n+1}(x) - C_{n-k}(x)}{2}\right) P_{n+1}(x) P_{n-k}(x)$$
$$-D_{n+1}(x)\, P_n(x)\, P_{n-k}(x) + P_{n+1}(x)(D_{n-k}(x)\, P_{n-(k+1)}(x) + \phi(x)\, P'_{n-k}(x)),$$

and for even $k \le n$,

$$\phi(x)\, P'_{n+1}(x)\, P_{n-k}(x) = \left(\frac{C_{n+1}(x) + C_{n-k}(x) - 2C_0(x)}{2}\right) P_{n+1}(x)\, P_{n-k}(x)$$
$$- D_{n+1}(x)\, P_n(x)\, P_{n-k}(x) - P_{n+1}(x)\left(D_{n-k}(x)\, P_{n-(k+1)}(x) + \phi(x)\, P'_{n-k}(x)\right).$$

In either case, for n either odd or even, when $k = n$ we obtain (4.9), but this implies that $\mathbf{u}$ is semiclassical. □

Before dealing with the next result, we fix some notation. Let $\{Q_n(x)\}_{n\ge 0}$ be a basis of $\mathbb{P}$. We define the vector $\mathbf{Q} := (Q_0(x), Q_1(x), Q_2(x), \ldots)^t$. Let N be the semi-infinite matrix such that $\chi'(x) = N\, \chi(x)$. Therefore,

$$N = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & \ldots \\ 0 & 2 & 0 & 0 & \ldots \\ 0 & 0 & 3 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

We denote by $\tilde{N}$ the semi-infinite matrix such that $\mathbf{Q}' = \tilde{N}\, \mathbf{Q}$. Observe that if S is a matrix of change of basis from the monomials $\chi(x)$ to $\mathbf{Q}$, that is, $\mathbf{Q} = S\, \chi(x)$, then $\tilde{N} = S\, N\, S^{-1}$.

If $\mathbf{Q}$ is semiclassical, we write (4.9) in matrix form as $\phi(x)\, \mathbf{Q}' = F\, \mathbf{Q}$, where F is a semi-infinite band matrix. Finally, for square matrices A and B of size n, we define its commutator as $[A, B] = AB - BA$.

Proposition 4.13 *Let* $\mathbf{u}$ *be a positive-definite semiclassical functional satisfying the Pearson equation* $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$*, and let* $\{Q_n(x)\}_{n\geq 0}$ *be the sequence of orthonormal polynomials associated with* $\mathbf{u}$*. Then,*

1. $[J, F] = \phi(J)$,
2. $\tilde{N}\,\phi(J)^t + \phi(J)\,\tilde{N}^t = \psi(J)$,
3. $F + F^t = \psi(J)$,

where J *is the Jacobi matrix associated with* $\{Q_n(x)\}_{n\geq 0}$.

Remark 4.14 This is the matrix representation of the Laguerre-Freud equations satisfied by the parameters of the three-term recurrence relation of semiclassical orthonormal polynomials. As a direct consequence, you can deduce nonlinear difference equations that the coefficients of the three-term recurrence relation satisfy. They are relate to discrete Painlevé equations. Some illustrative examples appear in [73].

Proof

1. Differentiating $x\,\mathbf{Q} = J\,\mathbf{Q}$ and then multiplying by $\phi(x)$, we get

$$J\phi(x)\,\mathbf{Q}' = \phi(x)\,\mathbf{Q} + x\,\phi(x)\,\mathbf{Q}'.$$

But $\phi(x)\,\mathbf{Q}' = F\mathbf{Q}$ and $\phi(x)\mathbf{Q} = \phi(J)\mathbf{Q}$. Hence,

$$JF\mathbf{Q} = \phi(J)\mathbf{Q} + x\,F\,\mathbf{Q} = (\phi(J) + FJ)\mathbf{Q},$$

and, since $\mathbf{Q}$ is a basis, the result follows.

2. From the Pearson equation

$$\begin{aligned}
0 &= \langle D(\phi\mathbf{u}), \mathbf{Q}\,\mathbf{Q}^t\rangle + \langle \psi\,\mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle = -\langle \phi\,\mathbf{u}, \mathbf{Q}'\,\mathbf{Q}^t + \mathbf{Q}\,(\mathbf{Q}')^t\rangle + \langle \psi\,\mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle \\
&= -\tilde{N}\langle \mathbf{u}, \phi(x)\,\mathbf{Q}\,\mathbf{Q}^t\rangle - \langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\phi(x)\rangle\tilde{N}^t + \langle \mathbf{u}, \psi(x)\,\mathbf{Q}\,\mathbf{Q}^t\rangle \\
&= -\tilde{N}\phi(J)\langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle - \langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle\phi(J)^t\,\tilde{N}^t + \psi(J)\langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle.
\end{aligned}$$

But $\langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle$ is equal to the identity matrix since $\{Q_n(x)\}_{n\geq 0}$ are orthonormal, and the result follows.

3. Similarly, from the Pearson equation

$$\begin{aligned}
0 &= \langle D(\phi\mathbf{u}), \mathbf{Q}\,\mathbf{Q}^t\rangle + \langle \psi\,\mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle = -\langle \phi\,\mathbf{u}, \mathbf{Q}'\,\mathbf{Q}^t + \mathbf{Q}\,(\mathbf{Q}')^t\rangle + \langle \psi\,\mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle \\
&= -\langle \mathbf{u}, \phi(x)\,\mathbf{Q}'\,\mathbf{Q}^t\rangle - \langle \mathbf{u}, \mathbf{Q}\,(\mathbf{Q}')^t\phi(x)\rangle + \langle \mathbf{u}, \psi(x)\,\mathbf{Q}\,\mathbf{Q}^t\rangle \\
&= -F\langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle - \langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle F^t + \psi(J)\langle \mathbf{u}, \mathbf{Q}\,\mathbf{Q}^t\rangle,
\end{aligned}$$

and the result follows. □

5 Examples of Semiclassical Orthogonal Polynomials

It is well know that the semiclassical functionals of class $s = 0$ are the classical linear functionals (Hermite, Laguerre, Jacobi, and Bessel) defined by an expression of the form

$$\langle \mathbf{u}, p \rangle = \int_E p(x) w(x) dx, \quad \forall p \in \mathbb{P},$$

where

Family	$\phi(x)$	$\psi(x)$	$w(x)$	E
Hermite	1	$2x$	e^{-x^2}	$\mathbb{R}$
Laguerre	x	$x - \alpha - 1$	$x^\alpha e^{-x}$	$(0, +\infty)$
Jacobi	$x^2 - 1$	$-(\alpha + \beta + 2)x + \beta - \alpha$	$(1 - x)^\alpha (1 + x)^\beta$	$(-1, 1)$
Bessel	x^2	$-2(\alpha x + 1)$	$x^\alpha e^{-2/x}$	Unit circle

The Hermite, Laguerre, and Jacobi functionals are positive-definite when $\alpha, \beta > -1$, and the Bessel functional is a quasi-definite linear functional that is not positive-definite.

If $\mathbf{u}$ is a semiclassical functional of class $s = 1$, we can distinguish two situations

$$(A)\ \deg(\psi) = 2,\ 0 \le \deg(\phi) \le 3; \qquad (\mathrm{B})\ \deg(\psi) = 1,\ \deg(\phi) = 3.$$

S. Belmehdi [10] exposed the canonical forms of the functionals of the class 1, up to linear changes of the variable, according to the degree of $\phi(x)$ and the multiplicity of its zeros.

(A) $\deg(\psi) = 2$	
$\deg(\phi) = 0$	1
$\deg(\phi) = 1$	x
$\deg(\phi) = 2$	x^2 $x^2 - 1$
$\deg(\phi) = 3$	x^3 $x^2(x - 1)$ $(x^2 - 1)(x - c)$

(B) $\deg(\psi) = 1$	
$\deg(\phi) = 3$	x^3 $x^2(x - 1)$ $(x^2 - 1)(x - c)$

Example Let $\mathbf{u}$ be the linear functional defined by (see [40])

$$\langle \mathbf{u}, p \rangle = \int_0^\infty p(x)\, x^\alpha e^{-x} dx + Mp(0), \quad \forall p \in \mathbb{P},$$

with $\alpha > -1$ and $M > 0$. Then $\mathbf{u}$ is a semiclassical functional of class $s = 1$ satisfying $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ with $\phi(x) = x^2$ and $\psi(x) = x(x - \alpha - 2)$.

The sequence of polynomials orthogonal with respect to the above functional is known in the literature as Laguerre-type orthogonal polynomials (see [39, 49], among others).

Example Let $\mathbf{u}$ be the linear functional defined by (see [10])

$$\langle \mathbf{u}, p\rangle = \int_{-1}^{1} p(x)\,(x-1)^{(a+b-2)/2}(x+1)^{(b-a-2)/2}e^{ax}dx, \quad \forall p \in \mathbb{P},$$

with $b > a$. Then $\mathbf{u}$ satisfies $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ with $\phi(x) = x^2 - 1$ and $\psi(x) = -ax^2 - bx$. The functional is semiclassical of class $s = 1$.

Example Let $\mathbf{u}$ be the linear functional defined by (see [12, 73])

$$\langle \mathbf{u}, p\rangle = \int_{0}^{\infty} p(x)\,x^{\alpha}e^{-x^2+tx}dx, \quad \forall p \in \mathbb{P},$$

with $\alpha > -1$ and $t \in \mathbb{R}$. In [12], it is shown that $\mathbf{u}$ is a semiclassical functional of class $s = 1$ satisfying $D(\phi\,\mathbf{u})+\psi\,\mathbf{u} = 0$ with $\phi(x) = x$ and $\psi(x) = 2x^2-tx-\alpha-1$.

Example Let $\mathbf{u}$ be the functional defined by (see [10])

$$\langle \mathbf{u}, p\rangle = \int_{0}^{N} p(x)\,x^{\alpha}e^{-x}dx, \quad \forall p \in \mathbb{P},$$

with $\alpha > -1$ and $N > 0$. The functional $\mathbf{u}$ is semiclassical of class $s = 1$ satisfying $D(\phi\,\mathbf{u}) + \psi\,\mathbf{u} = 0$ with $\phi(x) = (x - N)x$ and $\psi(x) = (x - \alpha)(x - N) + N - 2x$.

This functional is known in the literature as truncated gamma functional and the corresponding sequences of orthogonal polynomials are called truncated Laguerre orthogonal polynomials.

Semiclassical functionals can be constructed via discrete Darboux transformations. First, we need to prove the following theorem.

Theorem 5.1 *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be two linear functionals related by*

$$A(x)\mathbf{u} = B(x)\mathbf{v},$$

where $A(x)$ *and* $B(x)$ *are non-zero polynomials. Then* $\mathbf{u}$ *is semiclassical if and only if* $\mathbf{v}$ *is semiclassical.*

Proof Suppose that $\mathbf{u}$ is semiclassical satisfying $D(\phi_0\mathbf{u}) + \psi_0\mathbf{u} = 0$. Let $\phi_1(x) = A(x)\,B(x)\phi_0(x)$. Then,

$$\begin{aligned}\langle D(\phi_1\,\mathbf{v}), x^n\rangle =& \langle D(A\,B\,\phi_0\,\mathbf{v}), x^n\rangle = \langle D(A^2\,\phi_0\,\mathbf{u}), x^n\rangle = -\langle \phi_0\,\mathbf{u}, n\,A^2\,x^{n-1}\rangle \\ =& -\langle \phi_0\mathbf{u}, (A^2\,x^n)'\rangle + \langle \phi_0\,\mathbf{u}, (A^2)'x^n\rangle\end{aligned}$$

$$
\begin{aligned}
&=\langle A^2\,\psi_0\,\mathbf{u}, x^n\rangle - \langle 2\phi_0\,A'\,A\mathbf{u}, x^n\rangle \\
&=\langle \big(A\,\psi_0 - 2A'\,\phi_0\big)\,B\,\mathbf{v}, x^n\rangle.
\end{aligned}
$$

Therefore, $\mathbf{v}$ is semiclassical with $\psi_1(x) = \big(A(x)\psi_0(x) - 2\phi_0(x)A'(x)\big)\,B(x)$.

Similarly, if $\mathbf{v}$ is semiclassical, by interchanging the role of the functionals above, it follows that $\mathbf{u}$ is semiclassical. □

Corollary 5.2 *Any linear spectral transformation of a semiclassical functional is also a semiclassical functional.*

Remark 5.3

- For canonical Christoffel (3.1) and Geronimus (3.3) transformations, the class of the new functional depends on the location of the point a in terms of the zeros of $\phi(x)$.
- Uvarov transformations (3.6) of classical orthogonal polynomials generate semiclassical linear functionals. The so called Krall-type linear functionals appear when a Dirac measure, or mass point, is located at a zero of $\phi(x)$. The corresponding sequences of orthogonal polynomials satisfy, for some choices of the parameters (in the Laguerre case, for α a non negative integer number) higher order linear differential equations with order depending on α. It is an open problem to describe the sequences of orthogonal polynomials which are eigenfunctions of higher order differential operators. For order two (S. Bochner [11]) and four (H. L. Krall [43]), the problem has been completely solved.

Example The linear functional obtained from a Uvarov transformation of the Laguerre functional will be of class 1 if a mass point is located at $a = 0$, and will be of class 2 if a mass point is located at $a \neq 0$. See [49].

Other examples of semiclassical functionals of class 2 are also known.

Example Let $\mathbf{u}$ be the functional defined by

$$
\langle \mathbf{u}, p\rangle = \int_{\mathbb{R}} p(x)\,e^{-\frac{x^4}{4}-tx^2}dx, \quad \forall p \in \mathbb{P},
$$

where $t \in \mathbb{R}$. In this case, $\mathbf{u}$ is a semi-classical functional of class $s = 2$, with $\phi(x) = 1$ and $\psi(x) = 2tx + x^3$.

This is a particular case of the so called generalized Freud linear functionals [19, 20].

New semiclassical functionals can also be constructed through symmetrized functionals [17].

Definition 5.4 Let $\mathbf{u}$ be a linear functional. Its symmetrized functional $\mathbf{v}$ is defined by

$$
\langle \mathbf{v}, x^{2n}\rangle = \mu_n, \quad \langle \mathbf{v}, x^{2n+1}\rangle = 0, \quad n \geq 0.
$$

Given a functional with Stieltjes function $\mathcal{S}(z)$, the Stieltjes function $\tilde{\mathcal{S}}(z)$ of its symmetrized functional satisfies $\tilde{\mathcal{S}}(z) = z\mathcal{S}(z^2)$. The following holds for the semiclassical case.

Theorem 5.5 ([5]) *Let* $\mathbf{u}$ *be a semiclassical functional satisfying* $D(\phi\mathbf{u}) + \psi\mathbf{u} = 0$*, and let* $\mathcal{S}(z)$ *be its Stieltjes function, which satisfies* (4.7)

$$\phi(z)\mathcal{S}'(z) + (\psi(z) + \phi'(z))\mathcal{S}(z) = C(z).$$

The Stieltjes function $\tilde{\mathcal{S}}(z)$ *associated with the symmetrized linear functional* $\mathbf{v}$ *satisfies*

$$z\phi(z^2)\tilde{\mathcal{S}}'(z) + [2z^2(\psi(z^2) + \phi'(z^2)) + \phi(z^2)]\tilde{\mathcal{S}}(z) = 2z^3C(z^2).$$

Thus, the symmetrized functional of a semiclassical linear functional is semiclassical. The class of $\mathbf{v}$ *is either* $2s$*,* $2s+1$*, or* $2s+3$*, according to the coprimality of the polynomial coefficients in the ordinary linear differential equation satisfied by* $\tilde{\mathcal{S}}(z)$*.*

6 Analytic Properties of Orthogonal Polynomials in Sobolev Spaces

An inner product is said to be a Sobolev inner product if

$$\langle f, g\rangle_S := \int_{E_0} f(x)\, g(x)\, d\mu_0 + \sum_{k=1}^{m} \int_{E_k} f^{(k)}(x)\, g^{(k)}(x)\, d\mu_k,$$

where $(d\mu_0, \ldots, d\mu_m)$ is a vector of positive Borel measures and $E_k = \operatorname{supp} d\mu_k$, $k = 0, 1, \ldots, m$.

Using the Gram-Schmidt orthogonalization method for the canonical basis $\{x^n\}_{n\geq 0}$, one gets a sequence of monic orthogonal polynomials. Thus, the nth orthogonal polynomial is a minimal polynomial in terms of the Sobolev norm

$$||f||_S := \sqrt{\langle f, f\rangle_S}$$

among all monic polynomials of degree n.

Taking into account that $\langle x\, f, g\rangle_S \neq \langle f, x\, g\rangle_S$, these polynomials do not satisfy a three-term recurrence relation. Thus, a basic property of standard orthogonal polynomials is lost. A natural question is to compare analytic properties of these polynomials and the standard ones.

In 1947, D. C. Lewis [45] dealt with the following problem in the framework of polynomial least square approximation. Let $\alpha_0, \ldots, \alpha_p$ be monotonic, nondecreasing functions defined on $[a, b]$ and let f be a function on $[a, b]$ that satisfies certain regularity conditions. Determine a polynomial $P_n(x)$ of degree at most n that minimizes

$$\sum_{k=0}^{p} \int_a^b |f^{(k)}(x) - P_n^{(k)}(x)|^2 d\alpha_k(x).$$

Lewis did not use Sobolev orthogonal polynomials and gave a formula for the remainder term of the approximation as an integral of the Peano kernel. The first paper on Sobolev orthogonal polynomials was published by Althammer [3] in 1962, who attributed his motivation to Lewis's paper. These Sobolev orthogonal polynomials are orthogonal with respect to the inner product

$$\langle f, g\rangle_S = \int_{-1}^{1} f(x)\, g(x)\, dx + \lambda \int_{-1}^{1} f'(x)\, g'(x)\, dx, \quad \lambda > 0.$$

Observe that the first and second integral of this inner product involve the Lebesgue measure dx on $[-1, 1]$, which means that every point in $[-1, 1]$ is equally weighted.

Let $S_n(x; \lambda)$ denote the orthogonal polynomial of degree n with respect to the inner product $\langle\cdot, \cdot\rangle_S$, normalized by $S_n(1; \lambda) = 1$, and let $P_n(x)$ denote the n-th Legendre polynomial. The following properties hold for $S_n(x; \lambda)$:

1. $\langle S_n(x; \lambda)\}_{n\geq 0}$ satisfies a differential equation

$$\lambda\, S_n''(x; \lambda) - S_n(x; \lambda) = A_n\, P_{n+1}'(x) + B_n\, P_{n-1}'(x),$$

   where A_n and B_n are real numbers which are explicitly given.
2. $\{S_n(x; \lambda)\}_{n\geq 0}$ satisfies a recursive relation

$$S_n(x; \lambda) - S_{n-2}(x; \lambda) = a_n\, (P_n(x) - P_{n-2}(x)), \quad n = 1, 2, \ldots$$

3. $S_n(x; \lambda)$ has n real simple zeros in $(-1, 1)$.

For a more detailed account on the development of these results, we refer to [45, 62, 67]. The Sobolev-Legendre polynomials were also studied by Gröbner, who established a version of the Rodrigues formula in [33]. Indeed, he states that, up to a constant factor c_n,

$$S_n(x; \lambda) = c_n \frac{D^n}{1 - \lambda D^2} \left((x^2 - x)^n - \alpha_n (x^2 - x)^{n-1}\right)$$

where α_n are real numbers explicitly given in terms of λ and n.

In [3], Althammer also gave an example in which he replaced dx in the second integral in $\langle\cdot, \cdot\rangle_S$ by $w(x)dx$ with $w(x) = 10$ for $-1 \leq x < 0$ and $w(x) = 1$ for $0 \leq x \leq 1$, and made the observation that $S_2(x; \lambda)$ for this new inner product has one real zero outside of $(-1, 1)$.

In [13], Brenner considered the inner product

$$\langle f, g\rangle := \int_0^\infty f(x)\, g(x)\, e^{-x} dx + \lambda \int_0^\infty f'(x)\, g'(x)\, e^{-x} dx, \quad \lambda > 0,$$

and obtained results in a direction very similar to those of Althammer. Sobolev inner products when you replace the above weight by $x^\alpha e^{-x}$, $\alpha \geq -1$ has been studied in [51].

An important contribution in the early development of the Sobolev polynomials was made in 1972 by Schäfke and Wolf in [68], where they considered a family of inner products

$$\langle f, g\rangle_S = \sum_{j,k=0}^{\infty} \int_a^b f^{(j)}(x)\, g^{(k)}(x)\, v_{j,k}(x)\, w(x)\, dx, \tag{6.1}$$

where w and (a, b) are one of the three classical cases (Hermite, Laguerre, and Jacobi) and the functions $v_{j,k}$ are polynomials that satisfy $v_{j,k} = v_{k,j}$, $k = 0, 1, 2, \ldots$, and allow to write the inner product (6.1) as

$$\langle f, g\rangle_S = \int_a^b f(x)\, \mathcal{B}g(x)\, w(x)\, dx, \quad \text{with} \quad \mathcal{B}g := w^{-1} \sum_{j,k=0}^{\infty} (-1)^j D^j\, (w\, v_{j,k}\, D^k) g$$

by using an integration by parts. Under further restrictions on $v_{j,k}$, they are narrowed down to eight classes of Sobolev orthogonal polynomials, which they call simple generalizations of classical orthogonal polynomials.

The primary tool in the early study of Sobolev orthogonal polynomials is integration by parts. Schäfke and Wolf [68] explored when this tool is applicable and outlined potential Sobolev inner products. It is remarkable that their work appeared in such an early stage of the development of Sobolev orthogonal polynomials.

The study of Sobolev orthogonal polynomials unexpectedly became largely dormant for nearly two decades, from which it reemerged only when a new ingredient, *coherent pairs*, was introduced in [36].

6.1 Coherent Pairs of Measures and Sobolev Orthogonal Polynomials

The concept of coherent pair of measures was introduced in [36] in the framework of the study of the inner product

$$\langle f, g\rangle_\lambda = \int_a^b f(x)\, g(x)\, d\mu_0(x) + \lambda \int_a^b f'(x)\, g'(x)\, d\mu_1(x), \tag{6.2}$$

where $-\infty \le a < b \le \infty$, $\lambda \ge 0$, μ_0 and μ_1 are positive Borel measures on the real line with finite moments of all orders. Let $P_n(x; d\mu_i)$ denote the monic orthogonal polynomial of degree n with respect to $d\mu_i$, $i = 0, 1$.

Definition 6.1 The pair $\{d\mu_0, \ d\mu_1\}$ is called coherent if there exists a sequence of nonzero real numbers $\{\alpha_n\}_{n\ge 1}$ such that

$$P_n(x; d\mu_1) = \frac{P'_{n+1}(x; d\mu_0)}{n+1} + \alpha_n \frac{P'_n(x; d\mu_0)}{n}, \quad n \ge 1. \tag{6.3}$$

If $[a, b] = [-c, c]$ and $d\mu_0$ and $d\mu_1$ are both symmetric, then $\{d\mu_0, d\mu_1\}$ is called a symmetrically coherent pair if

$$P_n(x; d\mu_1) = \frac{P'_{n+1}(x; d\mu_0)}{n+1} + \alpha_n \frac{P'_{n-1}(x; d\mu_0)}{n-1}, \quad n \ge 2.$$

If $d\mu_1 = d\mu_0$, the measure $d\mu_0$ is said to be self-coherent (resp. symmetrically self-coherent).

For $n = 0, 1, 2, \ldots$, let

$$M_n(\lambda) = \begin{pmatrix} \langle 1, 1\rangle_\lambda & \langle 1, x\rangle_\lambda & \cdots & \langle 1, x^n\rangle_\lambda \\ \langle x, 1\rangle_\lambda & \langle x, x\rangle_\lambda & \cdots & \langle x, x^n\rangle_\lambda \\ \ldots & \ldots & \ddots & \ldots \\ \langle x^n, 1\rangle_\lambda & \langle x^n, x\rangle_\lambda & \cdots & \langle x^n, x^n\rangle_\lambda \end{pmatrix}.$$

Since $\det M_n(\lambda) > 0$ for all $n \ge 0$, then a sequence of monic orthogonal polynomials with respect to $\langle\cdot, \cdot\rangle_\lambda$ exists. Let $\{S_n(x; \lambda)\}_{n\ge 0}$ denote the sequence of monic Sobolev orthogonal polynomials with respect to $\langle\cdot, \cdot\rangle_\lambda$. In fact, the monic orthogonal polynomials are $S_0(x; \lambda) = 1$ and, for $n \ge 1$,

$$S_n(x; \lambda) = \frac{1}{\det M_{n-1}(\lambda)} \det \left(\begin{array}{cccc|c} & & & & \langle 1, x^n\rangle_\lambda \\ & & & & \langle x, x^n\rangle_\lambda \\ & M_{n-1}(\lambda) & & & \vdots \\ & & & & \langle x^{n-1}, x^n\rangle_\lambda \\ \hline 1 & x & \cdots & x^{n-1} & x^n \end{array}\right).$$

It is easy to see that

$$T_n(x) := \lim_{\lambda\to\infty} S_n(x; \lambda)$$

is a monic polynomial of degree n which satisfies

$$T_n'(x) = n\, P_{n-1}(x; d\mu_1) \quad \text{and} \quad \int_{\mathbb{R}} T_n(x)\, d\mu_0 = 0 \quad n \geq 1. \tag{6.4}$$

Theorem 6.2 ([36]) *If $\{d\mu_0,\ d\mu_1\}$ is a coherent pair, then*

$$S_n(x; \lambda) + \beta_{n-1}(\lambda)\, S_{n-1}(x; \lambda) = P_n(x; d\mu_0) + \hat{\alpha}_{n-1} P_{n-1}(x; d\mu_0),\ n \geq 2, \tag{6.5}$$

where $\hat{\alpha}_{n-1} = n\,\alpha_n/(n-1)$ and $\beta_{n-1}(\lambda) = \hat{\alpha}_{n-1}||P_{n-1}(x; d\mu_0)||^2_{d\mu_0}/||S_{n-1}(x; \lambda)||^2_\lambda$.

Proof According to (6.3) and (6.4), we see that

$$T_n(x) = P_n(x; d\mu_0) + \hat{\alpha}_{n-1} P_{n-1}(x; d\mu_0).$$

For $0 \leq j \leq n-2$, it follows from (6.4) that

$$\langle T_n(x), S_j(x; \lambda)\rangle_\lambda = \langle T_n(x), S_j(x; \lambda)\rangle_{d\mu_0} + n\lambda\langle P_{n-1}(x; d\mu_1), S_j'(x; \lambda)\rangle_{d\mu_1} = 0.$$

Considering the expansion of $T_n(x)$ in terms of the polynomials $S_j(x; \lambda)$, we see that

$$T_n(x) = S_n(x; \lambda) + \beta_{n-1}(\lambda)\, S_{n-1}(x; \lambda), \quad \text{where} \quad \beta_{n-1}(\lambda) = \frac{\langle T_n(x), S_{n-1}(x; \lambda)\rangle_\lambda}{||S_{n-1}(x; \lambda)||^2_\lambda}.$$

The expresion for $\beta_{n-1}(\lambda)$ follows from $\langle T_n'(x), S_{n-1}'(x; \lambda)\rangle_{d\mu_1} = 0$ as well as from the fact that both $P_{n-1}(x; d\mu_0)$ and $S_{n-1}(x; \lambda)$ are monic. □

The notion of coherent pairs can be extended to linear functionals $\{\mathbf{u}_0,\ \mathbf{u}_1\}$, if the relation (6.3) holds with $P_n(x; d\mu_i)$ replaced by $P_n(x; \mathbf{u}_i)$.

The following theorem was established in [63].

Theorem 6.3 *If $\{d\mu_0,\ d\mu_1\}$ is a coherent pair of measures, then at least one of them has to be classical (Laguerre, Jacobi).*

Together, [52, 63] give a complete list of coherent pairs. In the case when $\mathbf{u}_0$ and $\mathbf{u}_1$ are positive-definite linear functionals associated with measures $d\mu_0$ and $d\mu_1$, the coherent pairs are given as follows:

Laguerre Case

(1) $d\mu_0(x) = (x - \xi)x^{\alpha-1}e^{-x}dx$ and $d\mu_1(x) = x^\alpha e^{-x}dx$, where if $\xi < 0$, then $\alpha > 0$, and if $\xi = 0$ then $\alpha > -1$.

(2) $d\mu_0(x) = x^\alpha e^{-x}dx$ and $d\mu_1(x) = (x - \xi)^{-1}x^{\alpha+1}e^{-x}dx + M\,\delta(x - \xi)$, where if $\xi < 0$, $\alpha > -1$ and $M \geq 0$.

(3) $d\mu_0(x) = e^{-x}dx + M\delta(x)$ and $d\mu_1(x) = e^{-x}dx$, where $M \geq 0$.

Jacobi Case

(1) $d\mu_0(x) = |x-\xi|(1-x)^{\alpha-1}(1+x)^{\beta-1}dx$ and $d\mu_1(x) = (1-x)^{\alpha}(1+x)^{\beta}dx$, where if $|\xi| > 1$ then $\alpha > 0$ and $\beta > 0$, if $\xi = 1$ then $\alpha > -1$ and $\beta > 0$, and if $\xi = -1$ then $\alpha > 0$ and $\beta > -1$.
(2) $d\mu_0(x) = (1-x)^{\alpha}(1+x)^{\beta}dx$ and $d\mu_1(x) = |x-\xi|^{-1}(1-x)^{\alpha+1}(1+x)^{\beta+1}dx + M\delta(x-\xi)$, where $|\xi| > 1, \alpha > -1$ and $\beta > -1$ and $M \geq 0$.
(3) $d\mu_0(x) = (1+x)^{\beta-1}dx + M\delta(x-1)$ and $d\mu_1(x) = (1+x)^{\beta}dx$, where $\beta > 0$ and $M \geq 0$.
(4) $d\mu_0(x) = (1-x)^{\alpha-1}dx + M\delta(x+1)$ and $d\mu_1(x) = (1-x)^{\alpha}dx$, where $\alpha > 0$ and $M \geq 0$.

A similar analysis was also carried out for symmetrically coherent pairs in the work cited above. It lead to the following list of symmetrically coherent pairs.

Hermite Case

(1) $d\mu_0(x) = e^{-x^2}dx$ and $d\mu_1(x) = (x^2+\xi^2)^{-1}e^{-x^2}dx$, where $\xi \neq 0$.
(2) $d\mu_0(x) = (x^2+\xi^2)e^{-x^2}dx$ and $d\mu_1(x) = e^{-x^2}dx$, where $\xi \neq 0$.

Gegenbauer Case

(1) $d\mu_0(x) = (1-x^2)^{\alpha-1}dx$ and $d\mu_1(x) = (x^2+\xi^2)^{-1}(1-x^2)^{\alpha}dx$, where $\xi \neq 0$ and $\alpha > 0$.
(2) $d\mu_0(x) = (1-x^2)^{\alpha-1}dx$ and $d\mu_1(x) = (\xi^2-x^2)^{-1}(1-x^2)^{\alpha}dx + M\delta(x-\xi) + M\delta(x+\xi)$, where $|\xi| > 1, \alpha > 0$ and $M \geq 0$.
(3) $d\mu_0(x) = (x^2+\xi^2)(1-x^2)^{\alpha-1}dx$ and $d\mu_1(x) = (1-x^2)^{\alpha}dx$, where $\alpha > 0$.
(4) $d\mu_0(x) = (\xi^2-x^2)(1-x^2)^{\alpha-1}dx$ and $d\mu_1(x) = (1-x^2)^{\alpha}dx$, where $|\xi| \geq 1$ and $\alpha > 0$.
(5) $d\mu_0(x) = dx + M\delta(x-1) + M\delta(x+1)$ and $d\mu_1(x) = dx$, where $M \geq 0$.

6.1.1 Generalized Coherent Pairs

Identity (6.5) was deduced from definition (6.3) of coherent pairs. In the reverse direction, however, (6.3) does not follow from the identity (6.5), as observed in [38].

Let $\mathtt{S}_n(x)$ denote the left hand side of (6.5). Clearly $\mathtt{S}_n'$ can be expanded in terms of $\{P_k(x; d\mu_1)\}_{k\geq 0}$,

$$\mathtt{S}_n'(x) = n\,P_{n-1}(x; d\mu_1) + \sum_{k=0}^{n-2} d_{k,n}P_k(x; d\mu_1), \qquad d_{k,n} = \frac{\langle \mathtt{S}_n'(x), P_k(x; d\mu_1)\rangle_{d\mu_1}}{||P_k(x; d\mu_1)||^2_{d\mu_1}}, \qquad n \geq 1.$$

For $0 \leq j \leq n-2$, it follows directly from the definition of $\mathtt{S}_n$ that

$$\langle \mathtt{S}_n(x), P_j(x; d\mu_1)\rangle_{\lambda} = 0,$$

and it follows from (6.5) that $\langle \mathtt{S}_n(x), P_j(x; d\mu_1)\rangle_{d\mu_0} = 0$. Consequently, by the definition of $\langle\cdot,\cdot\rangle_\lambda$ we must have $\langle \mathtt{S}'_n, P_j(x; d\mu_1)\rangle_{d\mu_1} = 0$ for $0 \le j \le n-2$, which implies that $d_{k,n} = 0$ if $0 \le k \le n-2$. Hence,

$$\mathtt{S}'_n(x) = P'_n(x; d\mu_0) + \hat{a}_{n-1} P'_{n-1}(x; d\mu_0) = n\, P_{n-1}(x; d\mu_1) + d_{n-2,n} P_{n-2}(x; d\mu_1).$$

Recall that $\hat{\alpha}_n = (n+1)\alpha_n/n$. Setting $\beta_{n-2} = d_{n-2,n}/n$ and shifting the index from n to $n+1$, we conclude the following relation between $\{P_n(x; d\mu_0)\}_{n\ge 0}$ and $\{P_n(x; d\mu_1)\}_{n\ge 0}$,

$$P_n(x; d\mu_1) + \beta_{n-1}\, P_{n-1}(x; d\mu_1) = \frac{P'_{n+1}(x; d\mu_0)}{n+1} + \alpha_n \frac{P'_n(x; d\mu_0)}{n}, \quad n \ge 1. \tag{6.6}$$

Thus, in the reverse direction, (6.5) leads to (6.6) instead of (6.3).

Evidently, (6.6) is a more general relation than (6.3).

Definition 6.4 The pair $\{d\mu_0, d\mu_1\}$ is called a generalized coherent pair if (6.6) holds for $n \ge 1$, and this definition extends to linear functionals $\{\mathbf{u}_0, \mathbf{u}_1\}$.

Semiclassical orthogonal polynomials of class 1 (see Sect. 5) are involved in the analysis of generalized coherent pairs. The following theorem is established in [22].

Theorem 6.5 *If* $\{\mathbf{u}_0, \mathbf{u}_1\}$ *is a generalized coherent pair, then at least one of them must be semiclassical of class at most* 1.

All generalized coherent pairs of linear functionals are listed in [22].

On the other hand, given two sequences of monic orthogonal polynomials $\{P_n(x; d\mu_0)\}_{n\ge 0}$ and $\{P_n(x; d\mu_1)\}_{n\ge 0}$, where $d\mu_0$ and $d\mu_1$ are symmetric measures, such that the following relation holds

$$P_{n+1}(x; d\mu_1) + \beta_{n-1}\, P_{n-1}(x; d\mu_1) = \frac{P'_{n+2}(x; d\mu_0)}{n+1} + \alpha_n \frac{P'_n(x; d\mu_0)}{n}, \quad n \ge 1. \tag{6.7}$$

We introduce the following

Definition 6.6 The pair $\{d\mu_0, d\mu_1\}$ is called a symmetrically generalized coherent pair if (6.7) holds for $n \ge 1$, and this definition extends to linear functionals $\{\mathbf{u}_0, \mathbf{u}_1\}$.

Some examples of symmetrically generalized coherent pairs have been studied in [21], where $\mathbf{u}_0$ is associated with the Gegenbauer weight and

$$\mathbf{u}_1 = \frac{1-x^2}{1+qx^2}\mathbf{u}_1 + M_q\left[\delta(x+1/\sqrt{-q}) + \delta(x-1/\sqrt{-q})\right], \quad q \ge -1,$$

where $M_q \ge 0$ if $-1 \le q < 0$ and $M_q = 0$ if $q \ge 0$.

More recently, in [24] the authors obtain analytic properties of Sobolev orthogonal polynomials with respect to a symmetrically generalized coherent pair $\{\mathbf{u}_0, \mathbf{u}_1\}$, where $\mathbf{u}_0$ is the linear functional associated with the Hermite weight and $\mathbf{u}_1 = \frac{x^2+a^2}{x^2+b^2}\mathbf{u}_0$.

6.2 Sobolev-Type Orthogonal Polynomials

An inner product is said to be a Sobolev-type inner product if the derivatives appear only as function evaluations on a finite discrete set. More precisely, such an inner product takes the form

$$\langle f, g\rangle_S := \int_{\mathbb{R}} f(x)\, g(x)\, d\mu_0 + \sum_{k=1}^{m} \int_{\mathbb{R}} f^{(k)}(x)\, g^{(k)}(x)\, d\mu_k, \tag{6.8}$$

where $d\mu_0$ is a positive Borel measure on an infinite subset of the real line and $d\mu_k$, $k = 1, 2, \ldots, m$, are positive Borel measures supported on finite subsets of the real line. In most cases considered below, $d\mu_k = A_k\delta(x-c)$ or $d\mu_k = A_k\delta(x-a) + B_k\delta(x-b)$, where A_k and B_k are nonnegative real numbers. Orthogonal polynomials for such an inner product are called Sobolev-type orthogonal polynomials.

The first study was carried out for the classical weight functions. The Laguerre case was studied in [40, 41] with $d\mu_0 = x^\alpha e^{-x} dx$, $\alpha > -1$, and

$$d\mu_k = M_k\delta(x), \quad k = 1, 2, \ldots, m,$$

the nth Sobolev orthogonal polynomial, S_n, is given by

$$S_n(x) = \sum_{k=0}^{\min\{n,m+1\}} (-1)^k\, A_{n,k}\, L_{n-k}^{\alpha+k}(x), \quad n \geq 1,$$

where $A_{n,k}$ are real numbers determined by a linear system of equations. The Gegenbauer case was studied in [7, 8] with $d\mu_0 = (1-x^2)^{\lambda-1/2}dx + A(\delta(x-1) + \delta(x+1))$, $\lambda > -1/2$, and $m = 1$, $d\mu_1 = B(\delta(x-1) + \delta(x+1))$; the nth Sobolev orthogonal polynomial is given by

$$S_n(x) = a_{0,n}C_n^\lambda(x) + a_{1,n}x\, C_{n-1}^{\lambda+1}(x) + a_{2,n}x^2\, C_{n-2}^{\lambda+2}(x), \quad n \geq 2,$$

where $a_{0,n}$, $a_{1,n}$, and $a_{2,n}$ are appropriate real numbers. In both cases, the Sobolev orthogonal polynomials satisfy higher order (greater than three) recurrence relations.

When $M_k = 0$ for $k = 1, 2, \ldots, m-1$, and $d\mu_m = M_m\, \delta(x-c)$, the inner product (6.8) becomes

$$\langle f, g\rangle_m := \int_{\mathbb{R}} f(x)\, g(x)\, d\mu_0 + M_m\, f^{(m)}(c)\, g^{(m)}(c),$$

where $c \in \mathbb{R}$ and $M_m \geq 0$.

For $i, j \in \mathbb{N}_0$, define

$$\mathcal{K}_{n-1}^{(i,j)}(x, y) := \sum_{l=0}^{n-1} \frac{P_l^{(i)}(x)\, P_l^{(j)}(y)}{||P_l||_{d\mu_0}^2}, n \geq 1.$$

It was shown in [53] that

$$S_n(x) = P_n(x) - \frac{M_m\, P_n^{(m)}(c)}{1 + M_m\, \mathcal{K}_{n-1}^{(m,m)}(c, c)} \mathcal{K}_{n-1}^{(0,m)}(x, c), n \geq 1,$$

which extends the expression for $m = 0$ by A. M. Krall in [44]. From this relation, one deduces immediately that

$$S_{n+1}(x) + \alpha_n S_n(x) = P_{n+1}(x) + \beta_n P_n(x), \quad n \geq 0,$$

where α_n and β_n are constants that can be easily determined. This shows a similar structure to (6.5) derived for the Sobolev orthogonal polynomials in the case of coherent pairs.

The Sobolev polynomials $S_n(x)$ also satisfy a higher order recurrence relation

$$(x-c)^{m+1} S_n(x) = \sum_{j=n-m-1}^{n+m+1} c_{n,j}\, S_j(x), n \geq 0, \tag{6.9}$$

where $c_{n,n+m+1} = 1$ and $c_{n,n-m-1} \neq 0$.

If a sequence of polynomials satisfies a three-term recurrence relation, then it is orthogonal. The precise statement is known as Favard's theorem. For higher order recurrence relations, there are two types of results in this direction, both related to Sobolev orthogonal polynomials.

The first one gives a characterization of an inner product $\langle \cdot, \cdot \rangle$ for which the corresponding sequence of orthogonal polynomials satisfy a recurrence relation (6.9), which holds if the operation of multiplication by $M_{m,c} := (x-c)^{m+1}$ is symmetric, that is, $\langle M_{m,c}\, p, q\rangle = \langle p, M_{m,c}\, q\rangle$. It was proved in [26] that if $\langle \cdot, \cdot \rangle$ is an inner product such that $M_{m,c}$ is symmetric and it communtes with the operator $M_{0,c}$, that is, $\langle M_{m,c}\, p, M_{0,c}\, q\rangle = \langle M_{0,c}\, p, M_{m,c}\, q\rangle$, then there exists a nontrivial positive Borel measure $d\mu_0$ and a real, positive semi-definite matrix A of size $m+1$, such that the inner product is of the form

$$\langle p, q\rangle = \int_{\mathbb{R}} p(x)\, q(x)\, d\mu_0 + \left(p(c), p'(c), \dots, p^{(m)}(c)\right) A \left(q(c), q'(c), \dots, q^{(m)}(c)\right)^t. \tag{6.10}$$

A connection between such Sobolev orthogonal polynomials and matrix orthogonal polynomials was established in [27], by representing the higher order recurrence relation as a three-term recurrence relation with matrix coefficients for a family of matrix orthogonal polynomials defined in terms of the Sobolev orthogonal polynomials.

The second type of Favard type theorem was given in [28], where it was proved that the operator of multiplication by a polynomial h is symmetric with respect to the inner product (6.8) if and only if $d\mu_k$, $k = 1, 2, \dots, m$, are discrete measures whose supports are related to the zeros of h and its derivatives. Consequently, higher order recurrence relations for Sobolev inner products appear only in Sobolev inner products of the second type.

6.3 Asymptotics of Sobolev Orthogonal Polynomials

For standard orthogonal polynomials, three different types of asymptotics are considered: strong asymptotics, outer ratio asymptotics, and nth root asymptotics. All three have been considered in the Sobolev setting and we summarize the most relevant results in this section.

The first work on asymptotics for Sobolev orthogonal polynomials is [54] where the authors deal with the inner product

$$\langle f, g\rangle_S = \int_{-1}^{1} f(x)\, g(x)\, d\mu_0(x) + M_1\, f'(c)\, g'(c),$$

where $c \in \mathbb{R}$, $M_1 > 0$, and the measure $d\mu_0$ belongs to the Nevai class $M(0, 1)$. Using the outer ratio asymptotics for the ordinary orthogonal polynomials $P_n(x; d\mu_0)$ and the connection formula between $P_n(x; d\mu_0)$ and the Sobolev orthogonal polynomials $S_n(x)$, it was shown that, if $c \in \mathbb{R} \setminus \operatorname{supp} \mu_0$, then

$$\lim_{n\to\infty} \frac{S_n(z)}{P_n(z, d\mu_0)} = \frac{(\Phi(z) - \Phi(c))^2}{2\,\Phi(z)\,(z - c)}, \qquad \Phi(z) := z + \sqrt{z^2 - 1},$$

locally uniformly outside the support of the measure, where $\sqrt{z^2-1} > 0$ when $z > 1$. If $c \in \operatorname{supp} \mu_0$, then

$$\lim_{n\to\infty} \frac{S_n(z)}{P_n(z; d\mu_0)} = 1$$

outside the support of the measure.

The first extension of the above results was carried out in [1] for the Sobolev inner product (6.10) with a 2×2 matrix A. Under the same conditions on the measure, it was proved that

$$\lim_{n\to\infty} \frac{S_n(z)}{P_n(z; d\mu_0)} = \left(\frac{(\Phi(z)-\Phi(c))^2}{2\,\Phi(z)\,(z-c)}\right)^r, \qquad r := \operatorname{rank} A,$$

locally uniformly outside the support of the measure.

The second extension was given in [48] for the inner product

$$\langle f, g\rangle = \int_{\mathbb{R}} f(x)\, g(x)\, d\mu_0(x) + \sum_{j=1}^{N} \sum_{k=0}^{N_j} f^{(k)}(c_j)\, L_{j,k}(g; c_j),$$

where $d\mu_0 \in M(0,1)$, $\{c_k\}_{k=1}^{N} \in \mathbb{R}\backslash \operatorname{supp} \mu_0$, $j = 1, \ldots, N$, and $L_{j,k}$ is an ordinary linear differential operator acting on g such that L_{j,N_j} is not identically zero for $j = 1, \ldots, N$. Assuming that the inner product is quasi-definite so that a sequence of orthogonal polynomials exists, then on every compact subset in $\mathbb{C} \setminus \operatorname{supp} d\mu_0$,

$$\lim_{n\to\infty} \frac{S_n^{(\nu)}(z)}{P_n^{(\nu)}(z, d\mu_0)} = \prod_{j=1}^{m} \left(\frac{(\Phi(z)-\Phi(c))^2}{2\,\Phi(z)\,(z-c)}\right)^{I_j},$$

where I_j is the dimension of the square matrix obtained from the matrix of coefficients of L_{j,N_j} after deleting all zero rows and columns.

On the other hand, if both the measure $d\mu_0$ and its support Δ are regular, then techniques from potential theory can be used (see [47]) to derive the nth root asymptotics of the Sobolev orthogonal polynomials,

$$\limsup_{n\to\infty} ||S_n^{(j)}||_{\Delta}^{1/n} = C(\Delta), \qquad j \geq 0,$$

where $||\cdot||_{\Delta}$ denotes the uniform norm on the support of the measure and $C(\Delta)$ is its logarithmic capacity.

When the support of the measure in the inner product (6.10) is unbounded, the analysis has been focused on the case of the Laguerre weight function. A first study [4] considered the case when $c = 0$ and A is a 2×2 diagonal matrix (see also [50] for a survey of the unbounded case). Assuming that the leading coefficients of S_n

are standardized to be $(-1)^n/n!$, the following results on the asymptotic behavior of S_n were established:

(1) (Outer relative asymptotics) $\lim_{n\to\infty} \dfrac{S_n(z)}{L_n^{(\alpha)}(z)} = 1$ uniformly on compact subsets of the exterior of the positive real semiaxis.
(2) (Outer relative asymptotics for scaled polynomials) $\lim_{n\to\infty} \dfrac{S_n(nz)}{L_n^{(\alpha)}(nz)} = 1$ uniformly on compact subsets of the exterior of $[0, 4]$.
(3) (Mehler-Heine formula) $\lim_{n\to\infty} n^{-\alpha}\, S_n(z/n) = z^{-\alpha/2} J_{\alpha+4}(2\sqrt{z})$ uniformly on compact subsets of the complex plane, assuming that rank $A = 2$.
(4) (Inner strong asymptotics)

$$\frac{S_n(x)}{n^{\alpha/2}} = c_3(n) e^{x/2} x^{-\alpha/2} J_{\alpha+4}\left(2\sqrt{(n-2)x}\right) + O\left(n^{-\min\{\alpha+5,3/4\}}\right)$$

on compact subsets of the positive real semiaxis, where $\lim_{n\to\infty} c_3(n) = 1$.

If the point c is a negative real number, then the following outer relative asymptotics was established in [55],

$$\lim_{n\to\infty} \frac{S_n(z)}{L_n^{(\alpha)}(z)} = \left(\frac{\sqrt{-z} - \sqrt{-c}}{\sqrt{-z} + \sqrt{-c}}\right)^r, \qquad r = \operatorname{rank} A,$$

uniformly on compact subsets of the exterior of the real positive semiaxis.

When $c = 0$ and A is a non-singular diagonal matrix of size $m + 1$, the following asymptotic properties of the Sobolev orthogonal polynomials with respect to the inner product (6.10) were obtained in [2]:

(1) (Outer relative asymptotics) For every $\nu \in \mathbb{N}$, $\lim_{n\to\infty} \dfrac{S_n^{(\nu)}(z)}{(L_n^{(\alpha)})^{(\nu)}(z)} = 1$ uniformly on compact subsets of the exterior of the positive real semiaxis.
(2) (Mehler-Heine formula)

$$\lim_{n\to\infty} \frac{(-1)^n}{n!} \frac{S_n(z/n)}{n^{\alpha}} = (-1)^{m+1} z^{-\alpha/2} J_{\alpha+2m+2}(2\sqrt{z})$$

uniformly on compact subsets of the complex plane.

6.3.1 Continuous Sobolev Inner Products

Let $\{\mu_0,\ \mu_1\}$ be a coherent pair of measures and supp $\mu_0 = [-1, 1]$. Then the outer relative asymptotic relation for the Sobolev orthogonal polynomials with respect to (6.10) in terms of the orthogonal polynomials $P_n(x; d\mu_1)$ is (see [61])

$$\lim_{n\to\infty} \frac{S_n(z)}{P_n(z; d\mu_1)} = \frac{2}{\Phi'(z)}, \qquad \Phi(z) := z + \sqrt{z^2 - 1},$$

where $\sqrt{z^2-1} > 0$ when $z > 1$, uniformly on compact subsets of the exterior of the interval $[-1, 1]$.

When the measures μ_0 and μ_1 are absolutely continuous and belong to the Szegő class, the above result is also true [60].

For measures of coherent pairs that have unbounded support, asymptotic properties of the corresponding Sobolev orthogonal polynomials have been extensively studied in the literature (see [50] for an overview). The outer relative asymptotics, the scaled outer asymptotics, as well as the inner strong asymptotics of such polynomials have been considered for all families of coherent pairs and symmetrically coherent pairs.

The case where both measures in (6.2) correspond to the Freud weight, that is, $d\mu_0 = d\mu_1 = e^{-x^4}dx$, was studied in [15] (see also [30]), where the connection between Sobolev and standard orthogonal polynomials is given by

$$P_n(x; d\mu) = S_n(x; \lambda) + c_{n-2}(\lambda)S_{n-2}(x; \lambda), \quad n \ge 2.$$

7 Sobolev Orthogonal Polynomials of Several Variables

In contrast with the univariate case, Sobolev orthogonal polynomials of several variables have been studied only recently. In this section, we collect some results in this direction.

7.1 *Orthogonal Polynomials of Several Variables*

For $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ and $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}_0^d$, the (total) degree of the monomial

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$$

is, by definition, $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_d$. Let Π_n^d denote the linear space of polynomials in d variables of total degree at most n. It is known that $\dim \Pi_n^d = \binom{n+d}{n}$. Let $\Pi^d := \bigcup_{n\ge0} \Pi_n^d$ denote the space of all polynomials in d variables.

Let $\langle \cdot, \cdot \rangle$ be an inner product defined on $\Pi^d \times \Pi^d$. A polynomial $P \in \Pi_n^d$ is orthogonal if

$$\langle P, q \rangle = 0, \qquad \forall q \in \Pi_{n-1}^d.$$

For $n \in \mathbb{N}_0$, let $\mathcal{V}_n^d$ denote the space of orthogonal polynomials of total degree n. Then $\dim \mathcal{V}_n^d = \binom{n+d-1}{n}$. In contrast with the univariate case, the space $\mathcal{V}_n^d$ can have many different bases when $d \geq 2$. Moreover, the elements of $\mathcal{V}_n^d$ may not be mutually orthogonal.

For the structure and properties of orthogonal polynomials of several variables, we refer to [25]. We describe briefly a family of orthogonal polynomials as example.

A polynomial Y is said to be a spherical harmonic of degree n if it is a homogeneous polynomial such that $\Delta Y = 0$, where Δ is the Laplacian operator,

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}.$$

Let $\mathcal{H}_n^d$ denote the space of spherical harmonics of degree n. It is known that

$$a_n^d := \dim \mathcal{H}_n^d = \binom{n+d-1}{n} - \binom{n+d-3}{n-2}.$$

The elements of $\mathcal{H}_n^d$ are orthogonal with respect to polynomials of degree at most $n-1$ with respect to the inner product

$$\langle f, g \rangle_{\mathbb{S}^{d-1}} := \int_{\mathbb{S}^{d-1}} f(\xi)\, g(\xi)\, d\sigma(\xi),$$

where $d\sigma$ denotes the surface measure on $\mathbb{S}^{d-1}$.

For $\mu > -1$, let $w_\mu(x) = (1 - ||x||^2)^{\mu - 1/2}$ be the weight function defined on the unit ball $\mathbb{B}^d = \{x \in \mathbb{R}^d : \ ||x|| \leq 1\}$, where $|| \cdot ||$ denotes the Euclidean norm in $\mathbb{R}^d$. Orthogonal polynomials with respect to w_μ can be given in several different formulations. We give one basis of $\mathcal{V}_n^d(w_\mu)$ in terms of the classical Jacobi polynomials and spherical harmonics in the spherical coordinates $x = r\xi$, where $0 < r \leq 1$ and $\xi \in \mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : ||x|| = 1\}$.

For $0 \leq j \leq n/2$ and $1 \leq \nu \leq a_{n-2j}^d$, define

$$P_{j,\nu}^n(x) := P_j^{(\mu, n-2j+(d-2)/2)}(2||x||^2 - 1)\, Y_\nu^{n-2j}(x),$$

where $\{Y_\nu^{n-2j} \ : \ 1 \leq \nu \leq a_{n-2j}^d\}$ is an orthonormal basis of $\mathcal{H}_{n-2j}^d$. Then the set $\{P_{j,\ell}^n(x) \ : \ 0 \leq j \leq n/2,\ 1 \leq \ell \leq a_{n-2j}^d\}$ is a mutually orthogonal basis of $\mathcal{V}_n^d(w_\mu)$. The elements of $\mathcal{V}_n^d(w_\mu)$ are eigenfunctions of a second-order linear partial differential operator $\mathcal{D}_\mu$. More precisely, we have

$$\mathcal{D}_\mu P = -(n+d)(n+2\mu)\, P, \qquad \forall P \in \mathcal{V}_n^d(w_\mu), \tag{7.1}$$

where

$$\mathcal{D}_\mu := \Delta - \sum_{j=1}^{d} \frac{\partial}{\partial x_j} x_j \left[2\mu + \sum_{i=1}^{d} x_i \frac{\partial}{\partial x_i} \right]. \tag{7.2}$$

7.1.1 Sobolev Orthogonal Polynomials on the Unit Ball

The first work in this direction is [74] and deals with the inner product

$$\langle f, g \rangle_\Delta := \int_{\mathbb{B}^d} \Delta \left[(1 - ||x||^2)\, f(x) \right] \Delta \left[(1 - ||x||^2)\, g(x) \right] dx,$$

which arises from the numerical solution of the Poisson equation studied in [6]. The geometry of the ball and (7.3) suggest that one can look for a mutually orthogonal basis of the form

$$q_j(2||x||^2 - 1)\, Y_\nu^{n-2j}(x), \qquad Y_\nu^{n-2j} \in \mathcal{H}_{n-2j}^d, \tag{7.3}$$

where $q_j(x)$ is a polynomial of degree j in one variable. Such a basis was constructed in [74] for the space

$$\mathcal{V}_n^d(\Delta) = \mathcal{H}_n^d \bigoplus (1 - ||x||^2) \mathcal{V}_{n-2}(w_2).$$

The next inner product considered on the ball is defined by

$$\langle f, g \rangle_{-1} := \lambda \int_{\mathbb{B}^d} \nabla f(x) \cdot \nabla g(x)\, dx + \int_{\mathbb{S}^{d-1}} f(\xi)\, g(\xi)\, d\sigma(\xi),$$

where $\nabla f = (\partial_x f, \partial_y f)$ and $\lambda > 0$. An alternative way is to replace the integral over $\mathbb{S}^{d-1}$ by $f(0)\, g(0)$. A basis of the form (7.3) was constructed explicitly in [75] for the space $\mathcal{V}_n^d(\Delta)$ with respect to $\langle \cdot, \cdot \rangle_{-1}$, from which it follows that

$$\mathcal{V}_n^d(w_{-1}) = \mathcal{H}_n^d \bigoplus (1 - ||x||^2) \mathcal{V}_{n-2}(w_1). \tag{7.4}$$

The elements in $(1 - ||x||^2)\mathcal{V}_{n-2}(w_1)$ can be given in terms of the Jacobi polynomials $P_n^{(-1,b)}(x)$ of negative index, which explains the notation w_{-1}. Another interesting aspect of this case is that the polynomials in $\mathcal{V}_n^d(w_{-1})$ are eigenfunctions of the differential operator $\mathcal{D}_{-1}$, the limit case of (7.1).

For $k \in \mathbb{N}$, the operator $\mathcal{D}_{-k}$ in (7.2) makes perfect sense. The equation $\mathcal{D}_{-k} Y = \lambda_n Y$ was studied in [66], where a complete system of polynomial solutions was determined explicitly. For $k \geq 2$, however, it is not known if the solutions are

Sobolev orthogonal polynomials. Closely related to the case when $k = 2$ is the following inner product

$$\langle f, g\rangle_{-2} := \lambda \int_{\mathbb{B}^d} \Delta f(x)\, \Delta g(x)\, dx + \int_{\mathbb{S}^{d-1}} f(\xi)\, g(\xi)\, d\sigma(\xi), \qquad \lambda > 0.$$

An explicit basis for the space $\mathcal{V}_d^d(w_{-2})$ of the Sobolev polynomials with respect to $\langle\cdot,\cdot\rangle_{-2}$ was constructed in [66], from which it follows that

$$\mathcal{V}_d^d(w_{-2}) = \mathcal{H}_n^d \bigoplus (1-||x||^2)\mathcal{H}_{n-2}^d \bigoplus (1-||x||^2)^2 \mathcal{V}_{n-4}^d(w_2). \tag{7.5}$$

The elements in $(1 - ||x||^2)^2\mathcal{V}_{n-4}^d(w_2)$ can be given in terms of the Jacobi polynomials $P_n^{(-2,b)}$ of negative index.

It turns out that the Sobolev orthogonal polynomials for the last two cases can be used to study the spectral method for the numerical solutions of partial differential equations. This connection was established in [46], where, for $s \in \mathbb{N}$, the following inner product in the Sobolev space $W_p^s(\mathbb{B}^d)$ is defined

$$\langle f, g\rangle_{-s} := \langle \nabla^s f, \nabla^s g\rangle_{\mathbb{B}^d} + \sum_{k=0}^{\lceil \varrho/2\rceil -1} \lambda_k \langle \Delta^k f, \Delta^k g\rangle_{\mathbb{S}^{d-1}}.$$

Here λ_k, $k = 0, 1, \ldots, \lceil \varrho/2\rceil - 1$, are positive real numbers, and

$$\nabla^{2m} := \Delta^m \qquad \text{and} \qquad \nabla^{2m+1} := \nabla\Delta^m, \qquad m = 1, 2, \ldots$$

For $s > 2$, the space $\mathcal{V}_n^d(w_{-s})$ associated with $\langle\cdot,\cdot\rangle_{-s}$ cannot be decomposed as in (7.4) and (7.5). Nevertheless, an explicit mutually orthogonal basis was constructed in [46]. It requires considerable effort, and the basis uses a generalization of the Jacobi polynomials $P_n^{(\alpha,\beta)}$ for $\alpha, \beta \in \mathbb{R}$ that avoids the degree reduction when $-\alpha - \beta - n \in \{0, 1, \ldots, n\}$. The main result in [46] establishes an estimate for the polynomial approximation in the Sobolev space $W_p^s(\mathbb{B}^d)$. The proof relies on the Fourier expansion with respect to the Sobolev orthogonal polynomials associated with $\langle\cdot,\cdot\rangle_{-s}$.

Another Sobolev inner product considered on the unit ball is defined by

$$\langle f, g\rangle := \int_{\mathbb{B}^d} \nabla f(x)\cdot\nabla g(x)\, W_\mu(x)\, dx + \lambda \int_{\mathbb{B}^d} f(x)\, g(x)\, W_\mu(x)\, dx,$$

which is an extension of the Sobolev inner product (6.2) of coherent pairs where $d\mu_1 = d\mu_2$ correspond to the Gegenbauer weight in one variable. A mutually orthogonal basis was constructed in [65], which has the form (7.3) where the corresponding q_j is orthogonal with respect to a rather involved Sobolev product in one variable.

7.1.2 Sobolev Orthogonal Polynomials on Product Domains

On the product domain $[a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$, we define the weight function

$$\varpi(x, y) = w_1(x)\, w_2(y),$$

where w_i, $i = 1, 2$, is a weight function on $[a_i, b_i]$, $i = 1, 2$. With respect to ϖ, we consider the Sobolev inner product

$$\langle f, g\rangle_S := \int_{[a,b]^2} \nabla f(x, y) \cdot \nabla g(x, y)\varpi(x, y)\, dxdy + \lambda\, f(c_1, c_2)\, g(c_1, c_2),$$

where $\lambda > 0$, and (c_1, c_2) is a fixed point in $\mathbb{R}^2$.

Two cases are considered in [29]. The first one is the product of Laguerre weights for which

$$\langle f, g\rangle_S := \int_0^\infty \int_0^\infty \nabla f(x, y) \cdot \nabla g(x, y) w_\alpha(x)\, w_\beta(y) dxdy + \lambda_k f(0, 0)\, g(0, 0),$$

where $w_\alpha(x) = x^\alpha e^{-x}$, $\alpha > -1$. The Sobolev orthogonal polynomials are related to the polynomials $Q_{j,m}^{\alpha,\beta}(x, y)$ defined by

$$Q_{j,m}^{\alpha,\beta}(x, y) := Q_{m-j}^{\alpha}(x)\, Q_j^{\beta}(y) \quad \text{with} \quad Q_n^{\alpha}(x) := \hat{L}_n^{(\alpha)}(x) + n\, \hat{L}_{n-1}^{(\alpha)}(x),$$

where $\hat{L}_n^{(\alpha)}(x)$ denotes the nth monic Laguerre polynomial. The polynomial $Q_n^\alpha(x)$ is monic and satisfies $\frac{d}{dx} Q_n^\alpha(x) = n\, \hat{L}_{n-1}^{(\alpha)}(x)$. For $0 \le k \le n$, let $S_{n-k,k}^{\alpha,\beta}(x, y) = x^{n-k} y^k + \cdots$ be the monic Sobolev orthogonal polynomial of degree n. Define the column vectors

$$\mathbb{Q}_n^{\alpha,\beta} := (Q_{0,n}^{\alpha,\beta}, \ldots, Q_{n,n}^{\alpha,\beta})^t \quad \text{and} \quad \mathbb{S}_n^{\alpha,\beta} := (S_{0,n}^{\alpha,\beta}, \ldots, S_{n,n}^{\alpha,\beta})^t.$$

It was shown in [29] that there is a matrix B_{n-1} such that

$$\mathbb{Q}_n^{\alpha,\beta} = \mathbb{S}_n^{\alpha,\beta} + B_{n-1}\mathbb{S}_{n-1}^{\alpha,\beta}.$$

Notice that the matrix B_{n-1} and the norm $\langle \mathbb{S}_n^{\alpha,\beta}, \mathbb{S}_n^{\alpha,\beta}\rangle_S$ can both be computed by one recursive algorithm.

The above construction of orthogonal bases for the product domain works if w_1 and w_2 are self-coherent, that is, are classical weights (Jacobi, Laguerre, Hermite). The case when both are Gegenbauer weight functions was given as a second example in [29].

Acknowledgements The authors acknowledge the careful revision of the manuscript by the referees. Their comments and suggestions have contributed to improve its presentation.

The work of J. C. García-Ardila, F. Marcellán and M. Marriaga has been supported by Dirección General de Investigación Científica y Técnica, Ministerio de Economía, Industria y Competitividad of España, research project [MTM2015-65888-C4-2-P].

References

1. M. Alfaro, F. Marcellán, M.I. Rezola, A. Ronveaux, Sobolev type orthogonal polynomials: the nondiagonal case. J. Approx. Theory **83**, 737–757 (1995)
2. M. Alfaro, J.J. Moreno-Balcázar, A. Peña, M.L. Rezola, A new approach to the asymptotics of Sobolev type orthogonal polynomials. J. Approx. Theory **163**, 460–480 (2011)
3. P. Althammer, Eine Erweiterung des orthogonalitätsbergriffes bei polynomen und dere anwendung auf die beste approximation. J. Reine Angew. Math. **211**, 192–2014 (1962)
4. R. Álvarez-Nodarse, J.J. Moreno-Balcázar, Asymptotic properties of generalized Laguerre orthogonal polynomials. Indag. Math. (N.S.) **15**, 151–165 (2004)
5. J. Arvesú, M.J. Atia, F. Marcellán, On semiclassical linear functionals: the symmetric companion. Comm. Anal. Theory Contin. Fractions **10**, 13–29 (2002)
6. K. Atkinson, O. Hansen, Solving the nonlinear Poisson equation on the unit disk. J. Integral Equations Appl. **17**, 223–241 (2005)
7. H. Bavinck, H.G. Meijer, Orthogonal polynomials with respect to a symmetric inner product involving derivatives. Appl. Anal. **33**, 1003-117 (1989)
8. H. Bavinck, H.G. Meijer, On orthogonal polynomials with respect to an inner product involving derivatives: zeros and recurrence relations. Indag. Math. (N.S.) **1**, 7–14 (1990)
9. S. Belmehdi, *Formes Linéaires et Polynômes Orthogonaux Semi-Classiques de Class $s = 1$. Description et Classification* (Thèse d'État, Université Pierre et Marie Curie, Paris, 1990)
10. S. Belmehdi, On semiclassical linear functionals of class $s = 1$. Classification and integral representations. Indag. Math. (N.S.) **3**(3), 253–275 (1992)
11. S. Bochner, Über Sturm-Liouvillesche polynomsysteme. Math. Z. **29**(1), 730–736 (1929)
12. L. Boelen, W. Van Assche, Discrete Painlevé equations for recurrence coefficients of semiclassical Laguerre polynomials. Proc. Amer. Math. Soc. **138**, 1317–1331 (2010)
13. J. Brenner, Über eine erweiterung des orthogonaltäts bei plynomen, in ed. by G. Alex its, S.B. Stechkin, *Constructive Theory of Functions* (Akdémiai Kiadó, Budapest, 1972), pp. 77–83
14. M.I. Bueno, F. Marcellán, Darboux transformation and perturbation of linear functionals. Linear Algebra Appl. **384**, 215–242 (2004)
15. A. Cachafeiro, F. Marcellán, J.J. Moreno-Balcázar, On asymptotic properties of Freud Sobolev orthogonal polynomials. J. Approx. Theory **25**, 26–41 (2003)
16. T.S. Chihara, On co-recursive orthogonal polynomials. Proc. Amer. Math. Soc. **8**, 899–905 (1962)
17. T.S. Chihara, *An Introduction to Orthogonal Polynomials*. Mathematics and its Applications Series, vol. 13 (Gordon and Breach Science Publishers, New York, 1978)
18. E.B. Christoffel, Über die Gaussische Quadratur und eine Verallgemeinerung derselben. J. Reine Angew. Math. **55**, 61–82 (1858)
19. P.A. Clarkson, K. Jordaan, Properties of generalized Freud polynomials. J. Approx. Theory **225**, 148–175 (2018)
20. P.A. Clarkson, K. Jordaan, A. Kelil, A generalized Freud-weight. Studies Appl. Math **136**, 288–320 (2016)
21. E.X.L. de Andrade, C.F. Bracciali, A. Sri Ranga, Zeros of Gegenbauer-Sobolev orthogonal polynomials: beyond coherent pairs. Acta Appl. Math. **105**, 65–82 (2009)
22. A.M. Delgado, F. Marcellán, Companion linear functionals and Sobolev inner procuts: a case study. Methods Appl. Anal. **11**, 237–266 (2004)
23. M. Derevyagin, J.C. Garcia-Ardila, F. Marcellán, Multiple Geronimus transformations. Linear Algebra Appl. **454**, 158–183 (2014)

24. H. Dueñas Ruiz, F. Marcellán, A. Molano, Asymptotics of Sobolev orthogonal polynomials for Hermite (1,1) -coherent pairs. J. Math. Anal. Appl. **467**(1), 601–621 (2018)
25. C.F. Dunkl, Y. Xu, *Orthogonal Polynomials in Several Variables*. Encyclopedia of Mathematics and its Applications, vol. 155, 2nd edn. (Cambridge University Press, Cambridge, 2014)
26. A.J. Durán, A generalization of Favard's theorem for polynomials satisfying a recurrence ration. J. Approx. Theory **74**, 83–109 (1993)
27. A.J. Durán, W. Van Assche, Orthogonal matrix polynomials and higher order recurrence relations. Linear Algebra Appl. **219**, 261–280 (1995)
28. W.D. Evans, L.L. Littlejohn, F. Marcellán, C. Markett, A. Ronveaux, On recrrence relations for Sobolev orthogonal polynomials. SIAM J. Math. Anal. **26**, 446–467 (1995)
29. L. Fernández, F. Marcellán, T.E. Pérez, M.A. Piñar, Y. Xu, Sobolev orthogonal polynomials on product domain. J. Comput. Appl. Math. **284**, 202–215 (2015)
30. J.S. Geronimo, D.S. Lubinsky, F. Marcellán, Asymptotic for Sobolev orthogonal polynomials for exponential weights. Constr. Approx. **22**, 309–346 (2005)
31. Ya. L. Geronimus, On polynomials orthogonal with regard to a given sequence of numbers. Comm. Inst. Sci. Math. Mec. Univ. Kharkoff [Zapiski Inst. Mat. Mech.] **17**(4), 3–18 (1940) (Russian)
32. Ya. L. Geronimus, On polynomials orthogonal with respect to numerical sequences and on Hahn's theorem. Izv. Akad. Nauk. **4**, 215–228 (1940) (in Russian)
33. W. Gröbner, in Orthogonale polynommsysteme, die gleichzeitig mit $f(x)$ auch deren ableitung $f'(x)$ approximieren, ed. by L. Collatz, *Funktionalanalysis, Approximationstheorie, Numerische Mathematik*, ISNM, vol. 7 (Birkhäuser, Basel, 1967), pp. 24–32
34. W. Hahn, Über Differentialgleichunqen für Orthoqonalpolynome. Monat. Math. **95**, 269–274 (1983)
35. R.A. Horn, C.R. Johnson, *Matrix Analysis*, 2nd edn. (Cambridge University Press, Cambridge, 2013)
36. A. Iserles, J.M. Sanz-Serna, P.E. Koch, S.P. Nørsett, *Orthogonality and Approximation in a Sobolev Space*, in Algorithms for Approximation, II (Shrivenham, 1988) (Chapman and Hall, London, 1990), pp. 117–124
37. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
38. D.H. Kim, K.H. Kwon, F. Marcellán, G.J. Yoon, Sobolev orthogonality and coherent pairs of moment functionals: an inverse problem. Internat. Math. J. **2**, 877–888 (2002)
39. R. Koekoek, Koornwinder's Laguerre polynomials. Delft Prog. Rep. **12**, 393–404 (1988)
40. R. Koekoek, Generalizations of Laguerre polynomials. J. Math. Anal. Appl. **153**, 576–590 (1990)
41. R. Koekoek, H.G. Meijer, A generalization of Laguerre polynomials. SIAM J. Math. Anal. **24**, 768–782 (1993)
42. R. Koekoek, P. Lesky, R. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010)
43. H.L. Krall, On orthogonal polynomials satisfying a certain fourth order differential equation. Pa. State Coll. Stud. **6**, 1–24 (1940)
44. A.M. Krall, Orthogonal polynomials satisfying fourth order differential equations. Proc. Roy. Soc. Edinburgh, Sect. A **87**, 271–288 (1980/1981)
45. D.C. Lewis, Polynomial least square approximations. Amer. J. Math. **69**, 273–278 (1947)
46. H. Li, Y. Xu, Spectral approximation on the unit ball. SIAM J. Numer. Anal. **52**(6), 2647–2675 (2014)
47. G. López Lagomasino, H. Pijeira, Nth root asymptotics of Sobolev orthogonal polynomials. J. Approx. Theory **99**, 30–43 (1999)
48. G. López Lagomasino, F. Marcellán, W. Van Assche, Relative asymptotics for polynomials orthogonal with respect to a discrete Sobolev inner product. Constr. Approx. **11**, 107–137 (1995)

49. F. Marcellán, P. Maroni, Sur l'adjonction d'une masse de Dirac à une forme régulière et semi-classique. Ann. Mat. Pura Appl. **162**(4), 1–22 (1992)
50. F. Marcellán, J.J. Moreno-Balcázar, Asymptotics and zeros of Sobolev orthogonal polynomials on unbounded supports. Acta Appl. Math. **94**, 163–192 (2006)
51. F. Marcellán, T.E. Pérez, M.A. Piñar, Laguerre Sobolev orthogonal polynomials. J. Comput. Appl. Math. **71**, 245–265 (1996)
52. F. Marcellán, J. Petronilho, Orthogonal polynomials and coherent pairs: the classical case. Indag. Math. (N.S.) **6**, 287–307 (1995)
53. F. Marcellán, A. Ronveaux, On a class of polynomials orthogonal with respect to a discrete Sobolev inner product. Indag. Math. (N.S.) **1**, 451–464 (1990)
54. F. Marcellán, W. Van Assche, Relative asymptotics for orthogonal polynomials with a Sobolev inner product. J. Approx. Theory **72**, 193–209 (1993)
55. F. Marcellán, R. Xh. Zejnullahu, B. Xh. Fejzullahu, E. Huertas, On orthogonal polynomials with respect to certain discrete Sobolev inner product. Pacific J. Math. **257**, 167–188 (2012)
56. F. Marcellán, Y. Xu, On Sobolev orthogonal polynomials. Expo. Math. **33**(3), 308–352 (2015)
57. P. Maroni, Une caractérisation des polynômes orthogonaux semi-classiques, C. R. Acad. Sci. Paris Sér. I Math. **301**(6), 269–272 (1985)
58. P. Maroni, Le calcul des formes linéaires et les polynômes orthogonaux semi-classiques, in ed. by M. Alfaro, *Orthogonal Polynomials and Their Applications*, LNM 1329 (Springer, Berlin, 1988), pp. 279–290
59. P. Maroni, Une thèorie algébrique des polynômes orthogonaux. Application aux polynômes semiclassiques, in *Orthogonal Polynomials and Their Applications (Erice 1990)*. IMACS Annals of Computational and Applied Mathematics, vol. 9 (Baltzer, Basel, 1991), pp. 95–130
60. A. Martínez-Finkelshtein, Bernstein-Szegő's theorem for Sobolev orthogonal polynomials. Constr. Approx. **16**, 73–84 (2000)
61. A. Martínez-Finkelshtein, J.J. Moreno-Balcázar, T.E. Pérez, M.A. Piñar, Asymptotics of Sobolev orthogonal polynomials for coherent pairs. J. Approx. Theory **92**, 280–2893 (1998)
62. H.G. Meijer, A short history of orthogonal polynomials in a Sobolev space I. The non-discrete case, in: 31st Dutch mathematical conference, Groningen, 1995. Nieuw Arch. Wiskd. **14**(4), 93–112 (1996)
63. H.G. Meijer, Determination of all coherent pairs of functionals. J. Approx. Theory **89**, 321–343 (1997)
64. A.F. Nikiforov, V.B. Uvarov, *Special Functions of Mathematical Physics* (Birkhäuser Verlag, Basel 1988)
65. T.E. Pérez, M.A. Piñar, Y. Xu, Weighted Sobolev orthogonal polynomials on the unit ball. J. Approx. Theory **171**, 84–104 (2013)
66. M.A. Piñar, Y. Xu, Orthogonal polynomials and partial differential equations on the unit ball. Proc. Amer. Math. Soc. **37**, 2979–2987 (2009)
67. F.W. Schäfke, Zu den orthogonalplynomen von Althammer. J. Reine Angew. Math. **252**, 195–199 (1972)
68. F.W. Schäfke, G. Wolf, Einfache verallgemeinerte klassische orthogonal polynome. J. Reine Angew. Math. **262/263**, 339–355 (1973)
69. J. Shohat, A differential equation for orthogonal polynomials. Duke Math. J. **5**, 401–417 (1939)
70. G. Szegő, *Orthogonal Polynomials*, 4th edn., vol. 23. American Mathematical Society, Colloquium Publications (American Mathematical Society, Providence, 1975)
71. V.B. Uvarov, Relation between polynomials orthogonal with different weights. Doklady Akademii Nauk SSSR **126**(1), 33–36 (1959) (in Russian)
72. V.B. Uvarov, The connection between systems of polynomials that are orthogonal with respect to different distribution functions. USSR Comput. Math. Phys. **9**, 25–36 (1969)
73. W. Van Assche, *Orthogonal Polynomials and Painlevé Equations*. Australian Mathematical Society Lecture Series, vol. 27 (Cambridge University Press, Cambridge, 2018)
74. Y. Xu, A family of Sobolev orthogonal polynomials on the unit ball. J. Approx. Theory **138**, 232–241 (2006)

75. Y. Xu, Sobolev orthogonal polynomials defined via gradient on the unit ball. J. Approx. Theory **152**, 52–65 (2008)
76. G. Yoon, Darboux transforms and orthogonal polynomials. Bull. Korean Math. Soc. **39**, 359–376 (2002)
77. A. Zhedanov, Rational spectral transformations and orthogonal polynomials. J. Comput. Appl. Math. **85**, 67–86 (1997)

Two Variable Orthogonal Polynomials and Fejér-Riesz Factorization

J. S. Geronimo

Abstract We consider bivariate polynomials orthogonal on the bicircle with respect to a positive nondegenerate measure. The theory of scalar and matrix orthogonal polynomials is reviewed with an eye toward applying it to the bivariate case. The lexicographical and reverse lexicographical orderings are used to order the monomials for the Gram–Schmidt procedues and recurrence formulas are derived between the polynomials of different degrees. These formulas link the orthogonal polynomials constructed using the lexicographical ordering with those constructed using the reverse lexicographical ordering. Relations between the coefficients in the recurrence formulas are derived and used to give necessary and sufficient conditions for the existence of a positive measure. These results are then used to construct a class of two variable measures supported on the bicircle that are given by one over the magnitude squared of a stable polynomial. Applications to Fejér–Riesz factorization are also given.

Keywords Orthogonal polynomials on the bicircle · Fejér-Riesz factorization · Recurrence relations · Matrix orthogonal polynomials on the unit circle · Bivariate orthogonal polynomials · Matrix orthogonal polynomials moment problem · Recurrence coefficients

AMS Classification Numbers 42C05, 42C10

Based on lectures given for the AIMS-Volkswagen Foundation Workshop on Introduction to Orthogonal Polynomials and Applications.

J. S. Geronimo (✉)
Georgia Institute of Technology, Atlanta, GA, USA
e-mail: jeffrey.geronimo@math.gatech.edu

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_14

1 Introduction

I had the good fortune to have been invited to the AIMS-Volkswagen Foundation Workshop on "An Introduction to orthogonal polynomials and applications" in Douala, Cameroon to give lectures based on recent results in the theory of orthogonal polynomials on the bicircle. The notes below are based on these lectures. The theory as developed here has its roots in the work of Delsarte et al. [5] and the extension of the Fejér-Riesz theorem to the bivariate case by Geronimo and Woerdeman [8]. Development of the theory of bivariate orthogonal polynomials was presented in Geronimo and Woerdeman [9]. Further developments and extensions were made in the works by Bakonyi and Woerdeman [1], Geronimo and Iliev [7], Geronimo, Iliev, and Knese [10, 11], Knese [15, 16], and Landau and Landau [17].

The notes are organized as follows: In Sect. 2 the one variable Fejér-Riesz theorem is presented and then the theory of orthogonal polynomials on the unit circle is developed in order to give a proof of the Fejér-Riesz theorem that does not use the Fundamental Theorem of Algebra. Also proved in this section is Verblunsky's theorem. Little is new in this section and the results can be found in the books by Geronimus [12], Szegő [19], and Simon [18]. The main reason for this section is to review for the student the theory of orthogonal polynomials on the unit circle and develop a proof of the Fejér-Riesz theorem whose main elements will be extended to the bivariate case. The use of the Cauchy–Schwarz inequality in the proof of the Fejér-Riesz theorem seems to be due to Knese [16]. In Sect. 3 the theory of matrix orthogonal polynomials on the unit circle will be developed as initiated by Delsarte et al. [4]. As above little is new in this section. The normalization used here is different from the one imposed in [4] and Damanik et al. [3] and is chosen in order to make contact with the two variable theory. This complicates slightly the recurrence formulas in that the "recurrence coefficients" in the ascending recurrence formulas are different from those in the descending recurrence formulas. However these coefficients have their counterparts in the two variable theory. Here we also prove a matrix Fejér-Riesz theorem and a matrix Verblunsky theorem. Finally in Sect. 4 we consider the extension of the one variable theory of orthogonal polynomials to the bivariate case. An immediate problem is which ordering to use for the bivariate monomials and following [5] we use the lexicographical and reverse lexicographical orderings. In these orderings the moment matrix has a doubly Toeplitz structure which allows us to make contact with the theory of matrix orthogonal polynomials and develop more recurrences. The results in this section can be found in [7–10], and [11]. Using these recurrences a set of parameters is discussed and then an extension of the Fejér-Riesz theorem is presented.

2 Scalar Orthogonal Polynomials on the Unit Circle

In 1915–1918 the following theorem now known as the Fejér-Riesz theorem was proved,

Theorem 2.1 *Suppose q_n is a trigonometric polynomial of degree n. Then $q_n(\theta) > 0$ if and only if*

$$q_n(\theta) = |p_n(z)|^2, \qquad z = e^{i\theta},$$

where p_n is a polynomial of degree n in z which can be chosen nonzero for $|z| \leq 1$.

The most general statement of the Fejér-Riesz theorem only requires that q_n be nonnegative and follows from the one above by continuity. We will be content with the theorem as stated since it has a two variable analog which will be discussed below.

This theorem has many applications especially in signal processing, the trigonometric moment problem, and in the construction of wavelets. The proofs of the theorem typically use strongly the fundamental theorem of algebra so straightforward extensions to more than one variable are not obvious. Another proof based on the theory of polynomials orthogonal on the unit circle can be extended at least to the two variable case so it is this proof that will be developed.

Let μ be a positive Borel measure on $[-\pi, \pi]$ and let its Fourier coefficients c_k be given by

$$c_k = \int_{-\pi}^{\pi} e^{-ik\theta} d\mu(\theta).$$

The complex conjugate of the above coefficient is given by $\bar{c}_n = c_{-n}$ since μ is real. The Toeplitz matrix C_n constructed from these Fourier coefficients is the $(n+1) \times (n+1)$ matrix

$$C_n = \begin{bmatrix} c_0 & c_{-1} & \cdots & c_{-n} \\ c_1 & c_0 & \ddots & c_{-n+1} \\ \vdots & \ddots & \ddots & \vdots \\ c_n & c_{n-1} & \cdots & c_0 \end{bmatrix} \tag{2.1}$$

and it will be assumed that $C_n > 0$ (i.e. positive definite) for all $n \geq 0$ which is equivalent to

$$\int_{-\pi}^{\pi} |p(z)|^2 d\mu > 0, \qquad z = e^{i\theta}, \tag{2.2}$$

for every nonzero polynomial with complex coefficients. This can be seen by writing any polynomial of degree n, $p(z) = p_n z^n + \dots p_0$ as $p(z) = [p_n, \dots, p_0][z^n, \dots, 1]^T$ then substitute this into (2.2) to obtain

$$\int_{-\pi}^{\pi} |p(z)|^2 d\mu = [p_n, \dots, p_0] C_n [p_n, \dots, p_0]^\dagger.$$

Here † means complex conjugate transpose. The inner product

$$\langle p, q\rangle_\mu = \int_{-\pi}^{\pi} p(z)\overline{q(z)} d\mu = \overline{\langle q, p\rangle_\mu}, \qquad z = e^{i\theta},$$

can be used to construct a unique sequence of orthonormal polynomials $\phi_n,\ n = 0,\ 1,\ 2\dots$ satisfying

- $\phi_n(z) = \phi_{n,n} z^n + \dots + \phi_{n,0}$ with $\phi_{n,n} > 0$,
- $\langle \phi_n, \phi_k\rangle_\mu = \delta_{n,k}$.

Here $\delta_{n,k}$ is the Kronecker delta. The unique solution which can easily be checked to satisfy the above conditions is given by

$$\phi_n(z) = \frac{1}{\sqrt{\det C_n \det C_{n-1}}} \begin{vmatrix} c_0 & c_{-1} & \cdots & c_{-n} \\ c_1 & c_0 & \ddots & c_{-n+1} \\ \vdots & \ddots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_{-1} \\ 1 & z & \cdots & z^n \end{vmatrix}.$$

Important in the study of orthogonal polynomials on the unit circle is the so called reverse polynomial $\phi_n(z)^* = z^n \bar{\phi}_n(1/z)$. From the fact that

$$\langle \phi_n, z^i\rangle_\mu = 0, \qquad i = 0, \dots, n-1, \tag{2.3}$$

we find

$$\langle \phi_n^*, z^i\rangle_\mu = 0, \qquad i = 1, \dots, n. \tag{2.4}$$

The above equations characterize ϕ_n, respectively ϕ_n^*, up to multiplication by a constant.

The above conditions allow the orthogonal polynomials and their reverses to satisfy the following recurrence formulas for $k \ge 1$:

$$\phi_k(z) = a_k(z\phi_{k-1} - \alpha_k \phi_{k-1}^*(z)), \tag{2.5}$$

and

$$\phi_k^*(z) = a_k(\phi_{k-1}^* - \bar{\alpha}_k z\phi_{k-1}(z)). \tag{2.6}$$

The second equation is just the reverse of the first. To obtain the first note that if $a_k = \frac{\phi_{k,k}}{\phi_{k-1,k-1}}$ then $\phi_k(z) - a_k z\phi_{k-1}(z)$ is a polynomial of degree $k-1$ satisfying the orthogonality relations

$$\langle \phi_k - a_k z\phi_{k-1}, z^i\rangle_\mu = 0\,, \qquad i = 1, \dots, k-1.$$

Thus Eq. (2.4) shows $\phi_k - a_k z\phi_{k-1}$ is a constant times ϕ_k^* and we choose the constant to be equal to $-a_k\alpha_k$ to obtain Eq. (2.5). Since $\langle \phi_k(z), \phi_{k-1}^*(z)\rangle_\mu = 0$ Eq. (2.5) shows that

$$\alpha_k = \langle z\phi_{k-1}, \phi_{k-1}^*\rangle_\mu. \tag{2.7}$$

The orthonormality of ϕ_k gives

$$\begin{aligned} 1 = \langle \phi_k, \phi_k\rangle_\mu &= a_k^2\langle z\phi_{k-1} - \alpha_k\phi_{k-1}^*, z\phi_{k-1} - \alpha_k\phi_{k-1}^*\rangle_\mu \\ &= a_k^2(1 - \bar{\alpha}_k\langle z\phi_{k-1}, \phi_{k-1}^*\rangle_\mu - \alpha_k\overline{\langle z\phi_{k-1}, \phi_{k-1}^*\rangle_\mu} + |\alpha_k|^2) \\ &= a_k^2(1 - |\alpha_k|^2). \end{aligned} \tag{2.8}$$

This shows that the α_k which we call recurrence coefficients must have magnitude less than one. In the literature the parameters α_k go by other names such as reflection coefficients or Verblunsky coefficients. If ϕ_{k-1}^* is eliminated in Eq. (2.5) using (2.6) and (2.8) the result is,

$$a_k(\phi_k(z) + \alpha_k\phi_k^*(z)) = z\phi_{k-1}, \tag{2.9}$$

and its reverse

$$a_k(\phi_k^*(z) + \bar{\alpha}_k\phi_k(z)) = \phi_{k-1}^*, \tag{2.10}$$

The polynomial ϕ_n^* has some remarkable properties:

(i) It is stable, i.e., $\phi_n^*(z) \neq 0,\ |z| \leq 1$.
(ii) It has spectral matching i.e., $c_j^n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ij\theta}\frac{d\theta}{|\phi_n^*(z)|^2} = c_j$ for $|j| \leq n$.

The stability can be proved through the Christoffel–Darboux formula which is,

$$\frac{\phi_n^*(z)\overline{\phi_n^*(z_1)} - z\bar{z}_1\phi_n(z)\overline{\phi_n(z_1)}}{1 - z\bar{z}_1} = \sum_{k=0}^{n}\phi_k(z)\overline{\phi_k(z_1)}. \tag{2.11}$$

This follows in a straightforward manner by multiplying (2.6) by it complex conjugate and setting $z = z_1$ then subtracting (2.5) multiplied by its complex conjugate at $z = z_1$. The use of (2.8) gives

$$\phi_n^*(z)\overline{\phi_n^*(z_1)} - \phi_n(z)\overline{\phi_n(z_1)} = \phi_{n-1}^*(z)\overline{\phi_{n-1}^*(z_1)} - \phi_{n-1}(z)\overline{\phi_{n-1}(z_1)} + (1 - z\bar{z}_1)\phi_{n-1}(z)\overline{\phi_{n-1}(z_1)}.$$

Iteration of this equation yields

$$\phi_n^*(z)\overline{\phi_n^*(z_1)} - \phi_n(z)\overline{\phi_n(z_1)} = \phi_0(z)^*\overline{\phi_0^*(z_1)} - \phi_0(z)\overline{\phi_0(z_1)} + (1 - z\bar{z}_1)\sum_{k=0}^{n-1}\phi_k(z)\overline{\phi_k(z_1)}.$$

adding $(1 - z\bar{z}_1)\phi_n(z)\overline{\phi_n(z_1)}$ to both sides of the above equation gives the result since $\phi_0(z) = \frac{1}{\sqrt{c_0}} = \phi_0^*(z)$ which is a positive constant.

We now prove property (i) above,

Proof Set $z_1 = z$ with $|z| < 1$ then Eq. (2.11) implies

$$|\phi_n^*(z)|^2 \geq (1 - |z|^2)\sum_{k=0}^{n}|\phi_k(z)|^2 > (1 - |z|^2)|\phi_0(z)|^2 > 0,$$

because $\phi_0 = 1/\sqrt{c_0}$. If $\phi_n^*(z_0) = 0$ with $|z_0| = 1$ then $\phi_n(z_0) = 0$ so from (2.11), $0 = \sum_{k=0}^{n-1}\phi_k(z_0)\bar{\phi}_k(z_1)$ where z_1 is free. But this violates the linear independence of $\phi_k,\ k = 0, \ldots, n-1$. Thus $\phi_n^*(z)$ is stable.

We now prove (ii) spectral matching. This will be accomplished by showing that $\phi_k,\ k = 0, \ldots n$ and orthonormal with respect to the weight $d\mu_n = \frac{1}{2\pi}\frac{d\theta}{|\phi_n^*(e^{i\theta})|^2}$. Observe that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\frac{\phi_n(z)|^2}{|\phi_n^*(z)|^2}d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}d\theta = 1,$$

and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\phi_n(z)z^{-i}}{|\phi_n^*(z)|^2}d\theta = \frac{1}{2\pi i}\int_{|z|=1}\frac{z^{n-i-1}}{\phi_n^*(z)}dz,$$

where the above contour integration is taken in the counterclockwise direction. Since ϕ_n^* is nonzero for $|z| \leq 1$, $\frac{1}{\phi_n^*(z)}$ is analytic inside and on the unit circle Cauchy's theorem shows that the above integral is equal to zero for $i < n$. Thus

ϕ_n is an orthonormal polynomial associated with μ_n. Set $z = 0$ in (2.9) with $k = n$ then

$$\alpha_n = \frac{\phi_n(0)}{\phi_n^*(0)}. \tag{2.12}$$

Thus (2.8), (2.9), and (2.12) show we can construct ϕ_{n-1} just from ϕ_n. The use of (2.9) gives

$$\int_{-\pi}^{\pi} \phi_{n-1}(z)z^{-i}d\mu_n = a_n\left(\int_{-\pi}^{\pi} \phi_n(z)z^{-i-1}d\mu_n + \alpha_n \int_{-\pi}^{\pi} \phi_n^*(z)z^{-i-1}d\mu_n\right).$$

From the orthogonality properties of ϕ_n and ϕ_n^* (see (2.3) and (2.4) using μ_n) the above integral is equation to zero for $0 \le i < n-1$. Again using (2.9) in the equation below and expanding yields

$$\begin{aligned}\int_{-\pi}^{\pi} |\phi_{n-1}(z)|^2 d\mu_n = a_n^2 \int_{-\pi}^{\pi} |\phi_n^2(z)| d\mu_n + \bar{\alpha}_n a_n \int_{-\pi}^{\pi} \phi_n(z)\overline{\phi_n^*(z)} d\mu_n \\ + \alpha_n a_n \int_{-\pi}^{\pi} \bar{\phi}_n(z)\phi_n^*(z) d\mu_n + |\alpha_n|^2 a_n^2 \int_{-\pi}^{\pi} |\phi_n^*(z)|^2 d\mu_n.\end{aligned}$$

From (2.9) we find $\alpha_n = -\langle \phi_n, \phi_n^* \rangle$ which with (2.8) shows the above integral is equal to 1. We now proceed to ϕ_{n-2} etc. which shows that $\{\phi_k\}_{k=0}^n$ is a set of orthonormal polynomials associated with μ and μ_n. Since $e^{ij\theta} = z^j$ can be written in terms of ϕ_k with $0 \le k \le j$ and $e^{-ij\theta}$ can be obtained by complex conjugation of the above equation we see that the integrals in (ii) must be the same for $|j| \le n$ which gives the result. □

With the above results a Proof of Theorem 2.1 can be given,

Proof of Theorem 2.1 Suppose that q_n is a strictly positive trigonometric polynomial and let $d\mu = \frac{1}{2\pi}\frac{d\theta}{q_n}$. Construct the orthonormal polynomials associated with μ and let ϕ_n be the nth orthonormal polynomial. Now use the Cauchy–Schwarz inequality to obtain,

$$\begin{aligned}1 =& \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\phi_n^*(z)|}{\sqrt{q_n}} \frac{\sqrt{q_n}}{|\phi_n^*(z)|} d\theta \\ \le& \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\phi_n^*(z)|^2}{q_n} d\theta\right)^{1/2} \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{q_n}{|\phi_n^*(z)|^2} d\theta\right)^{1/2}.\end{aligned} \tag{2.13}$$

The first integral in the inequality is equal to 1 since $|\phi_n^*| = |\phi_n|$ and ϕ_n is an orthonormal polynomial. Also since $\frac{1}{|\phi_n^*|^2}$ and $\frac{1}{q_n}$ have the same first n moments we can replace $|\phi_n^*|^2$ in the second integral with q_n and so it too is equal to one. Thus we have equality which implies that $c\frac{|\phi_n^*(z)|}{\sqrt{q_n}} = \frac{\sqrt{q_n}}{|\phi_n^*(z)|}$ or $c|\phi_n^*|^2 = q_n$. But since

ϕ_n is an orthonormal polynomial with respect to $\frac{1}{2\pi}\frac{1}{q_n}$, $c = 1$ which completes the proof when q_n is strictly positive. □

We finish this section with Verblunsky's theorem which shows that there is a one-to-one correspondence between probability measures supported on the unit circle and sequences of recurrence coefficients and thus gives a parameterization of positive Borel measures on $[-\pi, \pi]$ satisfying (2.2). We first introduce the concept of weak convergence. A sequence of probability measures $\{\mu_n\}$ supported on the interval $[-\pi, \pi]$ is said to converge weakly to μ if

$$\int_{-\pi}^{\pi} f d\mu_n \to \int_{-\pi}^{\pi} f d\mu,$$

for every complex continuous function on $[-\pi, \pi]$.

If two measures μ and μ_1 with support on $[-\pi, \pi]$ are such that $\int_{-\pi}^{\pi} f d\mu = \int_{-\pi}^{\pi} f d\mu_1$ for every complex continuous function on $[-\pi, \pi]$ then $\mu_1 = \mu$.

With this we prove Verblunsky's theorem.

Theorem 2.2 *There is a one-to-one correspondence between infinite sequences $\{\alpha_i\}_{i=1}^{\infty}$ with $|\alpha_i| < 1$ and Borel probability measures on the unit circle satisfying (2.2).*

Proof Let μ be a positive Borel probability measure on $[-\pi, \pi]$ such that (2.2) holds and let $\{\phi_n\}$ be its associated orthonormal polynomials. If $d\mu_n = \frac{1}{2\pi}\frac{d\theta}{|\phi_n^*(z)|^2}$ then the recurrence coefficients $\{\alpha_k\}_{k=1}^{\infty}$ are constructed using (2.7) and the orthonormality of ϕ_k shows that the recurrence coefficients must satisfy (2.8) so $|\alpha_k| < 1$. Now given the recurrence coefficients $\{\alpha_k\}_{k=1}^{\infty}$ with $|\alpha_k| < 1$ we construct a set of polynomials $\{\phi_n\}_{n=0}^{\infty}$ using (2.5) and (2.6) with $\phi_0 = 1 = \phi_0^*$. Since these polynomials satisfy (2.11) ϕ_n^* is stable for each n. Define $d\mu_n = \frac{1}{2\pi}\frac{d\theta}{|\phi_n^*|^2}$ then the first $|k|$ trigonometric moments of μ_k are the same as μ_n for $n \geq k$. Thus we define μ on the set of trigonometric polynomials on $[-\pi, \pi]$ with complex coefficients by

$$\int_{-\pi}^{\pi} p(e^{i\theta}) d\mu = \lim_{n\to\infty} \int_{-\pi}^{\pi} p(e^{i\theta}) d\mu_n.$$

The above limit in fact does not change once $n \geq \deg p$ so that μ is given on this set of polynomials. But by the Weierstrass theorem these polynomials are dense in the set of complex continuous functions on $[-\pi, \pi]$. Thus integration with respect to μ can be extended to any complex continuous function f on $[-\pi, \pi]$ by $\lim_{n\to\infty} \int_{-\pi}^{\pi} f d\mu_n = \int_{-\pi}^{\pi} f d\mu$ which determines μ. Since each μ_n is a positive measure the integral of μ over any positive continuous function on $[-\pi, \pi]$ is positive and this gives the result. □

3 Matrix Orthogonal Polynomials on the Unit Circle

We now extend the theory discussed earlier to matrix orthogonal polynomials which will be important for our discussion of polynomials orthogonal on the bicircle and the extension of the Fejér-Riesz theorem to two variables.

Let Ω be an $(m+1)\times(m+1)$ positive hermitian matrix valued measure on $[-\pi,\pi]$. That is the entries in Ω are complex measures such that $X^\dagger \int_o d\Omega(\theta) X \geq 0$ for any $X \in \mathbb{C}^{m+1}$ and o any open set in $[-\pi,\pi]$. Let

$$C_j = \int_{-\pi}^{\pi} e^{-ij\theta} d\Omega(\theta) \tag{3.1}$$

be the jth $(m+1)\times(m+1)$ matrix Fourier coefficient associated with Ω and construct the block Toeplitz matrix

$$C_{n,m} = \begin{bmatrix} C_0 & C_{-1} & \cdots & C_{-n} \\ C_1 & C_0 & \ddots & C_{-n+1} \\ \vdots & \ddots & \ddots & \vdots \\ C_n & C_{n-1} & \cdots & C_0 \end{bmatrix}, \tag{3.2}$$

so $C_{n,m}$ is an $(n+1)(m+1)\times(n+1)(m+1)$ matrix. Note that $C_{-j} = C_j^\dagger$. We will assume that $C_{n,m} > 0$ for all n. This is equivalent to

$$\int_{-\pi}^{\pi} X(z) d\Omega X(z)^\dagger > 0, \qquad z = e^{i\theta}, \tag{3.3}$$

for $X(z)$ any nonzero polynomial in z with coefficients in $\mathbb{C}^{m+1}$ that is not identically zero. To see this suppose X is of at most degree n so $X(z) = X_0 + \cdots + X_n z^n$ with $X_i \in \mathbb{C}^{m+1}$. Write $X(z) = \hat{X} U(z)$ where $U(z) = [z^n I, z^{n-1} I, \cdots, I]^T$ and $\hat{X} = [X_n, X_{n-1}, \ldots, X_0]$, then

$$\int_{-\pi}^{\pi} X(z) d\Omega X^\dagger(z) = \hat{X} C_{n,m} \hat{X}^\dagger$$

which shows the equivalence of (3.2) being positive and (3.3). We now construct the orthonormal polynomials associated with Ω. Because of noncommutativity it is necessary to introduce two sets of polynomials. Let $\{R_i^m\}_{i=0}^\infty$ and $\{L_i^m\}_{i=0}^\infty$ be sequences of $(m+1)\times(m+1)$ complex valued matrix polynomials

$$R_i^m(z) = R_{i,i}^m z^i + R_{i,i-1}^m z^{i-1} + \cdots, \qquad i = 0, \ldots, n, \tag{3.4}$$

and

$$L_i^m(z) = L_{i,i}^m z^i + L_{i,i-1}^m z^{i-1} + \cdots, \qquad i = 0, \ldots, n, \tag{3.5}$$

satisfying

$$\int_{-\pi}^{\pi} R_i^m(z)^\dagger d\Omega(\theta) R_j^m(z) = \delta_{ij} I_{m+1}, \qquad z = e^{i\theta} \tag{3.6}$$

and

$$\int_{-\pi}^{\pi} L_i^m(z) d\Omega(\theta) L_j^m(z)^\dagger = \delta_{ij} I_{m+1}, \qquad z = e^{i\theta} \tag{3.7}$$

respectively, where I_{m+1} denotes the $(m+1) \times (m+1)$ identity matrix. The $\{R_n^m\}_{n\geq 0}$ and $\{L_n^m\}_{n\geq 0}$ are respectively right and left matrix orthogonal polynomials associated with Ω. The above relations uniquely determine the sequences $\{R_i^m\}_{i=0}^n$ and $\{L_i^m\}_{i=0}^n$ up to a unitary factor and this factor will be fixed by requiring $R_{i,i}^m$ and $L_{i,i}^m$ to be upper triangular matrices with positive diagonal entries. This is not the usual way to normalize the polynomials as it complicates slightly the theory below but it has the advantage of being related to the Cholesky factorization of $C_{n,m}$ which is discussed next and also being connected with the bivariate problem discussed later. For $n \geq i$ write

$$L_i^m(z) = [0 \cdots 0 \; L_{i,i}^m \; L_{i,i-1}^m \; \cdots \; L_{i,0}^m] \begin{bmatrix} z^n I_{m+1} \\ z^{n-1} I_{m+1} \\ \vdots \\ I_{m+1} \end{bmatrix}, \tag{3.8}$$

and

$$\hat{L}_n^m(z) = \begin{bmatrix} L_n^m(z) \\ L_{n-1}^m(z) \\ \vdots \\ L_0^m(z) \end{bmatrix} = L \begin{bmatrix} z^n I_{m+1} \\ z^{n-1} I_{m+1} \\ \vdots \\ I_{m+1} \end{bmatrix}, \tag{3.9}$$

where

$$L = \begin{bmatrix} L_{n,n}^m & L_{n,n-1}^m & \cdots & L_{n,0}^m \\ 0 & L_{n-1,n-1}^m & \cdots & L_{n-1,0}^m \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & L_{0,0}^m \end{bmatrix}. \tag{3.10}$$

In an analogous fashion write,

$$\hat{R}_n^m(z) = \begin{bmatrix} R_0^m(z) \\ R_1^m(z) \\ \vdots \\ R_n^m(z) \end{bmatrix} = \begin{bmatrix} I_{m+1} & \dots & z^n I_{m+1} \end{bmatrix} R, \tag{3.11}$$

where

$$R = \begin{bmatrix} R_{0,0}^m & R_{1,0}^m & \cdots & R_{n,0}^m \\ 0 & R_{1,1}^m & \cdots & R_{n,1}^m \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & R_{n,n}^m \end{bmatrix}. \tag{3.12}$$

By lower (respectively upper) Cholesky factor A (respectively B) of a positive definite matrix M we mean

$$M = AA^\dagger = BB^\dagger, \tag{3.13}$$

where A is a lower triangular matrix with positive diagonal elements and B is an upper triangular matrix with positive diagonal elements. With the above we have the following well known lemma

Lemma 3.1 *Let $C_{n,m}$ be a positive definite block Toeplitz matrix given by* (3.2) *then $L^\dagger$ is the lower Cholesky factor and R is the upper Cholesky factor of $C_{n,m}^{-1}$.*

Proof To obtain (3.13) note that (3.7) gives

$$I = \int_{-\pi}^{\pi} \hat{L}_n^m(z) d\Omega(\theta) \hat{L}_n^m(z)^\dagger = L \int_{-\pi}^{\pi} \begin{bmatrix} z^n I_{m+1} \\ z^{n-1} I_{m+1} \\ \vdots \\ I_{m+1} \end{bmatrix} d\Omega L^\dagger \begin{bmatrix} z^n I_{m+1} \\ z^{n-1} I_{m+1} \\ \vdots \\ I_{m+1} \end{bmatrix}^\dagger L^\dagger$$

$$= L \int_{-\pi}^{\pi} \begin{bmatrix} I_{m+1} & zI_{m+1} & \cdots & z^n I_{m+1} \\ z^{-1} I_{m+1} & I_{m+1} & \cdots & z^{n-1} I_{m+1} \\ \vdots & \vdots & \cdots & \vdots \\ z^{-n} I_{m+1} & z^{-n+1} I_{m+1} & \cdots & I_{m+1} \end{bmatrix} d\Omega = LC_{n,m}L^\dagger,$$

where I is the $(n+1)(m+1) \times (n+1)(m+1)$ identity matrix. Since $C_{n,m}$ is invertible we find,

$$C_{n,m}^{-1} = L^\dagger L.$$

The result for R follows in an analogous manner. □

From this formula and (3.10) we find,

$$L_n^m(z) = \left[(L_{n,n}^m{}^\dagger)^{-1}, 0, 0, \dots 0 \right] C_{n,m}^{-1} [z^n I_{m+1}, z^{n-1} I_{m+1}, \dots, I_{m+1}]^T, \tag{3.14}$$

and

$$R_n^m(z) = \left[I_{m+1}, z I_{m+1}, \dots, z^n I_{m+1} \right] C_{n,m}^{-1} \left[0, 0, \dots, 0, (\bar{R}_{n,n}^m)^{-1} \right]^T. \tag{3.15}$$

A consequence of (3.14) and (3.15) is that $L_{n,n}^m{}^\dagger$ is the lower Cholesky factor of $[I_{m+1}, 0, \cdots, 0] C_{n,m}^{-1} [I_{m+1}, 0, \cdots, 0]^T$ while $R_{n,n}^m$ is the upper Cholesky factor of $[0, \cdots, I_{m+1}] C_{n,m}^{-1} [0, \cdots, I_{m+1}]^T$.

Define a matrix inner product as

$$\langle F, G \rangle_\Omega = \int_{-\pi}^{\pi} F(\theta) d\Omega(\theta) G(\theta)^\dagger = \langle G, F \rangle_\Omega^\dagger, \tag{3.16}$$

whenever the above product is well defined and finite for example for F and G $(m+1) \times (m+1)$ complex matrix continuous functions. If $B_n(z)$ is an $(m+1) \times (m+1)$ matrix polynomial of degree n the reverse polynomial is given by $B_n^*(z) = z^n B_n^\dagger(1/z)$. Similar to the scalar case L_n^m (R_n^m) and L^{*m}_n (R^{*m}_n) are characterized up to a multiplication of an $(m+1) \times (m+1)$ matrix by the orthogonality relations

$$\langle L_n^m, z^i \rangle_\Omega = 0 = \langle z^{-i}, (R_n^m)^\dagger \rangle_\Omega, \qquad z = e^{i\theta},\ 0 \le i \le n-1, \tag{3.17}$$

and

$$\langle z^{-i}, (L^{*m}_n)^\dagger \rangle_\Omega = 0 = \langle R^{*m}_n, z^i \rangle_\Omega, \qquad z = e^{i\theta},\ 1 \le i \le n, \tag{3.18}$$

The above equation now allow recurrence formulas to be computed. If we write

$$\Pi = A_{n,m} L_n^m(z) - z L_{n-1}^m(z),$$

with

$$A_{n,m} = L_{n-1,n-1}^m (L_{n,n}^m)^{-1} \tag{3.19}$$

it follows that Π is an $(m+1) \times (m+1)$ matrix polynomial of degree $n-1$ in z satisfying

$$\langle \Pi, z^i \rangle_\Omega = A_{n,m} \langle L_n^m, z^i \rangle_\Omega - \langle L_{n-1}^m, z^{i-1} \rangle_\Omega = 0, \qquad 1 \le i \le n-1.$$

Thus from (3.18) Π is equal to an $(m+1)\times(m+1)$ matrix B times R_n^{m*} and we write $B=-E_{n,m}$ which gives

$$A_{n,m}L_n^m(z)=zL_{n-1}^m(z)-E_{n,m}R^{*m}_{n-1}(z), \tag{3.20}$$

so that the right matrix orthogonal polynomials are needed to obtain a recurrence formula. A similar argument gives

$$R_n^m(z)\hat{A}_{n,m}=zR_{n-1}^m(z)-L^{*m}_{n-1}(z)E_{n,m}. \tag{3.21}$$

That $E_{n,m}$ can be chosen in both formulas follows since (3.20) gives

$$E_{n,m}=\langle zL_{n-1}^m, R^{*m}_{n-1}\rangle_\Omega=\langle (L^{*m}_{n-1})^\dagger, (zR_{n-1}^m)^\dagger\rangle_\Omega. \tag{3.22}$$

which is what (3.21) yields. The second equality follows from $zL_{n-1}^m = z^n(L^{*m}_{n-1}(z))^\dagger$ and $(R^{*m}_{n-1}(z))^\dagger = z^{-n+1}R_{n-1}^m(z)$. As in the scalar case it follows from orthogonality that,

$$A_{n,m}A_{n,m}^\dagger=I_{m+1}-E_{n,m}E_{n,m}^\dagger \text{ and } \hat{A}_{n,m}^\dagger\hat{A}_{n,m}=I_{m+1}-E_{n,m}^\dagger E_{n,m}. \tag{3.23}$$

The above equations and the properties of $A_{n,m}$ and $\hat{A}_{n,m}$ show that $E_{n,m}$ is a strictly contractive matrix and that $A_{n,m}$ is the upper Cholesky factor of $I_{m+1}-E_{n,m}E_{n,m}^\dagger$. Similarly $\hat{A}_{n,m}^\dagger$ is the lower Cholesky factor of $I_{m+1}-E_{n,m}^\dagger E_{n,m}$.

The recurrence formulas (3.20) and (3.21) can be inverted in the following manner. Multiply the reverse of (3.21) on the right by $E_{n,m}$ to obtain

$$E_{n,m}\hat{A}_{n,m}^\dagger R^{*m}_n(z)=E_{n,m}R^{*m}_{n-1}(z)-zE_{n,m}E_{n,m}^\dagger L_{n-1}^m(z).$$

Add this equation to (3.20) then use (3.23) and eliminate $A_{n,m}$ to find,

$$L_n^m(z)+\tilde{E}_{n,m}R^{*m}_n(z)=zA_{n,m}^\dagger L_{n-1}^m(z), \tag{3.24}$$

where

$$\tilde{E}_{n,m}=A_{n,m}^{-1}E_{n,m}\hat{A}_{n,m}^\dagger. \tag{3.25}$$

In a similar manner we find

$$R_n^m(z)+L^{*m}_n(z)\tilde{E}_{n,m}=zR_{n-1}^m(z)\hat{A}_{n,m}^\dagger, \tag{3.26}$$

where we have the remarkable formula

$$\tilde{E}_{n,m}=A_{n,m}^\dagger E_{n,m}\hat{A}_{n,m}^{-1}. \tag{3.27}$$

To show that it is indeed $\tilde{E}_{n,m}$ in Eq. (3.27) integrate (3.24) against $(R^{*m}_n(z))^\dagger$ to obtain

$$\tilde{E}_{n,m} = -\langle L^m_n, R^{*m}_n\rangle_\Omega. \tag{3.28}$$

Since $L^m_n(z) = z^{-n}(L^{*m}_n(z))^\dagger$ and from the definition of R^{*m}_n we find that the above equation equals

$$\tilde{E}_{n,m} = -\langle (L^{*m}_n)^\dagger, (R^m_n)^\dagger\rangle_\Omega, \tag{3.29}$$

which gives the result. Note that using the orthogonality properties of the orthonormal matrix polynomials and the above integral for $\tilde{E}_{n,m}$ show that

$$A^\dagger_{n,m}A_{n,m} = I_{m+1} - \tilde{E}_{n,m}\tilde{E}^\dagger_{n,m} \text{ and } \hat{A}_{n,m}\hat{A}^\dagger_{n,m} = I_{m+1} - \tilde{E}^\dagger_{n,m}\tilde{E}_{n,m}. \tag{3.30}$$

From the structure of A^m_n and $\hat{A}^m_n$ we see that $A^\dagger_{n,m}$ is the lower Cholesky factor of $I_{m+1} - \tilde{E}_{n,m}\tilde{E}^\dagger_{n,m}$ while $\hat{A}_{n,m}$ is the upper Cholesky factor of $I_{m+1} - \tilde{E}^\dagger_{n,m}\tilde{E}_{n,m}$

The recurrences (3.20) and (3.21) and Eq. (3.23) give a matrix Christoffel–Darboux formula. From the reverse of (3.20) it follows that

$$\begin{aligned}
&L^{*m}_n(z)L^{*m}_n(z_1)^\dagger \\
&= L^{*m}_{n-1}(z)(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}L^{*m}_{n-1}(z_1)^\dagger - \bar{z}_1 L^{*m}_{n-1}(z)(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}E_{n,m}R^{*m}_{n-1}(z_1)^\dagger \\
&-zR^m_{n-1}(z)E^\dagger_{n,m}(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}L^m_n(z_1)^\dagger + z\bar{z}_1R^m_{n-1}(z)E^\dagger_{n,m}(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}E^\dagger_{n,m}R^m_{n-1}(z_1)^\dagger.
\end{aligned}$$

Likewise

$$\begin{aligned}
&R^m_n(z)R^m_n(z_1)^\dagger \\
&= z\bar{z}_1R^m_{n-1}(z)\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}R^m_{n-1}(z_1)^\dagger - zR^m_{n-1}(z)\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}E^\dagger_{n,m}L^{*m}_{n-1}(z_1)^\dagger \\
&- \bar{z}_1L^{*m}_{n-1}(z)E_{n,m}\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}L^m_{n-1}(z)^\dagger + L^{*m}_{n-1}(z)E_{n,m}\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}E^\dagger_{n,m}L^m_{n-1}(z_1)^\dagger.
\end{aligned}$$

The subtraction of the second equation from the first equation yields

$$\begin{aligned}
&L^{*m}_n(z)L^{*m}_n(z_1)^\dagger - R^m_n(z)R^m_n(z_1)^\dagger \\
&= L^{*m}_{n-1}(z)((A^\dagger_{n,m})^{-1}(A_{n,m})^{-1} - E_{n,m}\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}E^\dagger_{n,m})L^{*m}_{n-1}(z_1)^\dagger \\
&- \bar{z}_1L^{*m}_{n-1}(z)((A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}E_{n,m} - E_{n,m}\hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1})R^{*m}_{n-1}(z_1)^\dagger \\
&- zR^m_{n-1}(z)(E^\dagger_{n,m}(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1} - \hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1}E^\dagger_{n,m})L^m_n(z_1)^\dagger \\
&+ z\bar{z}_1R^m_{n-1}(z)(E^\dagger_{n,m}(A^\dagger_{n,m})^{-1}(A_{n,m})^{-1}E^\dagger_{n,m} - \hat{A}^{-1}_{n,m}(\hat{A}^\dagger_{n,m})^{-1})R^m_{n-1}(z_1)^\dagger.
\end{aligned}$$

Observe that (3.23) implies

$$\begin{aligned}E_{n,m}\hat{A}_{n,m}^{-1}(\hat{A}_{n,m}^{\dagger})^{-1}E_{n,m}^{\dagger} &= E_{n,m}(I_{m+1}-E_{n,m}^{\dagger}E_{n,m})^{-1}E_{n,m}^{\dagger}\\ &= (I_{m+1}-E_{n,m}E_{n,m}^{\dagger})^{-1}E_{n,m}E_{n,m}^{\dagger} = (A_{n,m}^{\dagger})^{-1}(A_{n,m})^{-1}E_{n,m}E_{n,m}^{\dagger},\end{aligned}$$

so that

$$(A_{n,m}^{\dagger})^{-1}(A_{n,m})^{-1} - E_{n,m}\hat{A}_{n,m}^{-1}(\hat{A}_{n,m}^{\dagger})^{-1}E_{n,m}^{\dagger} = I_{m+1}.$$

Also

$$E_{n,m}\hat{A}_{n,m}^{-1}(\hat{A}_{n,m}^{\dagger})^{-1} = (A_{n,m}^{\dagger})^{-1}(A_{n,m})^{-1}E_{n,m}$$

so the second term on the right hand side is equal to zero. Applying similar manipulations to the remaining terms yields

$$\begin{aligned}L^{*m}_{n}(z)L^{*m}_{n}(z_1)^{\dagger} - R_n^m(z)R_n^m(z_1)^{\dagger} &= L^{*m}_{n-1}(z)L^{*m}_{n-1}(z_1)^{\dagger} - R_{n-1}^m(z)R_{n-1}^m(z_1)^{\dagger}\\ &\quad + (1-z\bar{z}_1)R_{n-1}^m(z)R_{n-1}^m(z_1)^{\dagger}.\end{aligned}$$

Iteration of this equation to zero then adding $(1-z\bar{z}_1)R_n^m(z)\mathbb{R}_n^m(z_1)^{\dagger}$ to both sides of the above equation and using the fact that $L^{*m}_{0}(z)L^{*m}_{0}(z_1)^{\dagger} - R_0^m(z)R_0^m(z_1)^{\dagger} = C_0^{-1}-C_0^{-1}=0$ gives the first Christoffel–Darboux formula

$$L^{*m}_{n}(z)L^{*m}_{n}(z_1)^{\dagger} - z\bar{z}_1R_n^m(z)R_n^m(z_1)^{\dagger} = (1-z\bar{z}_1)\sum_{i=0}^{n}R_i^m(z)R_i^m(z_1)^{\dagger}. \tag{3.31}$$

Likewise the second Christoffel–Darboux formula is given by,

$$R^{*m}_{n}(z)^{\dagger}R^{*m}_{n}(z_1) - \bar{z}z_1L_n^m(z)^{\dagger}L_n^m(z_1) = (1-\bar{z}z_1)\sum_{i=0}^{n}L_i^m(z)^{\dagger}L_i^m(z_1). \tag{3.32}$$

We say $A \geq B$ for $(m+1)\times(m+1)$ matrices if $A-B$ is positive semidefinite. We say that a square matrix polynomial $A(z)$ has a zero at z_0 if there is a nonzero vector $u \in \mathbb{C}^{m+1}$ such that $A(z_0)u_0 = 0$. This is equivalent to $\det(A(z_0)) = 0$ which is equivalent to $\ker(A(z_0))$ being nontrivial. As in the case of scalar orthogonal polynomials on the unit circle L^{*m}_{n} and R^{*m}_{n} have the following remarkable properties.

Theorem 3.2 *Given L_n^m and R_n^m as above then for $n \geq 0$*

$$\det(R^{*m}_{n}(z)) \neq 0 \neq \det(L^{*m}_{n}(z)), \qquad |z| \leq 1. \tag{3.33}$$

If

$$W_n(z) = \left[L^{*m}_{n}(z)L^{*m}_{n}(z)^{\dagger}\right]^{-1}, \tag{3.34}$$

and

$$C^n_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ij\theta} W_n(e^{i\theta})d\theta,$$

then

$$C^n_j = C_j, \quad |j| \le n. \tag{3.35}$$

Furthermore from (3.37)

$$W_n(z) = \left[R^{*m}_{n}(z)^{\dagger} R^{*m}_{n}(z)\right]^{-1}, \qquad |z| = 1. \tag{3.36}$$

We say that $L^{*m}_{n}(z)$ and $R^{*m}_{n}(z)$ are stable and have spectral matching.

Proof With $z_1 = z$ Eq. (3.31) implies

$$L^{*m}_{n}(z)L^{*m}_{n}(z)^{\dagger} \ge (1-\bar{z}z_1)\sum_{i=0}^{n} R^m_i(z)R^m_i(z_1)^{\dagger} \ge (1-\bar{z}z_1)R^0_i(z_1)R^0_i(z)^{\dagger} > 0.$$

Thus $\det(L^m_n(z)) \ne 0$ for $|z| < 1$. For $|z| = 1$ (3.31) shows that

$$L^{*m}_{n}(z)L^{*m}_{n}(z)^{\dagger} = R^m_n(z)R^m_n(z)^{\dagger}, \tag{3.37}$$

so if there is a nonzero row vector so that $u_0L^{*m}_{n}(z_0) = 0$ then it follows that $u_0R^m_n(z_0) = 0$. Suppose now that $\det(L^{*m}_{n}(z_0)) = 0$ for $|z_0| = 1$ then there is a nonzero row vector u_0 so that $u_0L^{*m}_{n}(z_0) = 0$. Therefore from the above discussion it follows

$$u_0L^{*m}_{n}(z_0)(L^{*m}_{n}(z_1))^{\dagger} - z_0\bar{z}_1u_0R^m_n(z_0)R^m_n(z_1)^{\dagger} = 0,$$

hence

$$\sum_{i=0}^{n-1} u_0R^m_i(z_0)R^m_i(z_1)^{\dagger} = 0, \qquad z_1 \ne z_0.$$

Multiplying by $d\Omega$ and integrating, the continuity of the polynomials in the sum, yields,

$$0=\sum_{i=0}^{n-1} u_0 R_i^m(z_0)\int_{-\pi}^{\pi} R_i^m(z_1)^\dagger d\Omega(\theta_1)=u_0 R_0^m(z_0)\int_{-\pi}^{\pi} R_0^m(z_1)^\dagger d\Omega(\theta_1)=u_0 I_{m+1},$$

where (3.6) has been used. This is a contradiction so $\det(L^{*m}_n(z))\neq 0$ for $|z|\leq 1$ which gives the stability of $L^{*m}_n(z)$. The stability of $R^{*m}_n(z)$ follows in a similar manner from Eq. (3.32).

To show the spectral matching note that on the unit circle $W_n(z) = L_n^m(z)^{-1}((L_n^m(z))^\dagger)^{-1}$ so

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} L_n^m(z)W_n(z)(L_n^m(z))^\dagger d\theta=1.$$

Also

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} L_n^m(z)z^{-i}W_n(z)d\theta=\frac{1}{2\pi}\int_{-\pi}^{\pi} z^{n-i}(L^{*m}_n(z))^{-1}d\theta.$$

Writing the above as a contour integral counterclockwise around the unit circle and using the fact that $(L^{*m}_n(z))^{-1}$ is an analytic matrix function for $|z|\leq 1$ gives that the above integral is equal to zero for $0\leq i\leq n-1$ so L_n^m is an orthonormal polynomial associated with $W_n\frac{d\theta}{2\pi}$. Likewise with the use of (3.37) it is seen that R_n^m is an orthonormal polynomial with respect to $W_n\frac{d\theta}{2\pi}$. From (3.26) with $z=0$ we find that $\tilde{E}_{n,m}=-L^{*m}_n(0)^{-1}R_n^m(0)$. Furthermore $A_{n,m}$ can be obtained uniquely from (3.30). Thus L_{n-1}^m can be written in terms of L_n^m and R_n^m. Likewise the same holds true for R_{n-1}^m using (3.26) and (3.30). Thus it follows that

$$\langle L_{n-1}^m, z^i\rangle_\Omega=(A_{n,m}^\dagger)^{-1}(\langle L_n^m, z^{i+1}\rangle_\Omega+\tilde{E}_{n,m}\langle R^{*m}_n, z^{i+1}\rangle_\Omega)=0,$$

for $i=0,\dots n-2$. Substituting (3.24) into $\langle L_{n-1}^m, L_{n-1}^m\rangle$ then using (3.28) and (3.30) yields

$$\langle L_{n-1}^m, L_{n-1}^m\rangle_\Omega=(A_{n,m}^\dagger)^{-1}(I_{m+1}-\tilde{E}_{n,m}\tilde{E}_{n,m}^\dagger)A_{n,m}^{-1}=I_{m+1}.$$

This shows that L_{n-1}^m is an orthonormal matrix polynomial associated with $W_n\frac{d\theta}{2\pi}$. Similarly R_{n-1}^m can be shown to be an orthonormal matrix polynomial associated with $W_n\frac{d\theta}{2\pi}$ using (3.26), (3.29), (3.30), and (3.36). Continuing in this manner shows that $\{L_i^m\}_{i=0}^n$ and $\{R_i^m\}_{i=0}^n$ are orthonormal matrix polynomials associated with $W_n\frac{d\theta}{2\pi}$. Since $z^i I_{m+1}$ can be expanded in terms of L_k^m, $k=0,\dots,i$ the above argument shows that $C_{-j}^n=C_{-j}$ for $j=0,\dots n$. Taking the complex conjugate transpose gives the result for $j=0,\dots,-n$ and the result follows. □

Since we are dealing with a matrix measure it is not easy to use the Cauchy–Schwarz inequality as in the scalar case. In order to overcome this we will make use of an entropy principle. We begin with the following simple equality,

Theorem 3.3 *Let $L_n^m(z)$ be the left orthonormal polynomial discussed above and W_n be given by (3.34). Then*

$$\log(\det((L_{n,n}^m)^\dagger L_{n,n}^m)) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\det(W_n(\theta)))d\theta. \tag{3.38}$$

Proof From Eq. (3.34) it follows that

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\det(W_n(\theta)))d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log(|\det(L^{*m}_n(z))|^2)d\theta.$$

Since $L^{*m}_n(z)$ is nonzero inside and on the unit circle and is also analytic there so is $\det(L^{*m}_n(z))$, hence $\log(\det(L^{*m}_n(z)))$ is analytic for $|z| < 1$ and we can write $\log(|\det(L^{*m}_n(z))|^2) = \Re \log(\det(L^{*m}_n(z)))$ which is the real part of an analytic function and therefore a harmonic function. Thus by the mean value theorem for harmonic functions.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(|\det(L^{*m}_n(z))|^2)d\theta = \Re \log(\det(L^{*m}_n(0))) = \log(\det((L_{n,n}^m)^\dagger L_{n,n}^m)),$$

which gives the result. □

From Eq. (3.23) and the definition of $A_{n,m}$ we have

$$\begin{aligned}\det((L_{n,n}^m)^\dagger L_{n,n}^m) &= \det(I - E_{n,m}E_{n,m}^\dagger)^{-1}\det((L_{n-1,n-1}^m)^\dagger L_{n-1,n-1}^m)\\ &\geq \det((L_{n-1,n-1}^m)^\dagger L_{n-1,n-1}^m),\end{aligned}$$

since $E_{n,m}$ is a contraction. The use of the above theorem yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\det(W_n(\theta)))d\theta \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\det(W_{n-1}(\theta))). \tag{3.39}$$

There is an important result, Szegő's theorem for matrix measures, that can be developed from the above results. Write $\Omega = \Omega_{ac} + \Omega_s$ where Ω_{ac} is the absolutely continuous part of Ω and Ω_s is its singular part. Write $\frac{d\Omega_{ac}}{d\theta} = W(\theta)$ then Szegő's theorem says that

Theorem 3.4 *With the notation above*

$$\lim_{n\to\infty} \log(\det((L_{n,n}^m)^\dagger L_{n,n}^m)) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\det(W(\theta)))d\theta. \tag{3.40}$$

This gives rise to a *maximum entropy principle* (see [2]). Let S_n^m be the class of $(m+1)\times(m+1)$ matrix Borel measures Ω on the unit circle with $\int_{-\pi}^{\pi}\log(\det(W(\theta)))d\theta > -\infty$ and such that each $\Omega \in S_n^m$ has the same Fourier coefficients $C_i,\ |i|\le n$. With $\mathcal{E}(\Omega)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\det(W)d\theta$ we have,

Theorem 3.5 *There is a unique measure $\Omega_0 \in S_m^n$ that satisfies*

$$\mathcal{E}(\Omega_0)=\sup_{\Omega\in S_n^m}\mathcal{E}(\Omega)$$

and this measure is absolutely continuous with respect to Lebesgue measure and is given by $d\Omega_0 = W(\theta)\frac{d\theta}{2\pi}$ with $W(\theta)=Q_n^m(\theta)^{-1}$ where $Q_n^m(\theta)$ is a positive $(m+1)\times(m+1)$ matrix trigonometric polynomial of degree n.

Proof Any measure Ω in the class S_n^m has the same moments $C_j,\ |j|\le n$ and therefore the same left orthonormal matrix polynomials L_i^m for $0\le i\le n$. With $\frac{d\Omega_{ac}}{d\theta}=W(\theta)$ it follows from Theorem 3.4 and Eq. (3.39) that $\int_{-\pi}^{\pi}\log\det(W(\theta))d\theta \le \int_{-\pi}^{\pi}\log\det(W_n(\theta))d\theta$ where W_n is given by (3.34). Thus Ω such that $\frac{d\Omega}{d\theta}=W_n$ maximizes $\mathcal{E}$ on S_n^m and from (3.34) W^{-1} is a strictly positive $(m+1)\times(m+1)$ matrix trigonometric polynomial of degree n. Hence the result is proved. □

With the above it is now possible to prove the matrix analog of Verblunsky's theorem and a restricted Fejér-Riesz theorem. We begin with the Fejér-Riesz theorem.

Theorem 3.6 *Let $Q_n^m(\theta)$ be an $(m+1)\times(m+1)$ matrix trigonometric polynomial of degree n. Then $Q_n(\theta)>0$ if and only if there is a stable $(m+1)\times(m+1)$ matrix polynomial $P_n(z)$ of degree n in z such that*

$$Q_n(\theta)=P_n(z)P_n(z)^{\dagger},\qquad z=e^{i\theta}.$$

Proof Since Q_n is strictly positive let Ω be the positive Borel measure such that $d\Omega = Q_n^{-1}\frac{d\theta}{2\pi}$. Let L_n^m be the nth degree left orthonormal polynomial. Then Ω has the same Fourier coefficients as Ω_n where $d\Omega_n(\theta)=[L^{*m}_n(z)L^{*m}_n(z)^{\dagger}]^{-1}\frac{d\theta}{2\pi}$. So by the maximum entropy principle (Theorem 3.4) it follows that $Q_n(\theta)=L^{*m}_n(z)L^{*m}_n(z)^{\dagger}$. □

A measure Ω is said to be an $(m+1)\times(m+1)$ positive matrix probability measure if it is a positive $(m+1)\times(m+1)$ matrix measure such that $\int_{-\pi}^{\pi}d\Omega = I_{m+1}$.

We finish with the matrix version of Verblunsky's theorem.

Theorem 3.7 *There is a one-to-one correspondence between infinite sequences $\{E_i^m\}_{i=1}^{\infty}$ with each E_i^m an $(m+1)\times(m+1)$ strictly contractive matrix and $(m+1)\times(m+1)$ matrix Borel probability measures Ω on the unit circle such that (3.3) holds.*

Proof Starting with Ω we see that if (3.3) holds then the left and right orthonormal polynomials $\{L_n^m\}$ and $\{R_n^m\}$ can be constructed. Matrix recurrence coefficients are now found via (3.22) and the orthogonality properties of $L_n^m,\ L_{n-1}^m$ and R^{*m}_{n-1}

give Eq. (3.23) which shows that E_n^m is a strict contraction for $1 \leq n < \infty$. Now suppose that the $\{E_n^m\}$ are given, where E_n^m is a strict contraction for all $n \geq 1$. Beginning with $L_0^m = I_{m+1} = R_0^m$ use Eqs. (3.20) and (3.21) to construct $\{L_n^m\}_{n=0}^\infty$ and $\{R_n^m\}_{n=0}^\infty$. The Christoffel–Darboux equations (3.31) and (3.32) show that ${L^*}_n^m$ and ${R^*}_n^m$ are stable for all n and have spectral matching. Define Ω_n as $d\Omega_n(\theta) = W_n(\theta)\frac{d\theta}{2\pi}$ where W_n is given by Eq. (3.34). By the weak compactness of the $(m+1)\times(m+1)$ matrix valued probability measures (i.e. the Helly theorems) the Ω_n have a weak limit Ω which is a probability measure which satisfies (3.3). From the spectral matching of $\{\Omega_n\}$ the left and right matrix orthonormal polynomials associated with Ω are $\{L_n^m\}_{n=0}^\infty$ and $\{R_n^m\}_{n=0}^\infty$ respectively. Thus its recurrence coefficients are given by $\{E_n^m\}_{n=0}^\infty$ and the result follows. $\square$

4 Orthogonal Polynomials on the Bicircle

We now examine polynomials orthogonal on the bicircle. The first problem encountered is what ordering is to be used to order the monomials for the Gram–Schmidt procedure. Each ordering gives a different set of orthonormal polynomials and unlike the univariate case there does not appear to be a preferred ordering. Jackson suggested using the total degree ordering and this indeed may be the most preferable ordering to use to study the asymptotics of orthonormal polynomials. However the structure of the moment matrix makes it difficult to find recurrences among the polynomials. Delsarte et. al. [5] in their study of the planar least squares problem considered the lexicographical ordering. This ordering is also implicitly used in the work of Helson and Lowdenslager [13] in their study of the two dimensional prediction problem. In this lecture we will order the moments using the lexicographical (lex) or reverse lexicographical (revlex) orderings. Given a subset of $\mathbb{Z}^2$ the *lexicographical ordering* is defined by

$$(k,l) <_{lex} (k_1,l_1) \iff k < k_1 \text{ or } (k = k_1, \text{and } l < l_1).$$

The *reverse lexicographical ordering* is defined by

$$(k,l) <_{revlex} (k_1,l_1) \iff l < l_1 \text{ or } (l = l_1, \text{and } k < k_1).$$

Given a positive bivariate Borel measure μ supported on $[-\pi,\pi]\times[\pi,\pi]$ or $\mathbb{T}^2$ let $c_{j,k}$ be defined as

$$c_{j,k} = \int_{\mathbb{T}^2} e^{-ij\theta} e^{-ik\phi} d\mu(\theta,\phi), \tag{4.1}$$

and

$$\bar{c}_{j,k} = c_{-j,-k}.$$

If we start with the monomials $\{z^i w^j\}$, $0 \le i \le n$, $0 \le j \le m$, $z = e^{i\theta}$, $w = e^{i\phi}$, then the monomials lexicographically ordered are $1, w, \ldots, w^m, z, zw, \ldots, z^n w^m$ and the moment matrix $C_{n,m} = \int_{\mathbb{T}^2} [1, w, \ldots, z^n w^m]^\dagger [1, w, \ldots, z^n w^m] d\mu(\theta, \phi)$ is given by

$$C_{n,m} = \begin{bmatrix} C_0 & C_{-1} & \cdots & C_{-n} \\ C_1 & C_0 & \cdots & C_{-n+1} \\ \vdots & & \ddots & \vdots \\ C_n & C_{n-1} & \cdots & C_0 \end{bmatrix}, \tag{4.2}$$

where each C_i is an $(m + 1) \times (m + 1)$ Toeplitz matrix as follows

$$C_i = \begin{bmatrix} c_{i,0} & c_{i,-1} & \cdots & c_{i,-m} \\ c_{i,1} & c_{i,0} & \cdots & c_{i,-m+1} \\ \vdots & & \ddots & \vdots \\ c_{i,m} & & \cdots & c_{i,0} \end{bmatrix}, \qquad i = -n, \ldots, n. \tag{4.3}$$

Thus $C_{n,m}$ in the lexicographical ordering has a doubly Toeplitz structure. It is shown in Lemma 2.2 of [9], see also [5], that

$$J C_{n,m} J = C_{n,m}^T \tag{4.4}$$

where A^T denotes the transpose of the matrix A and J is an $(n + 1)(m + 1) \times (n + 1)(m + 1)$ matrix with ones along the antidiagonal and zeros elsewhere. If the reverse lexicographical ordering is used in place of the lexicographical ordering we obtain another moment matrix $\tilde{C}_{n,m}$ where the roles of n and m are interchanged. We will assume that $C_{n,m} > 0$ for all $n \ge 0$, $m \ge 0$ which is equivalent to

$$\int_{\mathbb{T}^2} |p(z, w)|^2 d\mu(\theta, \phi) > 0, \qquad z = e^{i\theta}, \ w = e^{i\phi} \tag{4.5}$$

for any nonzero polynomial in z and w with complex coefficients. As in the scalar case this can be seen by writing a polynomial of degree n in z and m in w as

$$\begin{aligned} p(z, w) &= p_{n,m} z^n w^m + p_{n,m-1} z^n w^{m-1} + \ldots + p_{0,0} \\ &= [p_{n,m}, p_{n,m-1}, \ldots, p_{0,0}][z^n w^m, z^n w^{m-1}, \ldots, 0]^T, \end{aligned}$$

in which case

$$\int_{\mathbb{T}^2} |p(z, w)|^2 d\mu(\theta, \phi) = [p_{n,m}, p_{n,m-1}, \ldots, p_{0,0}] C_{n,m} [p_{n,m}, p_{n,m-1}, \ldots, p_{0,0}]^\dagger.$$

As in the scalar case define an inner product as

$$\langle f, g\rangle_\mu = \int_{\mathbb{T}^2} f(\theta,\phi)\bar{g}(\theta,\phi)d\mu(\theta,\phi) = \overline{\langle g, f\rangle}_\mu. \tag{4.6}$$

We now construct the orthonormal polynomials by performing the Gram–Schmidt procedure on the lexicographically ordered monomials and define the orthonormal polynomials $\{\phi^l_{n,m}(z,w)\}_{n\ge 0, 0\le l\le m}$, by the equations,

$$\begin{aligned} &\langle \phi^l_{n,m} z^i w^j\rangle_\mu = 0, \quad 0 \le i < n \text{ and } 0 \le j \le m \text{ or } i = n \text{ and } 0 \le j < l,\\ &\langle \phi^l_{n,m}\phi^l_{n,m}\rangle_\mu = 1, \end{aligned} \tag{4.7}$$

and

$$\phi^l_{n,m}(z,w) = k^{n,l}_{n,m,l} z^n w^l + \sum_{(i,j)<_{\text{lex}}(n,l)} k^{i,j}_{n,m,l} z^i w^j. \tag{4.8}$$

With the convention $k^{n,l}_{n,m,l} > 0$, the above equations uniquely specify $\phi^l_{n,m}$. The following equation is a formula for $\phi^l_{n,m}$,

$$\phi^l_{n,m}(z,w) = \frac{1}{\sqrt{\det C^l_{n,m} \det \hat{C}^l_{n,m}}} \begin{vmatrix} c_{0,0} & c_{0,-1} & \cdots & c_{-n,-l}\\ c_{0,1} & c_{0,0} & \cdots & c_{-n,-l+1}\\ \vdots & & \ddots & \vdots\\ c_{n,l-1} & & \cdots & c_{n,-1}\\ 1 & & \cdots & z^n w^l \end{vmatrix},$$

where $C^l_{n,m}$ is the matrix in the above determinant except the last row is $c_{n,l},\dots,c_{0,0}$ and $\hat{C}^l_{n,m}$ is obtained from $C^l_{n,m}$ by removing its last row and column. Polynomials orthonormal with respect to μ but using the reverse lexicographical ordering will be denoted by $\tilde{\phi}^l_{n,m}$. They are uniquely determined by the above relations with the roles of n and m interchanged.

Set

$$\Phi_{n,m} = \begin{bmatrix} \phi^m_{n,m}\\ \phi^{m-1}_{n,m}\\ \vdots\\ \phi^0_{n,m} \end{bmatrix} = K_{n,m} \begin{bmatrix} z^n w^m\\ z^n w^{m-1}\\ \vdots\\ 1 \end{bmatrix}, \tag{4.9}$$

where the $(m+1) \times (n+1)(m+1)$ matrix $K_{n,m}$ is given by

$$K_{n,m} = \begin{bmatrix} k_{n,m,m}^{n,m} & k_{n,m,m}^{n,m-1} & \cdots & \cdots & \cdots & k_{n,m,m}^{0,0} \\ 0 & k_{n,m,m-1}^{n,m-1} & \cdots & \cdots & \cdots & k_{n,m,m-1}^{n,0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & k_{n,m,0}^{n,0} & k_{n,m,0}^{n-1,m} & \cdots & k_{n,m,0}^{0,0} \end{bmatrix}. \tag{4.10}$$

Similar as above denote

$$\tilde{\Phi}_{n,m} = \begin{bmatrix} \tilde{\phi}_{n,m}^{n} \\ \tilde{\phi}_{n,m}^{n-1} \\ \vdots \\ \tilde{\phi}_{n,m}^{0} \end{bmatrix}. \tag{4.11}$$

For the bivariate polynomials $\phi_{n,m}^{l}(z,w)$ above we define the reverse polynomials $\phi_{n,m}^{l}(z,w)$ by the relation

$$\phi^{*l}_{n,m}(z,w) = z^n w^m \bar{\phi}_{n,m}^{l}(1/z, 1/w). \tag{4.12}$$

With this definition $\phi^{*l}_{n,m}(z,w)$ is again a polynomial in z and w, and furthermore

$$\Phi^{*}_{n,m}(z,w) := z^n w^m \Phi_{n,m}^{\dagger}(1/z, 1/w) = \begin{bmatrix} \phi^{*m}_{n,m} \\ \phi^{*m-1}_{n,m} \\ \vdots \\ \phi^{*0}_{n,m} \end{bmatrix}. \tag{4.13}$$

An analogous procedure is used to define $\tilde{\phi}^{*l}_{n,m}$.

Let $\hat{\prod}^{n,m}$ be the linear span of $z^i w^j$, $0 \le i \le n, 0 \le j \le m$, $\hat{\prod}_k^{n,m}$ be the vector space of k dimensional vectors with entries in $\hat{\prod}^{n,m}$, and $\hat{\prod}_{m+1}^{m} = \hat{\prod}_{m+1}^{\infty,m}$.

The orthogonality relations (4.7) allow us to characterize $\Phi_{n,m}$.

Lemma 4.1 *Suppose $\Phi \in \hat{\prod}_k^{n,m}$. If Φ satisfies the orthogonality relations,*

$$\langle \Phi, z^i w^j \rangle_\mu = 0, \quad 0 \le i < n, \quad 0 \le j \le m, \tag{4.14}$$

then $\Phi = T\Phi_{n,m}$, where T is a $k \times (m+1)$ matrix. If $k = m+1$, T is upper triangular with positive diagonal entries, and if $\langle \Phi, \Phi \rangle_\mu = I_{m+1}$, then $T = I_{m+1}$. Likewise $\Phi^{*}_{n,m}(z,w)$ is characterized up to multiplication by a constant matrix by the relations

$$\langle \Phi^{*}, z^i w^j \rangle_\mu = 0, \qquad 0 < i \le n, \quad 0 \le j \le m. \tag{4.15}$$

Proof Let ϕ be a nonzero element of Φ, then (4.14) implies that ϕ is in the subspace spanned by linear combinations of the elements of $\Phi_{n,m}$ so that the first part of the Lemma follows. To see the second part note that the orthonormality of Φ and $\Phi_{n,m}$ imply that $TT^{\dagger} = I_{m+1}$. That T is upper triangular with positive diagonal entries implies that $T = I_{m+1}$. Equation (4.15) is obtained by use of the definition of $\Phi^*{}_{n,m}$ and the first part of the Lemma. □

Lemma 4.2 *Suppose $\tilde{\Phi} \in \hat{\prod}_k^{n,m}$. If $\tilde{\Phi}$ satisfies the orthogonality relations,*

$$\langle \tilde{\Phi}, z^i w^j \rangle_{\mu} = 0, \qquad 0 \le i \le n, \quad 0 \le j < m, \tag{4.16}$$

then $\tilde{\Phi} = T\tilde{\Phi}_{n,m}$, where T is an $k \times (n+1)$ matrix. If $k = n+1$, T is upper triangular with positive diagonal entries, and if $\langle \tilde{\Phi}, \tilde{\Phi} \rangle_{\mu} = I_{n+1}$, then $T = I_{n+1}$. Likewise $(\tilde{\Phi}^*_{n.m})^T(z, w)$ is characterized up to multiplication by a constant matrix by the relations

$$\langle \tilde{\Phi}^*, z^i w^j \rangle_{\mu} = 0, \quad 0 \le i \le n, \quad 0 < j \le m. \tag{4.17}$$

With the above we can make contact with matrix orthogonal polynomials.

Lemma 4.3 *Let $\Phi_{n,m}$ be given by (4.9). Then*

$$\Phi_{n,m}(z, w) = L_n^m(z)[w^m, w^{m-1}, \dots, 1]^T, \tag{4.18}$$

$$\Phi^*{}_{n,m}(z, w) = [1, w, \dots, w^m] J_m R^{*m}_n(z)^T J_m, \tag{4.19}$$

where J_m is the $(m+1) \times (m+1)$ matrix with ones on the antidiagonal and zeros everywhere else. Here $L_n^m(z)$ and $R_n^m(z)$ are respectively the left and right matrix polynomials defined in Eqs. (3.4)–(3.7) orthonormal with respect to the matrix measure

$$d\Omega(\theta) = \int_{-\pi}^{\pi} [w^m, \cdots, 1]^T [w^{-m}, \cdots, 1] d\mu(\theta, \phi) \qquad w = e^{i\phi}. \tag{4.20}$$

Furthermore

$$\begin{bmatrix} \Phi_{n,m}(z,w) \\ \Phi_{n-1,m}(z,w) \\ \vdots \\ \Phi_{0,m}(z,w) \end{bmatrix} = \begin{bmatrix} L_n^m(z) \\ L_{n-1}^m(z) \\ \vdots \\ L_0^m(z) \end{bmatrix} [w^m, w^{m-1}, \dots, 1]^T$$

$$= L \begin{bmatrix} z^n I_{m+1} \\ z^{n-1} I_{m+1} \\ \vdots \\ I_{m+1} \end{bmatrix} [w^m, w^{m-1}, \dots, 1]^T, \tag{4.21}$$

where $L^\dagger$ is the lower Cholesky factor of $C_{n,m}^{-1}$.

Note that the matrix measure $d\Omega$ has the structure of a Toeplitz matrix.

Proof If we substitute the equation

$$\Phi_{n,m} = \hat{L}_n(z)[w^m \ \cdots \ 1]^T$$

into (4.14), where $\hat{L}_n(z)$ is an $(m+1)\times(m+1)$ matrix polynomial of degree n, we find, for $j = 0, \ldots, n-1$,

$$0 = \left\langle \Phi_{n,m},\ z^j \begin{bmatrix} w^m \\ \vdots \\ 1 \end{bmatrix} \right\rangle_\mu = \int_{-\pi}^{\pi} \hat{L}_n^m(z) z^{-j} d\Omega(\theta), \qquad z = e^{i\theta}$$

and

$$I_{m+1} = \langle \Phi_{n,m}, \Phi_{n,m} \rangle_\mu = \int_{-\pi}^{\pi} \hat{L}_n^m(z) d\Omega(\theta) \hat{L}_n^m(z)^\dagger, \qquad z = e^{i\theta}.$$

Since $\hat{L}_{n,n}^m$ is upper triangular with positive diagonal entries (see Eq. (4.10)) Eqs. (3.17) and (3.7) show that $\hat{L}_n^m(z) = L_n^m(z)$. Equation (4.19) follows from the fact that the double Toeplitz structure of $C_{n,m}$ implies that

$$L_n^m(z)^T = J_m R_n^m J_m \tag{4.22}$$

(see equation (2.36) in [9] also [5]) where J_m is the $(m+1)\times(m+1)$ matrix with ones on the antidiagonal and zeros everywhere else. Equation (4.21) follows from (4.18) and the final assertion from Lemma 3.1. □

Since the left and right matrix orthonormal polynomials satisfy recurrences so will $\Phi_{n,m}$ and $\tilde{\Phi}_{n,m}$. However there is more structure in these polynomials giving rise to more recurrences which we now discuss.

Theorem 4.4 *Given $\{\Phi_{n,m}\}$ and $\{\tilde{\Phi}_{n,m}\}$, $0 \le n \le N$, $0 \le m \le M$, the following recurrence formulas hold*

$$A_{n,m}\Phi_{n,m} = z\Phi_{n-1,m} - \hat{E}_{n,m}\Phi^{*T}_{n-1,m}, \tag{4.23}$$

$$\Phi_{n,m} + A_{n,m}^\dagger \hat{E}_{n,m}(A_{n,m}^T)^{-1}\Phi^{*T}_{n,m} = A_{n,m}^\dagger z\Phi_{n-1,m}, \tag{4.24}$$

$$\Gamma_{n,m}\Phi_{n,m} = \Phi_{n,m-1} - \mathcal{K}_{n,m}\tilde{\Phi}_{n-1,m}, \tag{4.25}$$

$$\Gamma^1_{n,m}\Phi_{n,m} = w\Phi_{n,m-1} - \mathcal{K}^1_{n,m}(\tilde{\Phi}^*_{n-1,m})^T, \tag{4.26}$$

$$\Phi_{n,m} = I_{n,m}\tilde{\Phi}_{n,m} + \Gamma^\dagger_{n,m}\Phi_{n,m-1}, \tag{4.27}$$

$$\Phi^{*T}_{n,m} = I^1_{n,m}\tilde{\Phi}_{n,m} + (\Gamma^1_{n,m})^T\Phi^{*T}_{n,m-1}, \tag{4.28}$$

where

$$\hat{E}_{n,m} = \langle z\Phi_{n-1,m}, \Phi^{*T}_{n-1,m}\rangle = \hat{E}^T_{n,m} \in M^{m+1,m+1}, \tag{4.29}$$

$$A_{n,m} = \langle z\Phi_{n-1,m}, \Phi^T_{n,m}\rangle \in M^{m+1,m+1}, \tag{4.30}$$

$$\mathcal{K}_{n,m} = \langle \Phi_{n,m-1}, \tilde{\Phi}_{n-1,m}\rangle \in M^{m,n}, \tag{4.31}$$

$$\Gamma_{n,m} = \langle \Phi_{n,m-1}, \Phi_{n,m}\rangle \in M^{m,m+1}, \tag{4.32}$$

$$\mathcal{K}^1_{n,m} = \langle w\Phi_{n,m-1}, (\tilde{\Phi}^*_{n-1,m})^T\rangle \in M^{m,n}, \tag{4.33}$$

$$\Gamma^1_{n,m} = \langle w\Phi_{n,m-1}, \Phi_{n,m}\rangle \in M^{m,m+1}, \tag{4.34}$$

$$I_{n,m} = \langle \Phi_{n,m}, \tilde{\Phi}_{n,m}\rangle \in M^{m+1,n+1}, \tag{4.35}$$

$$I^1_{n,m} = \langle \Phi^{*T}_{n,m}, \tilde{\Phi}_{n,m}\rangle \in M^{m+1,n+1}. \tag{4.36}$$

Remark 4.5 Formulas similar to (4.23)–(4.28) hold for $\tilde{\Phi}_{n,m}$ and will be denoted by (4.23)–(4.28). In the rest of the lecture we use the same notation to denote the extension to $\tilde{\Phi}_{n,m}$ of existing formulas stated for $\Phi_{n,m}$.

Proof To prove Eq. (4.23) use the representation of $\Phi_{n,m}$ given in (4.18) and choose $A_{n,m}$ as in Eq. (3.19). Then $A_{n,m}\Phi_{n,m} - z\Phi_{n-1,m}$ is an $(m+1)$ column vector polynomial of degree $n-1$ in z and satisfies

$$\langle A_{n,m}\Phi_{n,m} - z\Phi_{n-1,m}, z^i w^j\rangle = 0,\ 0 < i \le n,\ 0 \le j \le m,$$

and the result follows from (4.15)

To prove (4.25) note that, because of the linear independence of the entries of $\Phi_{n,m}$, there is an $m \times (m+1)$ matrix $\Gamma_{n,m}$ such that $\Gamma_{n,m}\Phi_{n,m} - \Phi_{n,m-1} \in \hat{\prod}^{n-1,m}_m$. Furthermore

$$\langle \Gamma_{n,m}\Phi_{n,m} - \Phi_{n,m-1}, z^i w^j\rangle = 0, \qquad 0 \le i \le n-1 \quad 0 \le j \le m-1.$$

Thus (4.16) implies that

$$\Gamma_{n,m}\Phi_{n,m} - \Phi_{n,m-1} = H_{n,m}\tilde{\Phi}_{n-1,m},$$

and we set $H_{n,m} = -\mathcal{K}_{n,m}$. The remaining recurrence formulas follow in a similar manner. □

Remark 4.6 Formulas (4.23) and (4.24) follow from the theory of matrix orthogonal polynomials and so allow us to compute in the n direction along a strip of size $m+1$. This formula does not mix the polynomials in the two orderings. However, to increase m by one for polynomials constructed in the lexicographical ordering, the remaining relations show that orthogonal polynomials in the reverse lexicographical ordering must be used.

Using the orthogonality relations from Lemmas 4.1, 4.2 and Eq. (4.7) we find the following relations.

Lemma 4.7 *The following relations hold between the coefficients in the equations for $\tilde{\Phi}$ and Φ,*

$$\tilde{\mathcal{K}}_{n,m} = \mathcal{K}_{n,m}^{\dagger}, \ \tilde{I}_{n,m} = I_{n,m}^{\dagger}, \tag{4.37}$$

$$\tilde{I}_{n,m}^{1} = (I_{n,m}^{1})^{T}, \tilde{\mathcal{K}}_{n,m}^{1} = (\mathcal{K}_{n,m}^{1})^{T}, \tag{4.38}$$

Also

$$A_{n,m}A_{n,m}^{\dagger} = I_{m+1} - \hat{E}_{n,m}\hat{E}_{n,m}^{\dagger}, \tag{4.39}$$

$$\Gamma_{n,m}\Gamma_{n,m}^{\dagger} = I_{m} - \mathcal{K}_{n,m}\mathcal{K}_{n,m}^{\dagger}, \tag{4.40}$$

$$\Gamma_{n,m}^{1}(\Gamma_{n,m}^{1})^{\dagger} = I_{m} - \mathcal{K}_{n,m}^{1}(\mathcal{K}_{n,m}^{1})^{\dagger}, \tag{4.41}$$

$$I_{n,m}I_{n,m}^{\dagger} + \Gamma_{n,m}^{\dagger}\Gamma_{n,m} = I_{m+1}, \tag{4.42}$$

$$I_{n,m}^{1}(I_{n,m}^{1})^{\dagger} + (\Gamma_{n,m}^{1})^{\dagger}\Gamma_{n,m}^{1} = I_{m+1}. \tag{4.43}$$

Remark 4.8 The matrix $\Gamma_{n,m}$ has a zero in the entries $(i, j), i \geq j$ and has positive $(i, i+1)$ entries. Since $\Gamma_{n,m}\Gamma_{n,m}^{\dagger} = \Gamma_{n,m}U_m^{\dagger}U_m\Gamma_{n,m}^{\dagger}$ where U_m is the $m \times m+1$ matrix given by

$$U_m = \left[0, \ I_m\right], \tag{4.44}$$

we see that $\Gamma_{n,m}U_m^{\dagger}$ is the upper Cholesky factorization of the right hand side of (4.40). From this it is easy to obtain $\Gamma_{n,m}$. The matrix $\Gamma_{n,m}^{1}$ has zeroes in the entries $(i, j), i > j$ with positive (i, i) entries. The matrix $I_{n,m}$ has first row and column equal to zero except for a 1 in the $(1, 1)$ entry.

In the one variable case there is a one-to-one correspondence between the number of new Fourier coefficients C_n and the size of the recurrence coefficients $E_{n,m}$ in going from level $n-1$ to level n each being an $(m+1) \times (m+1)$ matrix. Unfortunately in the bivariate case this does not hold. More precisely given all the Fourier coefficients $c_{i,j}, 0 \leq i \leq n-1, \ 0 \leq j \leq m$ and for $i = n, 0 \leq j \leq m-1$ only two new Fourier coefficients $c_{n,m}$ and $c_{n,-m}$ are needed to go to the (n, m) level. However the coefficients in the recurrence formulas above are typically much larger. Thus there must be relations between the recurrence coefficients on the (n, m) level and those at the $(n, m-1)$ or $(n-1, m)$ level that determine most of the entries. This is indeed the case, for example,

Lemma 4.9 (Relations for $\mathcal{K}^1_{n,m}$) *For* $0 < n, m$,

$$\Gamma_{n,m-1}\mathcal{K}^1_{n,m} = \mathcal{K}^1_{n,m-1}(\tilde{A}^{-1}_{n-1,m})^T - \mathcal{K}_{n,m-1}\hat{\hat{E}}_{n-1,m}(\tilde{A}^{-1}_{n-1,m})^T, \tag{4.45}$$

$$\mathcal{K}^1_{n,m}(\tilde{\Gamma}_{n-1,m})^T = A^{-1}_{n,m-1}\mathcal{K}^1_{n-1,m} - A^{-1}_{n,m-1}\hat{E}_{n,m-1}\bar{\mathcal{K}}_{n-1,m}. \tag{4.46}$$

Proof To show (4.45) multiply (4.33) on the left by $\Gamma_{n,m-1}$ then use (4.25) to obtain

$$\Gamma_{n,m-1}\mathcal{K}^1_{n,m} = \langle w\Phi_{n,m-2}, (\tilde{\Phi}^*_{n-1,m})^T\rangle.$$

Now use (4.23) with n reduced by one and then Eqs. (4.33) and (4.31) yields (4.45). Equation (4.46) follows in a similar manner. □

It is shown in [9] that the above relations determine all the entries in $\mathcal{K}^1_{n,m}$ except for the $(1, 1)$ entry which we will denote as $\bar{u}_{n,m}$. Likewise we have,

Lemma 4.10 (Relations for $\mathcal{K}_{n,m}$) *For* $0 < n, m$,

$$\Gamma^1_{n,m-1}\mathcal{K}_{n,m} = \mathcal{K}_{n,m-1}(\tilde{A}^{-1}_{n-1,m})^\dagger - \mathcal{K}^1_{n,m-1}\hat{\hat{E}}^\dagger_{n-1,m}(\tilde{A}^{-1}_{n-1,m})^\dagger, \tag{4.47}$$

$$\mathcal{K}_{n,m}(\tilde{\Gamma}^1_{n-1,m})^\dagger = A^{-1}_{n,m-1}\mathcal{K}_{n-1,m} - A^{-1}_{n,m-1}\hat{E}_{n,m-1}\bar{\mathcal{K}}^1_{n-1,m}. \tag{4.48}$$

Proof To show Eq. (4.47) multiply (4.31) on the left by $\Gamma^1_{n,m-1}$ then use (4.26) with m reduced by one to obtain,

$$\Gamma^1_{n,m-1}\mathcal{K}_{n,m} = \langle w\Phi_{n,m-2}, \tilde{\Phi}_{n-1,m}\rangle.$$

Eliminating $\tilde{\Phi}_{n-1,m}$ using (4.23) then applying (4.31) and (4.33) gives (4.47). Equation (4.48) follows in a analogous manner. □

The above relations determine all of $(\Phi^{m-1}_{n,n})^{-1}\mathcal{K}_{n,m}((\tilde{\Phi}^{n-1}_{m,m})^\dagger)^{-1}$ except for the (n, m) entry which we will denote as $\bar{u}_{n,-m}$. Here $\Phi^{m-1}_{n,n}$ is the coefficient of $z^n[w^{m-1}, \dots 1]^T$ in $\Phi_{n,m-1}$ and $\tilde{\Phi}^{n-1}_{m,m}$ is the coefficient of $w^m[z^{n-1}, \dots 1]^T$ in $\tilde{\Phi}_{n-1,m}$ which are assumed to be known. The complex conjugates of $u_{n,m}$ and $u_{n,-m}$ are chosen since it is the complex conjugate of $c_{n,m}$ and $c_{n,-m}$ that appear in the Fourier expansion of the entries discussed above. Since $\mathcal{K}_{n,m}$ and $\mathcal{K}^1_{n,m}$ are contractions, $u_{n,m}$ and $u_{n,-m}$ must be in magnitude less than one and once they are fixed all the other recurrence coefficients on the (n, m) level may be computed. This gives an alternate parameterization of the two variable trigonometric moment problem other than the Fourier coefficients. A full list of equations giving the redundancies and an algorithm to compute the recurrence coefficients, taking into account the constraints imposed by Eqs. (4.39)–(4.43), is given in [9]. In the one variable matrix or scalar case there are constraints that the entries in the recurrence coefficients must satisfy. But from level $n-1$ to level n the constraints are independent so it is always possible to choose new parameters consistent with the constraints imposed. This is not the case for two variables. It may happen that all the constraints on the $(n-1, m)$ and $(m, n-1)$ level are satisfied but there is no choice of parameters that will satisfy the constraints on the (n, m) level. Theorem 6.2 in [9] gives a parameterization and conditions the parameters must satisfy in order for a positive measure on the bicircle to exist and is therefore a two variable Verblunsky type theorem.

As in the one variable case $\Phi_{n,m}$ satisfies a Christoffel–Darboux type formula.

Lemma 4.11 *Given* $\{\Phi_{n,m}\}$ *and* $\{\tilde{\Phi}_{n,m}\}$,

$$\Phi^*{}_{n,m}(z,w)\Phi^*{}^{\dagger}_{n,m}(z_1,w_1) - \bar{z}_1 z\Phi^T_{n,m}(z,w)\Phi^{\dagger}_{n,m}(z_1,w_1)^T \tag{4.49a}$$

$$= (1-\bar{z}_1 z)\Phi_{n,m}(z,w)^T\,\Phi^{\dagger}_{n,m}(z_1,w_1)^T$$
$$+\,\Phi^*{}_{n-1,m}(z,w)\Phi^*{}^{\dagger}_{n-1,m}(z_1,w_1) - \bar{z}_1 z\Phi^T_{n-1,m}(z,w)\Phi^{\dagger}_{n-1,m}(z_1,w_1)^T \tag{4.49b}$$

$$= (1-\bar{z}_1 z)\tilde{\Phi}_{n,m}(z,w)^T\,\tilde{\Phi}^{\dagger}_{n,m}(z_1,w_1)^T$$
$$+\,\Phi^*{}_{n,m-1}(z,w)\Phi^*{}_{n,m-1}(z_1,w_1)^T - \bar{z}_1 z\Phi_{n,m-1}(z,w)^T\,\Phi^{\dagger}_{n,m-1}(z_1,w_1)^T. \tag{4.49c}$$

Proof To obtain the first equality (4.49a) = (4.49b) follows by subtracting (3.31) with n reduced by one from the original equation to obtain

$$L^{*m}_{n}(z)L^{*m}_{n}(z_1)^{\dagger} - z\bar{z}_1 R^m_n(z)R^m_n(z_1)^{\dagger} = (1-z\bar{z}_1)R^m_n(z)R^m_n(z_1)^{\dagger}$$
$$L^{*m}_{n-1}(z)L^{*m}_{n-1}(z_1)^{\dagger} - z\bar{z}_1 R^m_{n-1}(z)R^m_{n-1}(z_1)^{\dagger}.$$

Multiply the above equation on the left by $[1,\ldots,w^m]$ and on the right by $[1,\ldots,w^m]^{\dagger}$. Equation (4.22) shows $(L^m_n(z_1)^{\dagger}L^m_n(z))^T = J_m R^m_n(z)R^m_n(z_1)^{\dagger}J_m$. Then using Eq. (4.18) and that $[1,\ldots,w^m]J_m = [w^m,\ldots,1]$ gives the stated equality. The equality (4.49a) = (4.49c) can be obtained in the following manner. Let

$$Z_{n,m}(z,w) = [1, w, \ldots, w^m][I_{m+1}, zI_{m+1}, \ldots, z^n I_{m+1}],$$

and $\tilde{Z}_{n,m}(z,w)$ be given by a similar formula with the roles of z and w, and n and m interchanged. Multiply Eq. (3.31) on the left by $[1, w, \ldots, w^m]$ and on the right by its complex conjugate transpose. Then performing the manipulations mentioned above give

$$\frac{\Phi^*{}_{n,m}(z,w)\Phi^*{}^{\dagger}_{n,m}(z_1,w_1) - \bar{z}_1 z\Phi^T_{n,m}(z,w)\Phi^{\dagger}_{n,m}(z_1,w_1)^T}{1-\bar{z}_1 z}$$
$$= Z_{n,m}(z,w)J(L^{\dagger}L)^T JZ_{n,m}(z_1,w_1)^{\dagger}$$

$$= Z_{n,m}(z,w)C_{n,m}^{-1}Z_{n,m}(z_1,w_1)^\dagger = \tilde Z_{n,m}(z,w)\tilde C_{n,m}^{-1}\tilde Z_{n,m}(z_1,w_1)^\dagger$$
$$= \tilde\Phi_{n,m}^T(z,w)\tilde\Phi_{n,m}^\dagger(z_1,w_1)^T + \tilde Z_{n,m-1}(z,w)\tilde C_{n,m-1}^{-1}\tilde Z_{n,m-1}(z_1,w_1)^\dagger,$$

where Lemma 4.3 and Eq. (4.4) have been used to obtain the first equality above. Switching back to the lexicographical ordering in the second term in the last equation then using Lemma 3.1 yields the result. □

The above formula has important consequences for $\phi_{n,m}^m$

We say that a polynomial $p(z,w)$ is stable if $p(z,w)\neq 0,\ |z|\leq 1,\ |w|\leq 1$. A polynomial p is of degree (n,m) if

$$p(z,w)=\sum_{i=0}^{n}\sum_{j=0}^{m}k_{i,j}z^iw^j,$$

with $k_{n,m}\neq 0$. Finally we say that the polynomial $p_{n,m}$ of degree (n,m) has the spectral matching property (up to (n,m) with respect to μ) if

$$\int_{\mathbb{T}^2} z^kw^j d\mu(\theta,\phi)=\frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\frac{z^kw^j}{|p_{n,m}(z,w)|^2}d\theta d\phi,\qquad z=e^{i\theta}\ ,w=e^{i\phi},$$

for $|k|\leq n,\ |j|\leq m$.

Lemma 4.12 *Suppose that $\mathcal{K}_{n,m}=0$, then,*

$$\phi^{*m}_{n,m}(z,w)\overline{\phi^{*m}_{n,m}(z_1,w_1)}-\phi^m_{n,m}(z,w)\overline{\phi^m_{n,m}(z_1,w_1)}$$
$$=(1-w\bar w_1)\Phi^*_{n,m-1}(z,w)\Phi^{*\dagger}_{n,m-1}(z_1,w_1)$$
$$+(1-z\bar z_1)\tilde\Phi_{n-1,m}(z,w)^T\tilde\Phi^\dagger_{n-1,m}(z_1,w_1). \tag{4.50}$$

Proof If $\mathcal{K}_{n,m}=0$ then (4.25) shows that $\Gamma_{n,m}\Phi_{n,m}(z,w)=\Phi_{n,m-1}(z,w)$. Thus $\Phi^*_{n,m}(z,w)\Gamma^\dagger_{n,m} = w\Phi^*_{n,m-1}(z,w)$. Also (4.40) implies that $(\Gamma_{n,m})_{(i,i+1)} = 1,\ i=1,\ldots,m$ with all other entries zero. Thus we find

$$\Phi^*_{n,m}(z,w)\Phi^*_{n,m}(z_1,w_1)^\dagger$$
$$=\phi^{*m}_{n,m}(z,w)\overline{\phi^{*m}_{n,m}(z_1,w_1)}+\Phi^*_{n,m}(z,w)\Gamma^\dagger_{n,m}\Gamma_{n,m}\Phi^*_{n,m}(z_1,w_1)^\dagger$$
$$=\phi^{*m}_{n,m}(z,w)\overline{\phi^{*m}_{n,m}(z_1,w_1)}+w\bar w_1\Phi^*_{n,m-1}(z,w)\Phi^*_{n,m-1}(z_1,w_1)^\dagger. \tag{4.51}$$

From (4.49c) we find

$$\phi^{*m}_{n,m}(z,w)\overline{\phi^{*m}_{n,m}(z_1,w_1)}-z\bar z_1\phi^m_{n,m}(z,w)\overline{\phi^m_{n,m}(z_1,w_1)}$$
$$=(1-w\bar w_1)\Phi^*_{n,m-1}(z,w)\Phi^{*\dagger}_{n,m-1}(z_1,w_1)$$
$$+(1-z\bar z_1)\tilde\Phi_{n,m}(z,w)^T\tilde\Phi^\dagger_{n,m}(z_1,w_1).$$

Using (4.25) and the fact that $\tilde{\phi}^n_{n,m}(z, w) = \phi^m_{n,m}$ gives the result. □

There are many criteria used to prove the stability of a bivariate polynomial. The one we will use was formulated by Strintzis (see [6]):

A polynomial $p(z, w)$ is nonzero for $|z| \leq 1$ and $|w| \leq 1$ if and only if $p(z, w) \neq 0$ for all $|z| = 1$ and all $|w| \leq 1$ and $p(z, w_0) \neq 0$ for all $|z| \leq 1$ and some $|w_0| \leq 1$.

We now prove,

Theorem 4.13 *If $\mathcal{K}_{n,m} = 0$ then ${\phi^*}_{n,m}$ is a stable polynomial.*

Proof Set $z_1 = z$ and $w_1 = w,\ |z| = 1$ in Eq. (4.50) to obtain

$$
\begin{aligned}
|{\phi^*}_{n,m}(z, w)|^2 &\geq (1 - |w|^2){\Phi^*}_{n-1,m}(z, w){\Phi^*}_{n-1,m}(z, w)^\dagger \\
&= (1 - |w|^2)[1, \ldots, w^m]L^{*m}_n(w)L^{*m}_n(w)^\dagger[1, \ldots, w^m]^\dagger.
\end{aligned}
$$

Thus if $\overline{\phi}_{n,m}(z_0, w_0) = 0$ in the region $|z| = 1$ and $|w| < 1$ then $\det(L^{*m}_n(z_0)) = 0$ with $|z_0| = 1$ which cannot happen from Theorem 3.2. In a similar fashion the roles of z and w can be interchanged and $\tilde{\Phi}_{n,m}$ can be used to show that ${\phi^*}_{n,m}(z, w) \neq 0$ for $|z| < 1$ and $|w| = 1$. Suppose that ${\phi^*}_{n,m}(z_0, w_0) = 0$ with $|z_0| = 1 = |w_0|$. Then $\phi_{n,m}(z_0, w_0) = 0$ also and with $z_1 = z_0$ (4.50) gives

$$
0 = (1 - w_0\bar{w}_1){\Phi^*}_{n,m}(z_0, w_0){\Phi^*}^\dagger_{n,m}(z_0, w_1)
$$

for all $|w_1| \leq 1$ with $w_1 \neq w_0$. Again this says that $\det(L^{*m}_n(z_0)) = 0$ which contradicts Theorem 3.2. □

We now examine the orthogonality properties of stable polynomials (see [11]). Let $p_{n,m} \in \mathbb{C}[z, w]$ be stable with degree n in z and m in w. We will use the following partial order on pairs of integers:

$$
(k, l) \leq (i, j) \text{ iff } k \leq i \text{ and } l \leq j.
$$

The notations $\nleq, \ngeq$ refer to the negations of the above partial order.

The polynomial $p_{n,m}$ is orthogonal to more monomials than the one variable theory might initially suggest. More precisely with

$$
d\mu_{n,m} = \frac{1}{4\pi^2}\frac{d\theta d\phi}{|p_{n,m}|^2},
$$

we have:

Lemma 4.14 *In $L^2(\mu_{n,m})$, $p_{n,m}$ is orthogonal to the set*

$$
\{z^i w^j : (i, j) \nleq (0, 0)\}
$$

and $p^*_{n,m}$ *is orthogonal to the set*

$$\{z^i w^j : (i, j) \ngeq (n, m)\}.$$

Proof Observe that since $1/p_{n,m}$ is holomorphic in $\overline{\mathbb{D}^2}$

$$\begin{aligned}\langle z^i w^j, p_{n,m}\rangle_{\mu_{n,m}} &= \int_{\mathbb{T}^2} z^i w^j \overline{p_{n,m}(z, w)} \frac{d\sigma}{|p_{n,m}(z, w)|^2} \\ &= \int_{\mathbb{T}^2} \frac{z^i w^j}{p_{n,m}(z, w)} d\sigma = 0 \text{ if } (i, j) \nleq (0, 0)\end{aligned}$$

by the mean value property (either integrating first with respect to z or w depending on whether $i > 0$ or $j > 0$). The claim about $p^*_{n,m}$ follows from the observation

$$\langle z^i w^j, p^*_{n,m}\rangle_{\mu_{n,m}} = \langle p_{n,m}, z^{n-i} w^{m-j}\rangle_{\mu_{n,m}}.$$

□

Write $p_{n,m}(z, w) = \sum_{i=0}^m p_i(z) w^i$. Then a straightforward computation gives

$$\frac{p(z, w)\overline{p(z, w_1)} - w\bar{w}_1 p^*(z, w)\overline{p*(z, w_1)}}{1 - w\bar{w}_1} = \begin{bmatrix} 1 & w & \cdots & w^m \end{bmatrix} T_p(z) \begin{bmatrix} 1 \\ \bar{w}_1 \\ \vdots \\ \bar{w}_1^m \end{bmatrix}, \tag{4.52}$$

where

$$\begin{aligned} T_p(z) = &\begin{bmatrix} p_0(z) & & \bigcirc \\ p_1(z) & \ddots & \\ \vdots & & \\ p_m(z) & \cdots & p_0(z) \end{bmatrix} \begin{bmatrix} \bar{p}_0(1/z) & \bar{p}_1(1/z) & \ldots & \bar{p}_m(1/z) \\ & & \ddots & \\ & & & \\ \bigcirc & & & \bar{p}_0(1/z) \end{bmatrix} \\ &- \begin{bmatrix} 0 & & & \bigcirc \\ \bar{p}_m(1/z) & \ddots & & \\ \vdots & & & \\ \bar{p}_1(1/z) & \cdots & \bar{p}_m(1/z) & 0 \end{bmatrix} \begin{bmatrix} 0 & p_m(z) & \ldots & p_1(z) \\ & \ddots & & \vdots \\ & & & p_m(z) \\ \bigcirc & & & 0 \end{bmatrix}. \end{aligned} \tag{4.53}$$

Since $p_{n,m}$ is stable it follows from the one variable Christoffel–Darboux formula or alternatively the Schur-Cohn test [8, p. 850] for stability that the $m + 1 \times m + 1$ matrix $T_p(z)$ is positive definite for $|z| = 1$.

Define the following parametrized version of a one variable Christoffel–Darboux kernel

$$
\begin{aligned}
L(z, w; \eta) &= z^n \frac{p_{n,m}(z, w)\overline{p_{n,m}(1/\bar{z}, \eta)} - p^*_{n,m}(z, w)\overline{p^*_{n,m}(1/\bar{z}, \eta)}}{1 - w\bar{\eta}} \\
&= z^n[1, \ldots, w^{m-1}]T_p(z)[1, \ldots, \eta^{m-1}]^\dagger \\
&= \sum_{j=0}^{m-1} a_j(z, w)\bar{\eta}^j,
\end{aligned}
\tag{4.54}
$$

where $a_j(z, w)$, $j = 0, \ldots, m-1$ are polynomials in (z, w), as the following lemma shows in addition to several other important observations.

Lemma 4.15 *Let $p_{n,m}(z, w)$ be a stable polynomial of degree (n, m). Then,*

(1) *L is a polynomial of degree $(2n, m-1)$ in (z, w) and a polynomial of degree $m-1$ in $\bar{\eta}$.*
(2) *$L(\cdot, \cdot; \eta)$ spans a subspace of dimension m as η varies over $\mathbb{C}$.*
(3) *L is symmetric in the sense that*

$$L(z, w; \eta) = z^{2n}(w\bar{\eta})^{m-1}\overline{L(1/\bar{z}, 1/\bar{w}; 1/\bar{\eta})},$$

so $a_k = a^*_{m-k-1}$.
(4) *L can be written as*

$$L(z, w; \eta) = p_{n,m}(z, w)A(z, w; \eta) + p^*_{n,m}(z, w)B(z, w; \eta)$$

where A, B are polynomials of degree $(n, m-1, m-1)$ in $(z, w, \bar{\eta})$.

Proof The numerator of L vanishes when $w = 1/\bar{\eta}$, so the factor $(1 - w\bar{\eta})$ divides the numerator. This gives (1).

For (2), when $|z| = 1$ use Eq. (4.54). Since $T_p(z) > 0$ for $|z| = 1$, $L(z, w; \eta)$ spans a set of polynomials of dimension m.

For (3), this is just a computation.

For (4), observe that (suppressing the dependence of p on n and m),

$$
\begin{aligned}
&z^n \frac{p(z, w)\overline{p(1/\bar{z}, \eta)} - p^*(z, w)\overline{p^*(1/\bar{z}, \eta)}}{1 - w\bar{\eta}} \\
&= p(z, w)\underbrace{\left(\frac{\bar{\eta}^m p^*(z, 1/\bar{\eta}) - \bar{\eta}^m p^*(z, w)}{1 - w\bar{\eta}}\right)}_{A(z,w;\eta)} + p^*(z, w)\underbrace{\left(\frac{\bar{\eta}^m p(z, w) - \bar{\eta}^m p(z, 1/\bar{\eta})}{1 - w\bar{\eta}}\right)}_{B(z,w;\eta)}.
\end{aligned}
\tag{4.55}
$$

□

We now show that L and $a_0, \ldots, a_{m-1}$ possess a great many orthogonality relations in $L^2(\mu_{n,m})$.

Theorem 4.16 *In $L^2(\mu_{n,m})$, each a_k is orthogonal to the set*

$$\begin{aligned} \mathcal{O}_k =& \{z^i w^j : i > n, j < 0\} \\ & \cup \{z^i w^j : 0 \le j < m, j \ne k\} \\ & \cup \{z^i w^j : i < n, j \ge m\} \\ & \cup \{z^i w^k : i \ne n\}. \end{aligned}$$

In $L^2(\mu_{n,m})$, $L(\cdot, \cdot; \eta)$ is orthogonal to the set

$$\begin{aligned} \mathcal{O} =& \{z^i w^j : i > n, j < 0\} \\ & \cup \{z^i w^j : i \ne n, 0 \le j < m\} \\ & \cup \{z^i w^j : i < n, j \ge m\}. \end{aligned} \tag{4.56}$$

Write

$$A(z, w; \eta) = \sum_{j=0}^{m-1} A_j(z, w)\bar{\eta}^j \qquad B(z, w; \eta) = \sum_{j=0}^{m-1} B_j(z, w)\bar{\eta}^j.$$

Recall Eq. (4.54) and Lemma 4.15 item (4). By examining coefficients of $\bar{\eta}^j$ in L

$$a_j = p_{n,m} A_j + p^*_{n,m} B_j.$$

Also, A_j and B_j have at most degree j in w. To see this, recall Eq. (4.55) and observe that

$$A(z, w; \eta) = \sum_j p^*_j(z)\bar{\eta}^j \frac{1 - (w\bar{\eta})^{m-j}}{1 - w\bar{\eta}}$$

which shows that $A_j(z, w)$ has degree at most j in w (i.e., powers of w only occur next to greater powers of η). The same holds for B.

Proof By Lemma 4.14, $p_{n,m}$ is orthogonal to

$$\{z^i w^j : (i, j) \nleq (0, 0)\}$$

and since A_k has degree at most n in z and k in w,

$$p_{n,m} A_k \text{ is orthogonal to } \{z^i w^j : (i, j) \nleq (n, k)\}.$$

Also,

$$p^*_{n,m}B_k \text{ is orthogonal to } \{z^i w^j : (i,j) \ngeq (n,m)\}$$

since the orthogonality relation for $p^*_{n,m}$ (also from Lemma 4.14) is unaffected by multiplication by holomorphic monomials.

Hence, $a_k = p_{n,m}A_k + p^*_{n,m}B_k$ is orthogonal to the intersection of these sets; namely,

$$\{z^i w^j : (i,j) \nleq (n,k) \text{ and } (i,j) \ngeq (n,m)\}. \tag{4.57}$$

Since

$$a_{m-k-1} \perp \{z^i w^j : (i,j) \nleq (n, m-k-1) \text{ and } (i,j) \ngeq (n,m)\}$$

and since $a_k = a^*_{m-k-1} = z^{2n}w^{m-1}\overline{a_{m-k-1}(1/\bar{z}, 1/\bar{w})}$,

$$\begin{aligned} a_k \perp &\{z^{2n-i}w^{m-j-1} : (i,j) \nleq (n, m-k-1) \text{ and } (i,j) \ngeq (n,m)\} \\ &= \{z^i w^j : (n,k) \nleq (i,j) \text{ and } (n,-1) \ngeq (i,j)\}. \end{aligned} \tag{4.58}$$

Hence, a_k is orthogonal to the union of the sets in (4.57) and (4.58). The set in (4.58) contains $\{z^i w^j : i < n, j \geq 0\}$ and the set in (4.57) contains $\{z^i w^j : i > n, j \leq m-1\}$. Also, the set in (4.57) contains $\{z^n w^j : k < j \leq m-1\}$ while the set in (4.58) contains $\{z^n w^j : 0 \leq j < k\}$. Combining all of this we get $a_k \perp \mathcal{O}_k$.

Finally, L is orthogonal to the intersection of $\mathcal{O}_0, \ldots, \mathcal{O}_{m-1}$. □

With the above we have the important Corollary

Corollary 4.17 *In $L^2(1/|p_{n,m}|^2 d\sigma)$, the polynomial a_k is uniquely determined (up to unimodular multiples) by the conditions:*

$$a_k \in \operatorname{span}\{z^i w^j : (0,0) \leq (i,j) \leq (2n, m-1)\},$$

$$\begin{aligned} a_k \perp &\{z^i w^j : (0,0) \leq (i,j) \leq (2n, m-1), j \neq k\} \\ &\cup \{z^i w^k : 0 \leq i \leq 2n, i \neq n\}, \end{aligned}$$

and

$$||a_k||^2 = \int_{-\pi}^{\pi} T_{k,k}(e^{i\theta}, e^{i\theta}) \frac{d\theta}{2\pi}.$$

Only the last equation has not been proved (see [10]). The previous results imply the useful, *Remark 4.18* (1) each a_k is explicitly given from the coefficients of $p_{n,m}$, (2) each a_k is determined by the orthogonality relations in Corollary 4.17 and (3) each satisfies the additional orthogonality relations from Theorem 4.16. One useful consequence of this is that the set

$$\{z^j a_k(z,w) : j \in \mathbb{Z}, 0 \le k < m\}$$

is dual to the monomials

$$\{z^{j+n} w^k : j \in \mathbb{Z}, 0 \le k < m\}$$

within the subspace

$$S = \overline{\text{span}}\{z^j w^k : j \in \mathbb{Z}, 0 \le k < m\}.$$

Namely,

$$\langle z^{j_1+n} w^{k_1}, z^{j_2} a_{k_2} \rangle_{u_{n,m}} = 0$$

unless $j_1 = j_2$ and $k_1 = k_2$. In particular, if $f \in S$, then

$$f \perp z^j a_k \text{ implies } \hat{f}(j+n,k) = 0. \tag{4.59}$$

We now look at the space generated by shifting the a_k's by powers of z.

Theorem 4.19 *With respect to $L^2(\mu_{n,m})$,*

$$\begin{aligned} &\overline{\text{span}}\{z^i a_j(z,w) : 0 \le i, 0 \le j < m\} \\ &= \overline{\text{span}}\{z^i w^j : 0 \le i, 0 \le j < m\} \ominus \overline{\text{span}}\{z^i w^j : 0 \le i < n, 0 \le j < m\} \end{aligned} \tag{4.60}$$

and this is orthogonal to the larger set

$$\overline{\text{span}}\{z^i w^j : i < n, j \ge 0\}.$$

Proof Since the a_k are polynomials of degree at most $m-1$ in w, it is clear that

$$\overline{\text{span}}\{z^i a_j(z,w) : 0 \le i, 0 \le j < m\} \subset \overline{\text{span}}\{z^i w^j : 0 \le i, 0 \le j < m\}.$$

By Theorem 4.16, the a_k are orthogonal to the spaces

$$\overline{\text{span}}\{z^i w^j : i < n, j \ge 0\} \supset \overline{\text{span}}\{z^i w^j : i < n, 0 \le j < m\},$$

and since these spaces are invariant under multiplication by $\bar{z}$, the polynomials $z^i a_k$ are also orthogonal to these spaces for all $i \geq 0$. So,

$$\overline{\text{span}}\{z^i a_j(z,w) : 0 \leq i, 0 \leq j < m\} \perp \overline{\text{span}}\{z^i w^j : i < n, j \geq 0\}.$$

Therefore,

$$\begin{aligned}&\overline{\text{span}}\{z^k a_j(z,w) : 0 \leq k, 0 \leq j < m\}\\ &\subset \overline{\text{span}}\{z^i w^j : 0 \leq i, 0 \leq j < m\} \ominus \overline{\text{span}}\{z^i w^j : 0 \leq i < n, 0 \leq j < m\}\end{aligned} \tag{4.61}$$

and this containment must in fact be an equality. Indeed, any f in

$$\overline{\text{span}}\{z^i w^j : 0 \leq i, 0 \leq j < m\}$$

which is orthogonal to $\{z^k a_j(z,w) : 0 \leq k, 0 \leq j < m\}$ satisfies $\hat{f}(i,j) = 0$ for $i \geq n$ and $0 \leq j < m$ by Remark 4.18 and Eq. (4.59). Such an f cannot also be orthogonal to the space $\overline{\text{span}}\{z^i w^j : 0 \leq i < n, 0 \leq j < m\}$ without being identically zero. □

Define

$$\begin{aligned}H =& \text{span}\{z^i w^j : (0,0) \leq (i,j) \leq (n, m-1)\}\\ &\ominus \text{span}\{z^i w^j : (0,0) \leq (i,j) \leq (n-1, m-1)\}.\end{aligned}$$

From above H is an m dimensional space of polynomials contained in the space (4.60) of the previous theorem. In particular,

Corollary 4.20

$$H \perp \overline{\text{span}}\{z^i w^j : i < n, j \geq 0\}. \tag{4.62}$$

With the above we now prove,

Theorem 4.21 *Let μ be a positive measure on the bicircle and $\phi^m_{n,m}$ its associated orthonormal polynomial of degree (n,m). If $\mathcal{K}_{n,m} = 0$ then $\phi^{*m}_{n,m}(z,w)$ is stable and has spectral matching i.e.*

$$\int_{\mathbb{T}^2} e^{ik\theta} e^{il\phi} d\mu(\theta,\phi) = (\frac{1}{2\pi})^2 \int_{\mathbb{T}^2} \frac{e^{-ik\theta} e^{-il\phi}}{|\phi^{*m}_{n,m}(e^{i\theta}, e^{i\phi})|^2} d\theta d\phi, \qquad |k| \leq n,\ |l| \leq m. \tag{4.63}$$

Conversely if ${\phi^*}^m_{n,m}$ is stable and

$$\int_{\mathbb{T}^2} e^{-ik\theta}e^{-il\phi}d\mu(\theta,\phi) = (\frac{1}{2\pi})^2\int_{\mathbb{T}^2}\frac{e^{-ik\theta}e^{-il\phi}}{|{\phi^*}_{n,m}(e^{i\theta},e^{i\phi})|^2}d\theta d\phi, \quad |k|\le n,\ |l|\le m,$$

then $\mathcal{K}_{n,m}=0$.

Proof If $\mathcal{K}_{n,m}=0$ then Theorem 4.13 says that ${\phi^*}_{n,m}(z,w)$ is stable. So it remains to show the spectral matching. Let us denote by ${\phi^*}_l(z)$ the coefficient of w^l in ${\phi^*}_{n,m}(z,w)$, i.e. we set

$${\phi^*}_{n,m}(z,w)=\sum_{l=0}^m {\phi^*}_l(z)w^l. \tag{4.64}$$

From (4.52) we have

$$\frac{{\phi^*}_{n,m}(z,w)\overline{{\phi^*}_{n,m}(z,w_1)}-w\bar{w}_1\phi_{n,m}(z,w)\overline{\phi_{n,m}(z,w_1)}}{1-w\bar{w}_1}$$
$$=\begin{bmatrix}1 & w & \cdots & w^m\end{bmatrix}T_{\phi^*}(z)\begin{bmatrix}1\\ \bar{w}_1\\ \vdots\\ \bar{w}_1^m\end{bmatrix}. \tag{4.65}$$

Equation (4.18) shows that

$${\Phi^*}_{n,m}(z,w){\Phi^*}_{n,m}(z,w_1)^\dagger=\begin{bmatrix}1 & w & \cdots & w^m\end{bmatrix}{L^*}^m_n(z){L^*}^m_n(z)^\dagger\begin{bmatrix}1\\ \bar{w}_1\\ \vdots\\ \bar{w}_1^m\end{bmatrix}. \tag{4.66}$$

Thus from Eq. (4.50) it follows that for $|z|=1$ we have

$$T_{\phi^*}(z)={L^*}^m_n(z){L^*}^m_n(z)^\dagger, \tag{4.67}$$

where $L^m_n(z)$ is the left orthonormal matrix polynomial associated with μ. Let

$$c_l^\theta=\frac{1}{2\pi}\int_{-\pi}^{pi}e^{-il\phi}\frac{d\phi}{|{\phi^*}_{n,m}(z,w)|^2},\qquad z=e^{i\theta},\ w=e^{i\phi}. \tag{4.68}$$

Then c_l^θ is the lth parametric moment associated with the (parametric) measure $d\mu_m^\theta(\phi)=\frac{1}{2\pi}\frac{d\phi}{|\phi_{n,m}(e^{i\theta},w)|^2},\ w=e^{i\phi}$. Since ${\phi^*}_{n,m}(z,w)$ is nonzero for $|z|=1$ and

$|w| \le 1$ we see that for fixed $z = e^{i\theta}$ on the unit circle, $\phi_m^{\theta}(w) = w^m \overline{p(e^{i\theta}, 1/\bar{w})}$ is an orthonormal polynomial of degree m with respect to the (parametric) measure above. From the Gohberg-Semencul formula [14]

$$C_m^{\theta} = \begin{bmatrix} c_0^{\theta} & c_{-1}^{\theta} & \cdots & c_{-m}^{\theta} \\ c_1^{\theta} & c_0^{\theta} & \cdots & c_{-m+1}^{\theta} \\ \vdots & & \ddots & \vdots \\ c_m^{\theta} & c_{m-1}^{\theta} & \cdots & c_0^{\theta} \end{bmatrix}.$$

Equation (4.67) gives

$$C_m^{\theta} = [L^{*m}_n(z)\, L^{*m}_n(z)^{\dagger}]^{-1}, \qquad z = e^{i\theta}. \tag{4.69}$$

The spectral matching part of theorem (3.2) shows that the matrix weight on the right-hand side of (4.69) generates the same left and right matrix-valued orthonormal polynomials $\{L_k^m(z)\}_{0\le k\le n}$ and $\{R_k^m(z)\}_{0\le k\le n}$ as Ω given in Eq. (4.20) and therefore

$$\int_{-\pi}^{\pi} z^k d\Omega(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta} C_m^{\theta} d\theta, \qquad z = e^{i\theta}.$$

From the first row of the last matrix equation we find

$$\int_{\mathbb{T}^2} e^{ik\theta} e^{il\phi} d\mu(\theta,\phi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta} c_{-l}^{\theta} d\theta, \tag{4.70}$$

which gives the first part of the theorem. To prove the second part notice that in order for $\mathcal{K}_{n,m} = 0$, $\Phi_{n,m-1} \perp \tilde{\Phi}_{n-1,m}$ so these polynomials must have more orthogonalities than those following from their definitions and Corollary 4.20 with $p(z,w) = \phi^{*m}_{n,m}$ shows this is so. □

We can now prove an extension of the Fejér-Riesz theorem to two variables.

Theorem 4.22 *Let $q_{n,m}(\theta,\phi)$ be a strictly positive bivariate trigonometric polynomial of degree n in θ and m in ϕ. Then $q_{n,m}(\theta,\varphi) = |p(z,w)|^2$ where $p(z,w)$ with $z = e^{i\theta}$, $w = e^{i\varphi}$ is of degree n in z and m in w and $p(z,w) \neq 0$ for $|z| \le 1$, $|w| \le 1$ if and only if the coefficient $\mathcal{K}_{n,m}$ associated with the measure $\frac{d\theta\, d\varphi}{4\pi^2 q_{n,m}(\theta,\varphi)}$ on $[-\pi,\pi]^2$ satisfy*

$$\mathcal{K}_{n,m} = 0.$$

Proof If the measure μ in the previous theorem is taken as

$$d\mu = \frac{1}{4\pi^2}\frac{d\theta d\phi}{q_{n,m}(\theta,\phi)}$$

then all of the assertions of the theorem follow except for the equality $q_{n,m} = |\phi^{*m}_{n,m}|^2$. But this follows from the Cauchy–Schwarz inequality as in the one variable case (see Eq. (2.13)). □

Measures or the form $d\mu = \frac{1}{4\pi^2}\frac{d\theta d\phi}{q_{n,m}(\theta,\phi)}$ where $q_{n,m}$ is a trigonometric polynomial have come to be called Bernstein-Szegő measures see [15].

Note that there are recent extensions of the above theorems to cases where p is just stable for $|z| \le 1$ and $|w| = 1$ or vice versa. Again for these cases the representation and factorization theorems hold if two polynomial subspaces become orthogonal to each other. The exact statement of the results can be found in the references below.

Acknowledgements I would like to thank Plamen Iliev for suggestions and corrections to this paper. I would also like to thank an anonymous reviewer for very carefully reading and commenting on these lectures.

References

1. M. Bakonyi, H.J. Woerdeman, *Matrix Completions, Moments, and Sums of Hermitian Squares* (Princeton University Press, Princeton, 2011)
2. J.P. Burg, Maximum entropy spectral analysis. Ph.D. Dissertation, Stanford University, Stanford (1975)
3. D. Damanik, A. Pushnitski, B. Simon, The analytic theory of matrix orthogonal polynomials. Surv. Approx. Theory **4**, 1–85 (2008)
4. Ph. Delsarte, Y.V. Genin, Y.G. Kamp, Orthogonal polynomial matrices on the unit circle. IEEE Trans. Circuits Syst. **25**(3), 149–160 (1978)
5. Ph. Delsarte, Y.V. Genin, Y.G. Kamp, Planar least squares inverse polynomials. I. Algebraic properties. IEEE Trans. Circuits Syst. **26**(1), 59–66 (1979)
6. Ph. Delsarte, Y.V. Genin, Y.G. Kamp, A simple proof of Rudin's multivariable stability theorem. IEEE Trans. Acoust. Speech Signal Process. **28**(6), 701–705 (1980)
7. J.S. Geronimo, P. Iliev, Fejér-Riesz factorizations and the structure of bivariate polynomials orthogonal on the bicircle. J. Eur. Math. Soc. **16**(9), 1849–1880 (2014)
8. J.S. Geronimo, H.J. Woerdeman, Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables. Ann. Math. **160**(3), 839–906 (2004)
9. J.S. Geronimo, H.J. Woerdeman, Two variable orthogonal polynomials on the bicircle and structured matrices. SIAM J. Matrix Anal. Appl. **29**(3), 796–825 (2007)
10. J.S. Geronimo, P. Iliev, G. Knese, Orthogonality relations for bivariate Bernstein-Szegő measures, in *Recent Advances in Orthogonal Polynomials, Special Functions, and Their Applications*. Contemporary Mathematics, vol. 578 (American Mathematical Society, Providence, 2012), pp. 119–131
11. J.S. Geronimo, P. Iliev, G. Knese, Polynomials with no zeros on a face of the bidisk. J. Funct. Anal. **270**, 3505–3558 (2016)
12. Ya. L. Geronimus, *Orthogonal Polynomials* (Consultants Bureau, New York, 1961)

13. H. Helson, D. Lowdenslager, Prediction theory and Fourier series in several variables. Acta Math. **99**, 165–202 (1958); II, Acta Math. **106**, 175– 213 (1961)
14. T. Kailath, A. Vieira, M. Morf, Inverses of Toeplitz operators, innovations, and orthogonal polyomials. SIAM Rev. **20**, 106–119 (1978)
15. G. Knese, Bernstein-Szegő measures on the two dimensional torus. Indiana Univ. Math. J. **57**(3), 1353–1376 (2008)
16. G. Knese, Polynomials with no zeros on the bidisk. Anal. PDE **3**(2) 109–149 (2010)
17. H.J. Landau, Z. Landau, On the trigonometric moment problem in two dimensions. Indag. Math. (N.S.) **23** (4), 1118–1128 (2012)
18. B. Simon, Orthogonal polynomials on the unit circle, in *Part 1, American Mathematical Society Colloquium Publications*, vol. 54 (American Mathematical Society, Providence, 2005)
19. G. Szegő, Orthogonal Polynomials, in *American Mathematical Society Colloquium Publications*, vol. 23 (American Mathematical Society, Providence, 1939 (revised edition 1959))

Exceptional Orthogonal Polynomials and Rational Solutions to Painlevé Equations

David Gómez-Ullate and Robert Milson

Abstract These are the lecture notes for a course on exceptional polynomials taught at the *AIMS-Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications* that took place in Douala (Cameroon) from October 5–12, 2018. They summarize the basic results and construction of exceptional poynomials, developed over the past 10 years. In addition, some new results are presented on the construction of rational solutions to Painlevé equation P_{IV} and its higher order generalizations that belong to the $A_{2n}^{(1)}$-Painlevé hierarchy. The construction is based on dressing chains of Schrödinger operators with potentials that are rational extensions of the harmonic oscillator. Some of the material presented here (Sturm-Liouville operators, classical orthogonal polynomials, Darboux-Crum transformations, etc.) are classical and can be found in many textbooks, while some results (genus, interlacing and cyclic Maya diagrams) are new and presented for the first time in this set of lecture notes.

Keywords Sturm-Liouville problems · Classical polynomials · Darboux transformations · Exceptional polynomials · Painlevé equations · Rational solutions · Darboux dressing chains · Maya diagrams · Wronskian determinants

Mathematics Subject Classification (2000) Primary 33C45; Secondary 34M55

D. Gómez-Ullate (✉)
Escuela Superior de Ingeniería, Universidad de Cádiz, Puerto Real, Spain

Departamento de Física Teórica, Universidad Complutense de Madrid, Madrid, Spain
e-mail: david.gomezullate@uca.es

R. Milson
Department of Mathematics and Statistics, Dalhousie University, Halifax, NS, Canada
e-mail: rmilson@dal.ca

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_15

1 Introduction

The past 10 years have witnessed an intense activity of several research groups around the concept of exceptional orthogonal polynomials. Although some isolated examples in the physics literature existed before [11], the systematic study of exceptional polynomials started in 2009, with the publication of two papers [23, 24]. The original approach to the problem was via a complete classification of exceptional operators by increasing codimension, which proved to be computationally untractable (and moreover, much later it was shown that codimension is not a very well defined concept). The term *exceptional* was originally intended to evoque very rare, almost exotic cases, as for low codimension exceptional families are almost unique [24]. Yet, shortly after the publication of these results, Quesne showed [56] that the exceptional families in [23] could be obtained by Darboux transformations, and Odake and Sasaki showed the way to generalize these examples to arbitrary codimension [49, 50]. New families emerged later associated to multiple Darboux transformations [25, 28, 31, 32, 51], and nowadays it is clear that exceptional polynomials are certainly not rare, as we are starting to understand the whole theory behind their construction and classification.

The role of Darboux transformations in the construction of exceptional polynomial families is an essential ingredient. It was conjectured in [27] that every exceptional polynomial system can be obtained from a classical one via a sequence of Darboux transformations, which has been recently proved in [20]. Explaining these results lies beyond the scope of these lectures, and the interested reader is advised to read [20] for an updated account on the structure theorems underlying the theory of exceptional polynomials and operators. We will limit ourselves in the following pages to introduce the main ideas and constructions, as a sort of primer to the subject.

For those interested in gaining deeper knowledge in the properties of exceptional orthogonal polynomials, there are a large number of references in the bibliography section that cover the main results published in the past 10 years, by authors like Durán [12–16], Sasaki and Odake [49–53], Marquette and Quesne [38–40], Kuijlaars and Bonneux [7, 37], etc. that cover aspects like recurrence relations, symmetries, asymptotics, admissibility and regularity of the weights, properties of their zeros and electrostatic interpretation, and applications in solvable quantum mechanical models, among others.

The connection between sequences of Darboux transformations and Painlevé type equations has been known for more than 20 years, since the works of Adler [2], and Veselov and Shabat, [62]. However, the Russian school of integrable systems was more concerned with uncovering relations between different structures rather than providing complete classifications of solutions to Painlevé equations. The Japanese school pioneered by Sato developed a scheme to understand integrable equations as reductions from the KP equations. Noumi and Yamada [46], and their collaborators developed the geometric theory of Painlevé equations, by studying the group of Bäcklund transformations that map solutions to solutions (albeit for different values of the parameters). Using this transformations to *dress* some very simple seed solutions they managed to build and classify large classes of rational solutions to P_{IV} and P_V [9, 41, 42, 54, 60], and to extend this symmetry approach to higher order equations, that now bear their name. It was later realised that determinantal representations of these rational solutions exist [34, 35, 42] and that they involve classical orthogonal polynomial entries. For an updated account of the relation between orthogonal polynomials and Painlevé equations, the reader is advised to read the recent book by van Assche [61].

Our aim is to merge these two approaches: the strength of the Darboux dressing chain formulation with a convenient representation and indexing to describe the whole set of rational solutions to P_{IV} and its higher order generalizations belonging to the A_{2N}-Painlevé hierarchy. This is achieved by indexing iterated Darboux transformations with Maya diagrams, originally introduced by Sato, and exploring conditions that ensure cyclicity after an odd number of steps. We tackle this problem by introducing the concepts of genus and interlacing of Maya diagrams, which allow us to classify and describe cyclic Maya diagrams. For every such cycle, we show how to build a rational solution to the A_{2N}-Painlevé system, by a suitable choice of Wronskian determinants whose entries are Hermite polynomials. This approach generalizes the solutions for P_{IV} (A_2-Painlevé) known in the literature as Okamoto and generalized Hermite polynomials. We illustrate the construction by providing the complete set of rational solutions to A_4-Painlevé, the next system in the hierarchy.

2 Darboux Transformations

In this section we describe Darboux transformations on Schrödinger operators and their iterations at a purely formal level (i.e. with no interest on the spectral properties).

Let $L = -D_{xx} + U(x)$ be a Schrödinger operator, and $\varphi = \varphi(x)$ a formal eigenfunction of L with eigenvalue λ, so that

$$L[\varphi] = -\varphi'' + U\varphi = \lambda\varphi$$

Note that we are not assuming any condition of φ at this stage, we do not care at this formal level whether φ is square integrable or not. The function φ is usually called the *seed function* for the transformation, and λ the *factorization energy*. For every choice of φ and λ, we can factorize L in the following manner

$$L - \lambda = (D_x + w)(-D_x + w) = BA, \qquad w = (\log\varphi)'$$

where $B = D_x + w$ and $A = -D_x + w$ are first order differential operators. The Darboux transform of L, that we call $\tilde{L}$, is defined by commuting the two factors:

$$\tilde{L} - \lambda = AB = (-D_x + w)(D_x + w)$$

Expanding the two factors, we can find the relation between U and its transform $\tilde{U}$:

$$U = w^2 + w' - \lambda, \quad \tilde{U} = w^2 - w' - \lambda \Rightarrow \tilde{U} = U - 2w' \tag{2.1}$$

or in terms of the seed function we have

$$\tilde{U} = U - 2(\log \varphi)''$$

Note that $\ker A = \langle \varphi \rangle$, i.e. $A[\varphi] = -\varphi' + w\varphi = 0$,and also that $\ker B = \left\langle \frac{1}{\varphi} \right\rangle$, i.e. $B\left[\frac{1}{\varphi}\right] = 0$. The main reason to introduce this transformation is that we have the following intertwining relations between L and $\tilde{L}$:

$$LB = B\tilde{L}, \qquad AL = \tilde{L}A \tag{2.2}$$

These relations mean that we can connect the eigenfunctions of L and $\tilde{L}$.

Exercise 1 Show that if ψ is an eigenfunction of L with eigenvalue E, then $\tilde{\psi} = A[\psi]$ is an eigenfunction of $\tilde{L}$ with the same eigenvalue. Likewise, if $\tilde{\varphi}$ is an eigenfunction of $\tilde{L}$ with eigenvalue μ, then $\varphi = B[\tilde{\varphi}]$ is an eigenfunction of L with the same eigenvalue.

By hypothesis, we have that $L[\psi] = E\psi$. Let $\tilde{\psi} = A[\psi]$. We see that

$$\tilde{L}[\tilde{\psi}] = \tilde{L}A[\psi] = AL[\psi] = A[E\psi] = \mu A[\psi] = E\tilde{\psi}$$

The converse transformation is proved in a similar way.

Note that if we try to apply the Darboux transformation A on φ we do not get any eigenfunction of $\tilde{L}$, because $A[\varphi] = 0$. However, the reciprocal of φ is a new eigenfunction of $\tilde{L}$, with eigenvalue λ, as

$$B\left[\frac{1}{\varphi}\right] = 0 \Rightarrow \tilde{L}\left[\frac{1}{\varphi}\right] = (AB + \lambda)\left[\frac{1}{\varphi}\right] = \lambda\left(\frac{1}{\varphi}\right)$$

2.1 Exact Solvability by Polynomials

The above transformation is purely formal and its main purpose is to connect the eigenfunctions and eigenvalues of two different Schrödinger operators L and $\tilde{L}$. Now, this would be of very little purpose if we cannot say anything about the spectrum and eigenfunctions of at least one of the operators. Typically, we use

Darboux transformations to generate new solvable operators from ones that we know to be solvable. But what do we mean by *solvable*?

In general, *solvable* means that we can describe the spectrum and eigenfunctions in a more or less explicit form, and in terms of known functions. It is still not clear what a *known function* is... so we'd rather narrow down the definition and define exact solvability by polynomials in the following manner.

Definition 2.1 A Schrödinger operator

$$L = -D_{xx} + U(x) \tag{2.3}$$

is said to be *exactly solvable by polynomials* if there exist functions $\mu(x)$, $z(x)$ such that *for all but finitely many* $k \in \mathbb{N}$, L has eigenfunctions (in the L^2 sense) of the form

$$\psi_k(x) = \mu(x) y_k(z(x))$$

where $y_k(z)$ is degree k polynomial in z.

This definition captures many of the Schrödinger operators that we know to be exactly solvable: those in which the eigenfunctions have a common prefactor $\mu(x)$ times a polynomial in a suitable variable $z(x)$. The prefactor $\mu(x)$ is responsible for ensuring the right asymptotic behaviour at the endpoints for all bound states, while the polynomials y_k represent a modulation that describes the excited states.

From the purpose of orthogonal polynomials, this kind of Schrödinger operators are directly related to classical orthogonal polynomials, since polynomials y_k are automatically orthogonal if L is a self-adjoint operator, i.e. with appropriate regularity and boundary conditions.

More specifically, classical orthogonal polynomials are related to the following Schrödinger operators:

1. Hermite polynomials to the harmonic oscillator
2. Laguerre polynomials to the isotonic oscillator
3. Jacobi polynomials to the Darboux-Pöschl-Teller potential.

We would like to apply Darboux transformations to these three families of Schrödinger operators that are exactly solvable by polynomials, in order to generate new operators, but we would like these new operators to also be exactly solvable by polynomials. This means that we have to impose certain restrictions on the type of seed functions of L that we are free to choose for the Darboux transformations. In general, the class of *rational Darboux transformation* is the subset of all possible Darboux transformations that preserve exact solvability by polynomials. Fortunately, there is a simple way to characterize seed functions for this subclass, which we describe below. But before we do so, let us introduce some jargon between differential operators.

2.2 Schrödinger and Algebraic Operators

If we are dealing with Schrödinger operators that are exactly solvable by polynomials, there are two operators that we will work with: on one hand, we have the Schrödinger operator $L = -D_{xx} + U(x)$, on the other hand, we have the algebraic operator $T = p(z)D_{zz} + q(z)D_z + r(z)$ which is the one that has polynomial eigenfunctions

$$T[y_k] = p(z)y_k'' + q(z)y_k' + r(z)y_k = \lambda_k y_k.$$

There is a connection between these two operators, which is not entirely bidirectional in general, but it is bidirectional if L is exactly solvable by polynomials.

Proposition 2.2 *Every second order linear differential operator $T = p(z)D_{zz} + q(z)D_z + r(z)$ can be transformed into a Schrödinger operator $L = -D_{xx} + U(x)$ by the following change of variables and similarity transformation:*

$$x = -\int^z (-p)^{-1/2} \tag{2.4}$$

$$L = \mu \circ T \circ \left(\frac{1}{\mu}\right)\Big|_{z=z(x)}, \qquad \mu = \exp \int^z \frac{q - \frac{p'}{2}}{2p} \tag{2.5}$$

where $z(x)$ is defined by inverting (2.4)

Exercise 2 Prove that if $T = p(z)D_{zz} + q(z)D_z + r(z)$, then the operator L defined by (2.4) and (2.5) is a Schrödinger operator $L = -D_{xx} + U(x)$, and find an expression for the potential U in terms of p, q and r.

Let us denote

$$\mu(z) = \exp \int^z \sigma, \qquad \sigma = \frac{q - \frac{p'}{2}}{2p}.$$

This, by applying the product rule for composition of differential operators, we have

$$\mu \circ D_z \circ \left(\frac{1}{\mu}\right) = D_z + \left(\log \frac{1}{\mu}\right)' = D_z - \sigma$$

$$\mu \circ D_{zz} \circ \left(\frac{1}{\mu}\right) = D_{zz} + 2\left(\log \frac{1}{\mu}\right)' D_z + \mu \left(\frac{1}{\mu}\right)'' = D_{zz} - 2\sigma D_z - \sigma' + \sigma^2$$

So collecting terms we have

$$\mu \circ T \circ \left(\frac{1}{\mu}\right) = pD_{zz} + \frac{1}{2}p'D_z + r - q\sigma - p\sigma' + p\sigma^2$$

Finally, the chain rule for differentiation leads to

$$\frac{d}{dz} = \frac{dx}{dz}\frac{d}{dx} \qquad \Rightarrow \qquad D_z = -(-p)^{1/2}D_x$$

$$\frac{d^2}{dz^2} = \left(\frac{dx}{dz}\right)^2 \frac{d^2}{dx^2} + \frac{d^2x}{dz^2}\frac{d}{dx} \qquad \Rightarrow \qquad D_{zz} = -\frac{1}{p}D_{xx} + \frac{1}{2}(-p)^{1/2}p'D_x$$

which inserted into the previous equation becomes

$$\mu \circ T \circ \left(\frac{1}{\mu}\right) = -D_{xx} + r + p\sigma^2 - p\sigma' - q\sigma$$

which is a Schrödinger operator. We identify the potential $U(x)$ to be

$$U(x) = r + p\sigma^2 - p\sigma' - q\sigma\Big|_{z=z(x)}, \qquad \sigma = \frac{q - \frac{p'}{2}}{2p} \tag{2.6}$$

Note that we can always go from T to L, but in general there is no prescribed way to go from L to T. This means that given a Schrödinger operator with some potential $L = -D_{xx} + U(x)$ it is a difficult question to know whether we can perform a change of variables and conjugation by a factor μ as in (2.4) and (2.5) such that T has polynomial eigenfunctions $y_k(z)$ (this is sometimes known as *algebraizing* a Schrödinger operator). Otherwise speaking, given a potential U it is hard to know whether it is exactly solvable by polynomials.[1]

2.3 Rational Darboux Transformations

Now we start from a given L which we know to be exactly solvable by polynomials, and we would like to perform a Darboux transformation in such a way that $\tilde{L}$ is still exactly solvable by polynomials. What conditions must the seed function φ satisfy for this to hold?

This question is not easy to answer in general. Let us look at one example.

Example 2.3 Consider the harmonic oscillator $L = -D_{xx} + x^2$. One possible choice for seed functions for rational Darboux transformations comes from choosing

[1] Solving this question is equivalent to classifying exceptional polynomials and operators, a question that we shall mention below.

φ among the bound states of L, i.e.

$$\varphi_k = \mathrm{e}^{-x^2/2} H_k(x), \qquad k \geq 0$$

where $H_n(x)$ is the n-th Hermite polynomial. We see that L is exactly solvable by polynomials, with $z(x) = x$ and $\mu(x) = \mathrm{e}^{-x^2/2}$. This implies that $p(z) = 1$ from (2.4), and from (2.5) we see that $q(z) = 2z$, so that $T = D_{zz} - 2zD_z$

We thus have

$$\begin{aligned} L &= -D_{xx} + x^2, \qquad & L[\varphi_k] &= (2k+1)\varphi_k, \qquad & k &= 0, 1, 2, \dots \\ T &= D_{zz} - 2zD_z, \qquad & T[H_k] &= 2kH_k, \qquad & k &= 0, 1, 2, \dots \end{aligned}$$

But this is not the only possible choice [21, 22]. Note that the Schrödinger operator $L = -D_{xx} + x^2$ is not only invariant under the transformation $x \to -x$ but it only picks a sign when we perform the transformation $x \to \mathrm{i}x$, so there is another set of eigenfunctions

$$\tilde{\varphi}_k = \mathrm{e}^{x^2/2} \tilde{H}_k(x), \quad k \geq 1$$

where $\tilde{H}_k(x) = \mathrm{i}^{-k} H_k(\mathrm{i}x)$ is called the conjugate Hermite polynomial. Note that these eigenfunctions are obtained by exploiting a discrete symmetry of the equation, and their eigenvalues are negative:

$$L[\tilde{\varphi}_k] = -(2k+1)\tilde{\varphi}_k, \quad k = 1, 2, \dots$$

Because the pre-factor is now a positive gaussian, the functions blow up at $\pm\infty$ and they are not square integrable (in the physics literature they are sometimes called *virtual states*). But for the purposes of using them as seed functions for Darboux transformations, they are perfectly valid. These two sets of eigenfunctions exhaust all possible seed functions for rational Darboux transformations of the harmonic oscillator. Rather than two families of functions, each of them indexed by natural numbers, it will be useful to consider them as one single family indexed by integers:

$$\varphi_n(x) = \begin{cases} \mathrm{e}^{-x^2/2} H_n(x), & \text{if } n \geq 0 \\ \mathrm{e}^{x^2/2} \tilde{H}_{-n-1}(x), & \text{if } n < 0 \end{cases} \tag{2.7}$$

2.4 Iterated or Darboux-Crum Transformations

Now that we know how to apply Darboux transformations to pass from L to $\tilde{L}$ there is no reason why we should stop there... we can apply the transformation once again, and in general as many times as we want. Let's do it once more. Suppose that

φ_1 and φ_2 are two formal eigenfunctions of L:

$$L[\varphi_1] = \lambda_1\varphi_1, \qquad L[\varphi_2] = \lambda_2\varphi_2$$

We first use seed function φ_1 to transform L into $\tilde{L}$, so that

$$\tilde{U} = U - 2(\log\varphi_1)''$$

and the eigenfunctions are related by

$$\tilde{\psi} = A[\psi] = -\psi' + \frac{\varphi_1'}{\varphi_1}\psi = \frac{\mathrm{Wr}[\varphi_1, \psi]}{\varphi_1}.$$

But now we observe that $\tilde{\varphi}_2$ given by

$$\tilde{\varphi}_2 = A[\varphi_2] = \frac{\mathrm{Wr}[\varphi_1, \varphi_2]}{\varphi_1}$$

is a formal eigenfunction of $\tilde{L}$, so we can use it to Darboux transform $\tilde{L}$ into $\tilde{\tilde{L}}$, so

$$\tilde{\tilde{U}} = \tilde{U} - 2(\log\tilde{\varphi}_2)'' = U - 2(\log\mathrm{Wr}[\varphi_1, \varphi_2])''$$

If ψ are the eigenfunctions of L and $\tilde{\psi} = \mathrm{Wr}[\varphi_1, \psi]/\varphi_1$ are the eigenfunctions of $\tilde{L}$, the eigenfunctions of $\tilde{\tilde{L}}$ are given by

$$\tilde{\tilde{\psi}} = A[\tilde{\psi}] = \frac{\mathrm{Wr}[\tilde{\varphi}_2, \tilde{\psi}]}{\tilde{\varphi}_2} = \frac{\mathrm{Wr}\left[\varphi_1^{-1}\mathrm{Wr}[\varphi_1, \varphi_2], \varphi_1^{-1}\mathrm{Wr}[\varphi_1, \psi]\right]}{\varphi_1^{-1}\mathrm{Wr}[\varphi_1, \varphi_2]} = \frac{\mathrm{Wr}[\varphi_1, \varphi_2, \psi]}{\mathrm{Wr}[\varphi_1, \varphi_2]}$$

where we have used two identities satisfied by Wronskian determinants, namely

$$\mathrm{Wr}[gf_1, \dots, gf_n] = g^n\,\mathrm{Wr}[f_1, \dots, f_n]$$

and

$$\mathrm{Wr}[f_1, \dots, f_n, g, h] = \frac{\mathrm{Wr}\left[\mathrm{Wr}[f_1, \dots, f_n, g], \mathrm{Wr}[f_1, \dots, f_n, h]\right]}{\mathrm{Wr}[f_1, \dots, f_n]}.$$

It is not hard to iterate this argument and prove by induction the following result, known as Darboux-Crum formula.

Proposition 2.4 *Let $\varphi_1, \dots, \varphi_n$ be a set of n formal eigenfunctions of a Schrödinger operator L. We can perform an n-step Darboux transformation with these seed* *eigenfunctions, to obtain a chain of Schrödinger operators*

$$L = L_0 \to L_1 \to \cdots \to L_n.$$

The Schrödinger operator of L_n is given by

$$L_n = -D_{xx} + U_n = -D_{xx} + U - 2\,(\log \mathrm{Wr}[\varphi_1, \ldots, \varphi_n])''\,.$$

If ψ is a formal eigenfunction of L with eigenvalue E, then

$$\psi^{(n)} = \frac{\mathrm{Wr}[\varphi_1, \ldots, \varphi_n, \psi]}{\mathrm{Wr}[\varphi_1, \ldots \varphi_n]} \tag{2.8}$$

is a formal eigenfunction of L_n with the same eigenvalue.

Example 2.5 Coming back to the harmonic oscillator of Example 2.3, we saw that seed functions for rational Darboux transformations are in one-to-one correspondence with the integers. If we want to perform a multi-step Darboux transformation, we need to fix a multi-index that specifies the set of seed functions to be used. For instance, corresponding to the multi-index $(-3, -2, 1, 4)$ we would have, according to (2.7) the Darboux-Crum transformation acting on a function ψ would be

$$\psi^{(4)} = \frac{\mathrm{Wr}\left[\mathrm{e}^{x^2/2}\tilde{H}_2, \mathrm{e}^{x^2/2}\tilde{H}_1, \mathrm{e}^{-x^2/2}H_1, \mathrm{e}^{-x^2/2}H_4, \psi\right]}{\mathrm{Wr}\left[\mathrm{e}^{x^2/2}\tilde{H}_2, \mathrm{e}^{x^2/2}\tilde{H}_1, \mathrm{e}^{-x^2/2}H_1, \mathrm{e}^{-x^2/2}H_4\right]}$$

In the following sections we will see how the polynomial part of these functions essentially defines exceptional Hermite polynomials, and how these Wronskians enjoy very particular symmetry properties that admit an elegant combinatorial description in terms of Maya diagrams.

3 The Bochner Problem: Classical and Exceptional Polynomials

After having reviewed the notion of Darboux-Crum transformations, in this section we will introduce the concept of exceptional orthogonal polynomials, as orthogonal polynomial systems that arise from Sturm-Liouville problems with *exceptional degrees*, i.e. gaps in their degree sequence. But before we do so, we need to review some basic facts about Sturm-Liouville problems, and introduce Bochner's theorem, that characterizes the classical orthogonal polynomial systems of Hermite, Laguerre and Jacobi as polynomial eigenfunctions (with no missing degrees) of a Sturm-Liouville problem.

3.1 Sturm-Liouville Problems

A Sturm-Liouville problem is a second-order boundary value problem of the form

$$-(P(z)y')' + R(z)y = \lambda W(z)y, \quad y = y(z), \tag{3.1}$$

$$\begin{aligned} \alpha_0 y(a) + \alpha_1 y'(a) &= 0 \\ \beta_0 y(b) + \beta_1 y'(b) &= 0 \end{aligned} \tag{3.2}$$

where $I = (a, b)$ is an interval, where λ is a spectral parameter, where $P(z), W(z), R(z)$ are suitably smooth real-valued functions with $P(z), W(z) > 0$ for $z \in I$.[2]

Dividing (3.1) by $W(z)$ re-expresses the underlying differential equation in an operator form:

$$-T[y] = \lambda y, \tag{3.3}$$

where

$$T[y] = p(z)y'' + q(z)y' + r(z)y, \tag{3.4}$$

and where

$$\begin{aligned} p(z) &= \frac{P(z)}{W(z)} & P(z) &= \exp \int \frac{q(z)}{p(z)} dz \\ q(z) &= \frac{P'(z)}{W(z)} & W(z) &= \frac{P(z)}{p(z)}, \\ r(z) &= -\frac{R(z)}{W(z)} & R(z) &= -r(z)W(z) \end{aligned} \tag{3.5}$$

If $y_1(z), y_2(z)$ are two sufficiently smooth real-valued functions, then integration by parts gives Lagrange's identity:

$$\int (T[y_1]y_2 - T[y_2]y_1)(z)\, W(z)dz = P(z)(y_1'(z)y_2(z) - y_2'(z)y_1(z)). \tag{3.6}$$

[2]In the case of an unbounded interval with $a = -\infty$ and/or $b = +\infty$, or if solutions $y(z)$ of (3.1) have no defined value at the endpoints, one has to consider the asymptotics of the corresponding solutions and impose boundary conditions of a more general form:

$$\begin{aligned} \alpha_0(z)y(z) + \alpha_1(z)y'(z) &\to 0 \quad \text{as } z \to a^- \\ \beta_0(z)y(z) + \beta_1(z)y'(z) &\to 0 \quad \text{as } z \to b^+ \end{aligned}$$

where $\alpha_0(z), \alpha_1(z), \beta_0(z), \beta_1(z)$ are continuous functions defined on I.

Suppose that the boundary conditions entail (1) the square integrability of eigenfunctions with respect to $W(z)dz$ over the interval I; and (2) the vanishing of the right side of (3.6) at the endpoints of the interval. With some suitable regularity assumptions on $P(z)$, $W(z)$, $R(z)$ one can then show that the eigenvalues of $-T$ can be ordered so that $\lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots \to \infty$.

If $y_i, y_j,\ \ i \neq j$ are two eigenfunctions corresponding to eigenvalues λ_i, λ_j, respectively, then (3.6) reduces to

$$(\lambda_i - \lambda_j) \int_I y_i(z) y_j(z) W(z) dz = P(z)(y_i'(z) y_j(z) - y_j'(z) y_i(z))\Big|_{a^+}^{b^-} = 0. \qquad (3.7)$$

Therefore, the eigenfunctions are orthogonal with respect to the inner product

$$\langle f, g \rangle_W = \int_I f(z) g(z) W(z) dz.$$

Example 3.1 Let's work out the weight and boundary conditions for the Hermite differential equation

$$y'' - 2zy + \lambda y, \quad y = y(z). \qquad (3.8)$$

We apply (3.5) and rewrite the above in Sturm-Liouville form

$$-(W(z)y')' = \lambda W(z) y, \quad y \in \mathrm{L}^2(\mathbb{R}, W dz) \qquad (3.9)$$

where the weight has the form

$$W(z) = \exp\left(\int^z (-2z) \right) = e^{-z^2}$$

In this case, the boundary conditions are that $e^{-z^2} y(z)^2$ be integrable near $\pm\infty$.

A basis of solutions to (3.8) are

$$\phi_0(z; \lambda) = \Phi\left(-\frac{\lambda}{4}, \frac{1}{2}, z^2 \right) \qquad (3.10)$$

$$\phi_1(z; \lambda) = z\, \Phi\left(\frac{1}{2} - \frac{\lambda}{4}, \frac{3}{2}, z^2 \right) \qquad (3.11)$$

where

$$\Phi(a, c, x) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n n!} x^n,$$

is the confluent hypergeometric function. This function has the asymptotic behaviour

$$\Phi(a, c, x) = \frac{\Gamma(c)}{\Gamma(a)} e^x x^{a-c} \left(1 + O(|x|^{-1})\right), \quad x \to +\infty,$$

This implies that

$$e^{-z^2}\phi_0(z;\lambda)^2 = \frac{\pi e^{z^2} z^{-2-\lambda}}{\Gamma(-\lambda/4)^2}\left(1 + O(z^{-2})\right), \quad z \to \pm\infty,$$
$$e^{-z^2}\phi_1(z;\lambda)^2 = \frac{\pi e^{z^2} z^{-2-\lambda}}{4\Gamma(1/2-\lambda/4)^2}\left(1 + O(z^{-2})\right), \quad z \to \pm\infty.$$

are not integrable for generic values of λ near $z = \pm\infty$. We now introduce two other solutions of (3.8),

$$\psi_R(z;\lambda) = \Psi\left(-\frac{\lambda}{4}, \frac{1}{2}, z^2\right), \quad z > 0 \tag{3.12}$$

$$\psi_L(z;\lambda) = \Psi\left(-\frac{\lambda}{4}, \frac{1}{2}, z^2\right), \quad z < 0 \tag{3.13}$$

where

$$\Psi(a, c; x) = \frac{\Gamma(1-c)}{\Gamma(a-c+1)}\Phi(a, c; x) + \frac{\Gamma(c-1)}{\Gamma(a)}\Phi(a-c+1, 2-c; x), \quad x > 0. \tag{3.14}$$

Note that $\psi_R(z)$ and $\psi_L(z)$ are different functions, because Ψ is a branch of a multi-valued function defined by taking a branch cut over the negative real axis. However, ψ_L, ψ_R may be continued to solutions of (3.8) over all of $\mathbb{R}$ by means of connection formulae (3.15), below.

We have the asymptotics

$$x^a\Psi(a, c; x) = 1 + O(x^{-1}), \quad x \to +\infty$$
$$e^{-z^2}\psi_R(z;\lambda)^2 = e^{-z^2} z^{\lambda}\left(1 + O(z^{-2})\right), \quad z \to +\infty$$
$$e^{-z^2}\psi_L(z;\lambda)^2 = e^{-z^2} z^{\lambda}\left(1 + O(z^{-2})\right), \quad z \to -\infty$$

Hence, ψ_R, ψ_L each satisfy a one-sided boundary conditions at $\pm\infty$.

From (3.14) we get the connection formulae

$$\begin{aligned}\psi_R(z;\lambda) &= \frac{\sqrt{\pi}}{\Gamma(1/2-\lambda/4)}\phi_0(z;\lambda) - \frac{2\sqrt{\pi}}{\Gamma(-\lambda/4)}\phi_1(z;\lambda),\\ \psi_L(z;\lambda) &= \frac{\sqrt{\pi}}{\Gamma(1/2-\lambda/4)}\phi_0(z;\lambda) + \frac{2\sqrt{\pi}}{\Gamma(-\lambda/4)}\phi_1(z;\lambda).\end{aligned} \tag{3.15}$$

Therefore, our boundary conditions amount to imposing the condition that ψ_L be proportional to ψ_R. By inspection of (3.15), this can happen in exactly two ways: $\psi_L = \psi_R$ and $\psi_L = -\psi_R$. The first case occurs when $\Gamma(-\lambda/4) \to \infty$ that is when $\lambda/2 = 2n,\ n = 1, 2, \ldots$. The second possibility occurs when $\Gamma(1/2 - \lambda/4) \to \infty$ which occurs when $\lambda/2 = 2n + 1,\ n = 1, 2, \ldots$. In the first case, we recover the even Hermite polynomials; in the second the odd Hermite polynomials. This last observation can be restated as the following identity

$$2^{-n} H_n(z) = \sqrt{\pi}\left(\frac{\phi_0(z; 2n)}{\Gamma(1/2 - n/2)} - \frac{2\phi_1(z; 2n)}{\Gamma(-n/2)}\right), \quad n = 0, 1, 2, \ldots.$$

Therefore the Hermite polynomials are precisely the solutions of (3.8) that satisfy the boundary conditions of (3.9), namely they are the only solutions of (3.8) that are square-integrable with respect to e^{-z^2} over all of $\mathbb{R}$.

3.2 Classical Orthogonal Polynomials

The notion of a Sturm-Liouville system with polynomial eigenfunctions is the cornerstone idea in the theory of classical orthogonal polynomials. The reason is simple: if the eigenfunctions of a Sturm-Liouville problem (3.1) are polynomials, then they will be orthogonal with respect to the corresponding weight $W(z)$.

The following three types of polynomials—bearing the names of Hermite, Laguerre, and Jacobi—are known as the classical orthogonal polynomials.

- Hermite polynomials obey the following 3-term recurrence relation:

$$2zH_n = H_{n+1} + 2nH_{n-1}, \qquad H_{-1} = 0,\ H_0 = 1. \tag{3.16}$$

They are orthogonal with respect to

$$W_{\mathrm{H}}(z) = e^{-z^2}, \quad z \in (-\infty, \infty),$$

and satisfy the following differential equation

$$y'' - 2zy' + 2ny = 0, \quad y = H_n(z), \quad n \in \mathbb{N}, \tag{3.17}$$

- Laguerre polynomials $L_n = L_n^{(\alpha)}(z)$ have one parameter α, and satisfy the following 3-term recurrence relation:

$$2zL_n = (n + 1)L_{n+1} - (2n + \alpha + 1)L_n + (n + \alpha)L_{n-1}, \qquad L_{-1} = 0, L_0 = 1. \tag{3.18}$$

For $\alpha > -1$, Laguerre polynomials are orthogonal with respect to

$$W_{\mathrm{L}} = e^{-z} z^{\alpha}, \quad z \in (0, \infty).$$

They satisfy the following differential equation

$$zy'' + (\alpha + 1 - z)y' + ny = 0, \quad y = L_n^{(\alpha)}(z), \quad n \in \mathbb{N}, \tag{3.19}$$

- Jacobi polynomials $P_n = P_n^{(\alpha,\beta)}(z)$ have two parameters, α, β and are defined by:

$$\begin{aligned} zP_n =& \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} P_{n+1} \\ &+ \frac{\left(\beta^2 - \alpha^2\right)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} P_n \\ &+ \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} P_{n-1}, \qquad P_{-1} = 0, \; P_0 = 1 \end{aligned} \tag{3.20}$$

These polynomials obey the differential equation

$$(1 - z^2)P_n'' + (\beta - \alpha - z(\alpha + \beta + 2))P_n' + n(\alpha + \beta + n + 1)P_n = 0 \tag{3.21}$$

For $\alpha, \beta > -1$ they are orthogonal with respect to

$$W_{\mathrm{H}} = (1 - z)^{\alpha}(1 + z)^{\beta}, \quad z \in (-1, 1).$$

Exercise 3 Rewrite the above differential equations in Sturm-Liouville form. In each case, work out the boundary conditions that pick out the polynomial solutions.

The class of Sturm-Liouville problems with polynomial eigenfunctions was studied and classified by Solomon Bochner in the following fundamental result. Bochner's Theorem was subsequently refined by Lesky to show that the three classical families of Hermite, Laguerre, and Jacobi give a full classification of such Sturm-Liouville problem.

Theorem 3.2 (Bochner) *Suppose that an operator*

$$T[y] = p(z)y'' + q(z)y' + r(z)y \tag{3.22}$$

admits eigenpolynomials of every degree; that is, there exist polynomials $y_k(z)$ with $\deg y_k = k$ *and constants λ_k such that*

$$-T[y_k] = \lambda_k y_k, \quad k = 0, 1, 2, \dots. \tag{3.23}$$

Then, necessarily p, q, r are polynomials with

$$\deg p \le 2, \quad \deg q \le 1, \quad \deg r = 0.$$

Moreover, if these polynomials are the orthogonal eigenfunctions of a Sturm-Liouville system, then up to an affine transformation of the independent variable z, they are the classical polynomials of Hermite, Laguerre, and Jacobi.

Proof Applying (3.23) to $k = 0, 1, 2$, we obtain

$$\begin{aligned} -\lambda_0 y_0 &= r \\ -\lambda_1 y_1 &= q y_1' + r y_1 \\ -\lambda_2 y_2 &= p y_2'' + q y_2' + r y_2. \end{aligned}$$

By inspection, r is a constant, while q, p are polynomials with $\deg q \le 1$ and $\deg p \le 2$.

Up to an affine transformation $z \mapsto sz + t$, the leading coefficient $p(z)$ can assume one of the following normal forms:

$$1, z, z^2, 1 - z^2, 1 + z^2.$$

Write $q(z) = az + b$. Applying (3.5), the corresponding weights have the form

$$\begin{aligned} &\text{(i)} && W(z) = e^{\frac{b^2}{2a}} e^{\frac{a}{2}(z+b/a)^2} \\ &\text{(ii)} && W(z) = e^{az} z^{b-1} \\ &\text{(iii)} && W(z) = e^{-\frac{b}{z}} z^{a-2}, \\ &\text{(iv)} && W(z) = (1-z)^{-(a+b)/2-1}(1+z)^{(b-a)/2-1} \\ &\text{(v)} && W(z) = e^{b \arctan(z)}(1+z^2)^{a/2-1}. \end{aligned}$$

- For normal form (i), the case $a = 0$ is excluded. If not, the resulting operator would be strictly degree lowering, which would preclude the existence eigenpolynomials of degrees ≥ 2. Since $p(z) = 1$ is invariant with respect to scaling and translation, no generality is lost by setting $b = 0$, $a = \pm 2$. The case of $a = -2$ corresponds to the classical Hermite polynomials. The case $a = 2$ can be excluded because there is no choice of boundary conditions that result in the vanishing of the right side of (3.7).
- For the normal form (ii), note that $p(z) = z$ is preserved by scaling transformations. Hence, without loss of generality we can take $a = -1$. This case corresponds to the classical Laguerre polynomials.
- Normal form (iii) is a bit tricky. The case $b = 0$ can be ruled out because of the absence of suitable boundary conditions. The analysis of $b < 0$ and $b > 0$ is the

same, so suppose that $b > 0$. Here the only possible boundary conditions are at the endpoints of the interval $(0, \infty)$. If $a < 0$ then a finite number of polynomials can be made to be square integrable with respect to the weight in question. These constitute the so-called Bessel orthogonal polynomials, which however fall outside the range of our definition—we require that *all* y_k are square-integrable with respect to $W(z)$.

- Normal form (iv) corresponds to the Jacobi orthogonal polynomials.
- Normal form (v) corresponds to the so-called twisted Jacobi (also called Romanovsky) polynomials. If $a < 0$ then a finite number of initial degrees are square-integrable with respect to the indicated weight over the interval $(-\infty, \infty)$. As above, this violates our requirement that *all* the y_k be square-integrable with respect to $W(z)dz$. □

3.3 Exceptional Polynomials and Operators

We now modify the assumption of Bochner's Theorem 3.2 to arrive at the following.

Definition 3.3 We say that $T[y] = p(z)y'' + q(z)y' + r(z)y$ is an *exceptional operator* if it admits polynomial eigenfunctions for a cofinite number of degrees; that is, there exist polynomials $y_k(z),\ \ k \notin \mathbb{N} \setminus \{d_1, \ldots, d_m\}$ with $\deg y_k = k$ and with $d_1, \ldots, d_m \in \mathbb{N}$ a finite number of exceptional degrees, and constants λ_k such that

$$-T[y_k] = \lambda_k y_k, \quad k \in \mathbb{N} \setminus \{d_1, \ldots, d_m\}.$$

Moreover, if it is possible to impose boundary conditions so that the polynomials y_k become eigenfunctions of the corresponding Sturm-Liouville problem, then we call the $\{y_k\}_{k \notin \{d_1,\ldots,d_m\}}$ *exceptional orthogonal polynomials*.

The relaxed assumption that permits for a finite number of missing degrees allows to escape the constraints of Bochner's theorem and characterizes a large and interesting new class of operators and polynomials.

Example 3.4 We next show an example of codimension 2 exceptional Hermite polynomials. Recall the classical Hermite polynomials defined by (3.16). Introduce a family of exceptional Hermite polynomials defined by

$$\widehat{H}_n = \frac{\mathrm{Wr}[H_1, H_2, H_n]}{8(n-1)(n-2)} = H_n + 4nH_{n-2} + 4n(n-3)H_{n-4}, \quad n \neq 1, 2 \tag{3.24}$$

where the $H_i(z)$ are classical Hermite polynomials and where Wr denotes the usual Wronskian determinant:

$$\mathrm{Wr}[H_1, H_2, H_n] = \begin{vmatrix} H_1 & H_1' & H_1'' \\ H_2 & H_2' & H_2'' \\ H_n & H_n' & H_n'' \end{vmatrix}.$$

Exercise 4 Using the following identity for the classical Hermite polynomials:

$$H_n' = 2nH_{n-1}$$

and the 3-term recurrence relation (3.16) reduce the Wronskian expression

$$\widehat{H}_n = \frac{\mathrm{Wr}[H_1, H_2, H_n]}{8(n-1)(n-2)}, \quad n \neq 1, 2$$

to the right-hand side expression shown in (5.7).

Observe that $\deg \widehat{H}_n = n$. We call the resulting sequence of polynomials *exceptional* because the degree sequence $\deg \widehat{H}_n$ is missing two the degrees—the exceptional degrees $n = 1$ and $n = 2$. We call the $\widehat{H}_n(z)$ *exceptional Hermite* polynomials because they furnish polynomial solutions of the following modified version of the Hermite differential equation:

$$y'' - \left(2z + \frac{8z}{1+2z^2}\right) y' + 2ny = 0, \quad y = \widehat{H}_n(z),\ n \neq 1, 2. \tag{3.25}$$

At first glance, the exceptional modification of Hermite's differential equation (3.25) has a rather peculiar form; indeed it is slightly paradoxical that a differential equation with rational coefficients admits polynomial solutions. However, some of the underlying structure of the equation comes to light once we "clear denominators" and re-express (3.25) using the following, bilinear formulation:

$$\left(\eta y'' - 2\eta' y' + \eta'' y\right) - 2z\left(\eta y' - \eta' y\right) + 2(n-2)\,\eta y = 0 \tag{3.26}$$

where

$$\eta = \mathrm{Wr}[H_1, H_2] = 4 + 8z^2$$

Now the equation is bilinear in η, which is fixed and $y = y(z)$ the dependent variable, and nearly symmetric with respect to the two variables.

We can also rewrite expression (3.25) using Sturm-Liouville form, as

$$\left(\widehat{W} y'\right)' = \lambda \widehat{W} y, \tag{3.27}$$

where

$$\widehat{W}(z) = \frac{e^{-z^2}}{\eta(z)^2}, \quad \lambda = -2n.$$

Table 1 Degrees in the exceptional recurrence relations

n	RHS degrees in relation (3.28)	RHS degrees in relation (3.29)
0	3	4,0
3	6,4,0	7,5,3
4	7,5,3	8,6,4,0
5	8,6,4	9,7,5,3
6	9,7,5,3	10,8,6,4
7	10,8,6,4	11,9,7,5,3
$\vdots$	$\vdots$	$\vdots$
$n \geq 7$	$n+3, n+1, n-1, n-3$	$n+4, n+2, n, n-2, n-4$

As before, the Sturm-Liouville form implies the orthogonality of the eigenpolynomials:

$$\int_{-\infty}^{\infty} \widehat{H}_m(z)\widehat{H}_n(z)\widehat{W}(z)dz = 0, \quad m \neq n,\ m, n \neq 1, 2$$

It is also possible to show that the exceptional polynomials satisfy recurrence relations. However, now there are multiple relations of higher order:

$$4z(3+2z^2)\widehat{H}_n = \widehat{H}_{n+3} + 6n\,\widehat{H}_{n+1} + 12n(n-3)\,\widehat{H}_{n-1} + 8n(n-4)(n-5)\,\widehat{H}_{n-3}, \tag{3.28}$$

$$\begin{aligned} 16z^2(1+z^2)\widehat{H}_n = \widehat{H}_{n+4} + 8n\widehat{H}_{n+2} + 4(6n^2 - 14n + 1)\widehat{H}_n + \\ + 32n(n-3)(n-4)\widehat{H}_{n-2} + 16n(n-3)(n-5)(n-6)\widehat{H}_{n-4} \end{aligned} \tag{3.29}$$

Table 1 lists the degrees of the exceptional polynomials involved in the above recurrence at the values of $n = 0, 3, 4, \ldots$. By inspection, $\widehat{H}_0$ determines $\widehat{H}_3, \widehat{H}_4, \widehat{H}_6, \ldots, \widehat{H}_{2k}$, $k \geq 2$. Relation (3.28) with $n = 5$ then determines $\widehat{H}_5$. After that the $\widehat{H}_{2k+1}$, $k \geq 3$ are established. Observe that $\widehat{H}_n(z)$, $n \geq 7$ are determined by both relations (3.28) and (3.29). Remarkably, the relations are coherent, in the sense that both relations give the same value of $\widehat{H}_n(z)$, $n \geq 7$. This may be explained by the fact that the finite-order difference operators that describe the RHS of (3.28) and (3.29) commute with one another.

Exercise 5 Verify the recurrence relations (3.28) and (3.29) using a computer algebra system and show that the finite-difference operators that define their right-hand sides commute.

Finally, many of the properties of exceptional polynomials are explained by the fact that there is a hidden relation between them and their classical counterparts. Let us define second order operators

$$T[y] = y'' - 2zy',$$
$$\hat{T}[y] = y'' - \left(2z + \frac{8z}{1+2z^2}\right) y'$$

and re-express the classical and exceptional Hermite differential equations in operator form, respectively, as

$$-T[H_n] = 2nH_n,\ n \in \mathbb{N} \quad -\hat{T}[\widehat{H}_n] = 2n\widehat{H}_n,\ n \neq 1, 2.$$

Let us also introduce the second order operator

$$A[y] = \mathrm{Wr}[H_1, H_2, y] = 4(1+2z^2)y'' - 16zy' + 16y.$$

Exercise 6 Verify that the three differential operators T, $\hat{T}$ and A satisfy the following (second-order) intertwining relation:

$$\hat{T}A = AT. \tag{3.30}$$

Note that in the intertwining relations (2.2), operator A is first order, corresponding to a single-step Darboux transformation. In this case, A is a second-order differential operator that comes from a two-step Darboux-Crum transformation with seed functions H_1 and H_2. In general, up to a normalization constant, the exceptional polynomials are given by applying the intertwiner A to the classical polynomials:

$$\widehat{H}_n \propto A[H_n].$$

If we take the intertwining relation as proven, we obtain that

$$\hat{T}[A[H_n]] = (\hat{T}A)[H_n] = (AT)[H_n] = -2nA[H_n].$$

Thus, the intertwining relation "explains" why the $\widehat{H}_n$ are eigenpolynomials of the exceptional operator $\hat{T}$. This is essentially the same argument as the one used in Exercise 1, albeit with a higher-order intertwiner A.

4 Symmetric Painlevé Equations and Darboux Dressing Chains

And now for something completely different [45], or maybe not? The set of six nonlinear second order Painlevé equations $\mathrm{P_I}, \ldots, \mathrm{P_{VI}}$ have attracted considerable interest in the past 100 years [8, 33, 60]. They have the defining property that their solutions have no movable branch points. The Painlevé equations, whose solutions are called Painlevé transcendents, are now considered to be the nonlinear analogues of special functions, cf. [8]. These functions, in general, are transcendental in the sense that they cannot be expressed in terms of previously known functions. However, the Painlevé equations, except P_I, also possess special families of solutions that can be expressed via rational functions, algebraic functions or the classical special functions, such as Airy, Bessel, parabolic cylinder, Whittaker or hypergeometric functions, for special values of the parameters.

However, rather than studying the Painlevé second order scalar equations, we will follow Noumi and Yamada since it will prove to be more useful to rewrite these equations as a system of first order equations, which will allow us not only to understand the symmetry properties better, but also to generalize these system to higher order equations with the same desired properties.

Definition 4.1 We define the A_2-Painlevé system as the following system of three coupled nonlinear ODEs

$$\begin{aligned} f_0' + f_0(f_1 - f_2) &= \alpha_0, \\ f_1' + f_1(f_2 - f_0) &= \alpha_1, \\ f_2' + f_2(f_0 - f_1) &= \alpha_2, \end{aligned} \tag{4.1}$$

subject to the condition

$$(f_0 + f_1 + f_2)' = \alpha_0 + \alpha_1 + \alpha_2 = 1. \tag{4.2}$$

where $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{C}$ are complex parameters and $f_i = f_i(z)$ are complex functions.

If the parameters take on arbitrary values, the general solution of this equation is transcendental. We are interested in this lecture to find solutions to (4.1) where the functions $f_i = f_i(z)$ are rational functions of z. A solution of (4.1) will be a tuple of the form $(f_0, f_1, f_2|\alpha_0, \alpha_1, \alpha_2)$.

The reason why this system is relevant is that by eliminating two of the functions, we can reduce system (4.1) to a single second order nonlinear ODE, that we will call P_{IV} because it is the fourth equation in the list of six Painlevé equations, namely:

$$y'' = \frac{1}{2y}(y')^2 + \frac{3}{2}y^3 + 4ty^2 + 2(t^2 - a)y + \frac{b}{y} \tag{4.3}$$

Exercise 7 Show that if the tuple $(f_0, f_1, f_2|\alpha_0, \alpha_1, \alpha_2)$ is a solution to (4.1), then $y = y(t)$ is a solution to (4.3), where:

$$f_0 = -cy, \quad z = -\frac{t}{c}, \quad c = \sqrt{\frac{-1}{2}}, \quad a = 2(\alpha_1 - \alpha_2), \quad b = -2\alpha_0^2 \tag{4.4}$$

We first take the derivative of the first equation in (4.1):

$$f_0'' + f_0'(f_1 - f_2) + f_0(f_1' - f_2') = 0 \tag{4.5}$$

Next subtract the third from the second equation to obtain

$$f_1' - f_2' = \alpha_1 - \alpha_2 - 2f_1f_2 + f_0(f_1 + f_2)$$

and insert it into the previous equation, to get

$$f_0'' + f_0'(f_1 - f_2) + (\alpha_1 - \alpha_2)f_0 - 2f_1f_2f_0 + (f_1 + f_2)f_0^2 = 0. \tag{4.6}$$

From the first equation in (4.1) and the normalization $f_0 + f_1 + f_2 = z$, we have

$$f_1 - f_2 = \frac{\alpha_0 - f_0'}{f_0} \tag{4.7}$$

$$f_1 + f_2 = z - f_0 \tag{4.8}$$

Now bearing in mind that $4f_1f_2 = (f_1 + f_2)^2 - (f_1 - f_2)^2$ we have also

$$4f_1f_2 = (z - f_0)^2 - \left(\frac{\alpha_0 - f_0'}{f_0}\right)^2 \tag{4.9}$$

Inserting (4.7), (4.8) and (4.9) into (4.6), and after some cancellations and grouping terms we arrive at

$$f_0'' = \frac{f_0'^2}{2f_0} + \frac{3}{2}f_0^3 - 2zf_0^2 + \left(\frac{z^2}{2} + \alpha_2 - \alpha_1\right)f_0 - \frac{\alpha_0^2}{2f_0} \tag{4.10}$$

which after the rescaling of variable, function and parameters shown in (4.4) leads finally to (4.3).

Now that we know the equivalence between solutions of (4.1), that we will call $\mathrm{sP_{IV}}$, the symmetric form of $\mathrm{P_{IV}}$, it will be easier to work with the system than with the equation. In particular, Noumi and Yamada showed [47] that system (4.1) in invariant under a symmetry group, which acts by Bäcklund transformations on a tuple of functions and parameters. This symmetry group is the affine Weyl group $A_2^{(1)}$, generated by the operators $\{\boldsymbol{\pi}, \mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2\}$ whose action on the tuple $(f_0, f_1, f_2|\alpha_0, \alpha_1, \alpha_2)$ is given by:

$$\begin{aligned}
\mathbf{s}_k(f_j) &= f_j - \frac{\alpha_k\delta_{k+1,j}}{f_k} + \frac{\alpha_k\delta_{k-1,j}}{f_k}, \\
\mathbf{s}_k(\alpha_j) &= \alpha_j - 2\alpha_j\delta_{k,j} + \alpha_k(\delta_{k+1,j} + \delta_{k-1,j}), \\
\boldsymbol{\pi}(f_j) &= f_{j+1}, \qquad \boldsymbol{\pi}(\alpha_j) = \alpha_{j+1}
\end{aligned} \tag{4.11}$$

where $\delta_{k,j}$ is the Kronecker delta and $j, k = 0, 1, 2 \mod (3)$.

The technique to generate rational solutions is to first identify a number of very simple rational *seed solutions*, and then successively apply the Bäcklund transformations (4.11) to generate families of rational solutions.

Exercise 8 Check that the tuple $(z, 0, 0|1, 0, 0)$ satisfies (4.1). This is one possible *seed solution*. Now use the Bäcklund transformations $\mathbf{s}_0$ and $\mathbf{s}_1\mathbf{s}_0$ to generate two solution tuples, and check explicitly that the obtained solutions solves (4.1)

It is obvious that $(z, 0, 0|1, 0, 0)$ satisfies (4.1). From (4.11), the action of $\mathbf{s}_0$ on the generic tuple $(f_0, f_1, f_2|\alpha_0, \alpha_1, \alpha_2)$ is given by

$$\mathbf{s}_0(f_0) = f_0, \qquad \mathbf{s}_0(\alpha_0) = -\alpha_0 \tag{4.12}$$

$$\mathbf{s}_0(f_1) = f_1 - \frac{\alpha_0}{f_0}, \qquad \mathbf{s}_0(\alpha_1) = \alpha_1 + \alpha_0 \tag{4.13}$$

$$\mathbf{s}_0(f_2) = f_2 + \frac{\alpha_0}{f_0}, \qquad \mathbf{s}_0(\alpha_0) = \alpha_2 + \alpha_0 \tag{4.14}$$

So we have then that $\mathbf{s}_0(z, 0, 0|1, 0, 0) = \left(z, \frac{-1}{z}, \frac{1}{z}| - 1, 1, 1\right)$, and we can readily verify that this tuple satisfies (4.1). In a similar manner, we see that

$$\mathbf{s}_1\mathbf{s}_0(z, 0, 0|1, 0, 0) = \mathbf{s}_1\left(z, \frac{-1}{z}, \frac{1}{z}\Big| - 1, 1, 1\right) = \left(0, -\frac{1}{z}, z + \frac{1}{z}\,\Big|\, 0, -1, 2\right)$$

which is also seen to satisfy (4.1).

In this way we can iteratively apply Bäcklund transformations on a small set of seed solutions and generate many rational solutions to (4.1). This is a beautiful approach, pioneered by the Japanese school, and the transformations (4.11) have a nice geometric interpretation in terms of reflection groups acting on the space of parameters $(\alpha_0, \alpha_1, \alpha_2)$. Note however that the solutions obtained by dressing a given seed solution are hard to write in closed form, and in general the whole procedure is more an algorithm to generate solutions than an explicit enumeration of them. If we ask ourselves how many poles the rational solution $\mathbf{s}_1^6\,\mathbf{s}_0^3(z, 0, 0|1, 0, 0)$ has, this might be a difficult question to answer with this representation.

For this reason, we will not pursue this approach henceforth in these notes, and we refer the interested reader to Noumi's book [46] to learn the geometric theory of Painlevé equations, and their connections with other topics in integrable systems (τ-functions, Hirota bilinear equations, Jacobi-Trudi formulas, reductions from KP equation, etc.).

We will concentrate in these lectures on alternative representations of the rational solutions, most notably the determinantal representations [34, 35].

Once we are aware of the symmetry structure of (4.1), the system admits a natural generalization to any number of equations, known as the $A_N^{(1)}$-Painlevé or the Noumi-Yamada system. The even case ($N = 2n$) is considerably simpler (for reasons that will be explained later), and it is the one we will focus on this notes.

Definition 4.2 We define the $A_{2n}^{(1)}$-Painlevé system (or Noumi-Yamada system) as the following system of $2n+1$ coupled nonlinear ODEs

$$f_i' + f_i\left(\sum_{j=1}^{n} f_{i+2j-1} - \sum_{j=1}^{n} f_{i+2j}\right) = \alpha_i, \qquad i = 0, \dots, 2n \mod (2n+1) \tag{4.15}$$

subject to the normalization condition

$$(f_0 + \cdots + f_{2n})' = \alpha_0 + \cdots + \alpha_{2n} = 1. \tag{4.16}$$

The symmetry group of this higher order system is the affine Weyl group $A_{2n}^{(1)}$, acting by Bäcklund transformations as in (4.11). The system has the Painlevé property, and thus can be considered a proper higher order generalization of sP_{IV} (4.1), which corresponds to $n = 1$.

The goal of this lecture is to develop a systematic procedure to describe rational solutions to system (4.11), providing an explicit representation of the solutions in terms of Wronskian determinants whose entries are Hermite polynomials. This is an alternative approach to the dressing of seed solutions by Bäcklund transformations described above.

4.1 Darboux Dressing Chains

The theory of dressing chains, or sequences of Schrödinger operators connected by Darboux transformations was developed by Adler [2], and Veselov and Shabat [62]. The connection between dressing chains and Painlevé equations was already shown in [2] and it has been exploited by some authors [5, 6, 38–40, 43, 57–59, 63]. This section follows mostly the early works of Adler, Veselov and Shabat.

Consider the following sequence of Schrödinger operators

$$L_i = -D_z^2 + U_i, \qquad D_z = \frac{d}{dz}, \quad U_i = U_i(z), \quad i \in \mathbb{Z} \tag{4.17}$$

where each operator is related to the next by a Darboux transformation, i.e. by the following factorization

$$\begin{aligned} L_i &= (D_z + w_i)(-D_z + w_i) + \lambda_i, \quad w_i = w_i(z), \\ L_{i+1} &= (-D_z + w_i)(D_z + w_i) + \lambda_i. \end{aligned} \tag{4.18}$$

It follows that the functions w_i satisfy the Riccati equations

$$w_i' + w_i^2 = U_i - \lambda_i, \quad -w_i' + w_i^2 = U_{i+1} - \lambda_i. \tag{4.19}$$

Equivalently, w_i are the log-derivatives of ψ_i, the seed function of the Darboux transformation that maps Ł$_i$ to Ł$_{i+1}$

$$L_i\psi_i = \lambda_i\psi_i, \qquad \text{where } w_i = \frac{\psi_i'}{\psi_i}. \tag{4.20}$$

Using (4.17) and (4.18), the potentials of the dressing chain are related by

$$U_{i+1} = U_i - 2w_i', \tag{4.21}$$

$$U_{i+n} = U_i - 2\left(w_i' + \cdots + w_{i+n-1}'\right), \quad n \geq 2 \tag{4.22}$$

If we eliminate the potentials in (4.19) and set

$$a_i = \lambda_i - \lambda_{i+1} \tag{4.23}$$

the following chain of coupled equations is obtained

$$(w_i + w_{i+1})' + w_{i+1}^2 - w_i^2 = a_i, \quad i \in \mathbb{Z}$$

Before continuing, note that this infinite chain of equations has the evident reversal symmetry

$$w_i \mapsto -w_{-i}, \qquad a_i \mapsto -a_{-i}. \tag{4.24}$$

This infinite chain of equations closes and becomes a finite dimensional system of ODEs if a cyclic condition is imposed on the potentials of the chain

$$U_{i+p} = U_i + \Delta, \quad i \in \mathbb{Z} \tag{4.25}$$

for some $p \in \mathbb{N}$ and $\Delta \in \mathbb{C}$. If this holds, then necessarily $w_{i+p} = w_i$, $a_{i+p} = a_i$, and

$$\Delta = -(a_0 + \cdots + a_{p-1}). \tag{4.26}$$

Definition 4.3 A p-cyclic Darboux dressing chain (or factorization chain) with shift Δ is a sequence of p functions $w_0, \ldots, w_{p-1}$ and complex numbers $a_0, \ldots, a_{p-1}$ that satisfy the following coupled system of p Riccati-like ODEs

$$(w_i + w_{i+1})' + w_{i+1}^2 - w_i^2 = a_i, \qquad i = 0, 1, \ldots, p-1 \mod (p) \tag{4.27}$$

subject to the condition (4.26).

Note that transformation

$$w_i \mapsto -w_{-i}, \quad a_i \mapsto -a_{-i}, \quad \Delta \mapsto -\Delta \tag{4.28}$$

projects the reversal symmetry to the finite-dimensional system (4.27). Moreover, for $j = 0, 1 \ldots, p-1$ we also have the cyclic symmetry

$$w_i \mapsto w_{i+j}, \quad a_i \mapsto a_{i+j}, \quad \Delta \mapsto \Delta \qquad i = 0, \ldots p-1 \mod (p)$$

In the classification of solutions to (4.27) it will be convenient to regard two solutions related by a reversal symmetry or by a cyclic permutation as being equivalent.

Adding the p equations (4.27) we immediately obtain a first integral of the system

$$\sum_{j=0}^{p-1} w_j = \tfrac{1}{2} z \sum_{j=0}^{p-1} a_j = -\tfrac{1}{2}\Delta z.$$

The equivalence between the A_{2n}-Painlevé system (4.15) and the cyclic dressing chain (4.27) is given by the following proposition.

Proposition 4.4 *If the tuple of functions and complex numbers* $(w_0, \ldots, w_{2n} | a_0, \ldots, a_{2n})$ *satisfies a* $(2n+1)$*-cyclic Darboux dressing chain with shift* Δ *as per Definition* 4.3*, then the tuple* $\left(f_0, \ldots, f_{2n} \,\middle|\, \alpha_0, \ldots, \alpha_{2n}\right)$ *with*

$$f_i(z) = c\,(w_i + w_{i+1})\,(cz)\,, \qquad i = 0, \ldots, 2n \mod (2n+1), \tag{4.29}$$

$$\alpha_i = c^2 a_i, \tag{4.30}$$

$$c^2 = -\frac{1}{\Delta} \tag{4.31}$$

solves the A_{2n}*-Painlevé system* (4.15) *with normalization* (4.16).

Proof The linear transformation

$$f_i = w_i + w_{i+1}, \qquad i = 0, \ldots, 2n \mod (2n+1) \tag{4.32}$$

is invertible (only in the odd case $p = 2n+1$), the inverse transformation being

$$w_i = \tfrac{1}{2} \sum_{j=0}^{2n} (-1)^j f_{i+j}, \qquad i = 0, \ldots, 2n \mod (2n+1) \tag{4.33}$$

They imply the relations

$$w_{i+1} - w_i = \sum_{j=0}^{2n-1} (-1)^j f_{i+j+1}, \qquad i = 0, \ldots, 2n \mod (2n+1). \tag{4.34}$$

Inserting (4.32) and (4.34) into the equations of the cyclic dressing chain (4.27) leads to the A_{2n}-Painlevé system (4.15). For any constant $c \in \mathbb{C}$, the scaling transformation

$$f_i \mapsto cf_i, \quad z \mapsto cz, \quad \alpha_i \mapsto c^2\alpha_i$$

preserves the form of the equations (4.15). The choice $c^2 = -\frac{1}{\Delta}$ ensures that the normalization (4.16) always holds, for dressing chains with different shifts Δ. □

Remark 4.5 $(2n)$-cyclic dressing chains and A_{2n-1}-Painlevé systems are also related, but the mapping is given by a rational rather than a linear function. A full treatment of this even cyclic case (which includes P_V and its higher order hierarchy) is considerably harder and shall be treated elsewhere.

The problem now becomes that of finding and classifying cyclic dressing chains, i.e. Schrödinger operators and sequences of Darboux transformations that reproduce the initial potential up to an additive shift Δ after a fixed given number of transformations.

The theory of exceptional polynomials is intimately related with families of Schrödinger operators connected by Darboux transformations [20, 27]. Constructing cyclic dressing chains on this class of potentials becomes a feasible task, and knowledge of the effect of rational Darboux transformations on the potentials suggests that the only family of potentials to be considered in the case of odd cyclic dressing chains are the rational extensions of the harmonic oscillator [26], which are exactly solvable potentials whose eigenfunctions are expressible in terms of exceptional Hermite polynomials.

Each potential in this class can be indexed by a finite set of integers (specifying the sequence of Darboux transformations applied on the harmonic oscillator that lead to the potential), or equivalently by a Maya diagram, which becomes very useful representation to capture a notion of equivalence and relations of the type (4.25).

As mentioned before, the fact that all rational odd cyclic dressing chains (and equivalently rational solutions to the A_{2n}-Painlevé system) must *necessarily* belong to this class remains an open question. We conjecture that this is indeed the case, and no rational solutions other than the ones described in the following sections exist.

5 Rational Extensions of the Harmonic Oscillator

5.1 Maya Diagrams

In this section we construct odd cyclic dressing chains on potentials belonging to the class of rational extensions of the harmonic oscillator. Every such potential is represented by a Maya diagram, a rational Darboux transformation acting on this class will be a flip operation on a Maya diagram and cyclic Darboux chains correspond to cyclic Maya diagrams. With this representation, the main problem of constructing rational cyclic Darboux chains becomes purely algebraic and combinatorial.

Following Noumi [46], we define a Maya diagram in the following manner.

Definition 5.1 A Maya diagram is a set of integers $M \subset \mathbb{Z}$ that contains a finite number of positive integers, and excludes a finite number of negative integers. We will use $\mathcal{M}$ to denote the set of all Maya diagrams.

Definition 5.2 Let $m_1 > m_2 > \cdots$ be the elements of a Maya diagram M arranged in decreasing order. By assumption, there exists a unique integer $s_M \in \mathbb{Z}$ such that $m_i = -i + s_M$ for all i sufficiently large. We define s_M to be the index of M.

We visualize a Maya diagram as a horizontally extended sequence of ▣ and □ symbols with the filled symbol ▣ in position i indicating membership $i \in M$. The defining assumption now manifests as the condition that a Maya diagram begins with an infinite filled ▣ segment and terminates with an infinite empty □ segment.

Definition 5.3 Let M be a Maya diagram, and

$$M_- = \{-m-1 : m \notin M, m < 0\}, \qquad M_+ = \{m : m \in M\,, m \geq 0\}.$$

Let $s_1 > s_2 > \cdots > s_p$ and $t_1 > t_2 > \cdots > t_q$ be the elements of M_- and M_+ arranged in descending order.

We define the *Frobenius symbol* of M to be the double list $(s_1, \ldots, s_p \mid t_q, \ldots, t_1)$.

It is not hard to show that $s_M = q - p$ is the index of M. The classical Frobenius symbol [3, 4, 55] corresponds to the zero index case where $q = p$. If M is a Maya diagram, then for any $k \in \mathbb{Z}$ so is

$$M + k = \{m + k : m \in M\}.$$

The behaviour of the index s_M under translation of k is given by

$$M' = M + k \quad \Rightarrow \quad s_{M'} = s_M + k. \tag{5.1}$$

We will refer to an equivalence class of Maya diagrams related by such shifts as an *unlabelled Maya diagram*. One can visualize the passage from an unlabelled to a labelled Maya diagram as the choice of placement of the origin.

A Maya diagram $M \subset \mathbb{Z}$ is said to be in standard form if $p = 0$ and $t_q > 0$. Visually, a Maya diagram in standard form has only filled boxes ▣ to the left of the origin and one empty box □ just to the right of the origin. Every unlabelled Maya diagram permits a unique placement of the origin so as to obtain a Maya diagram in standard form.

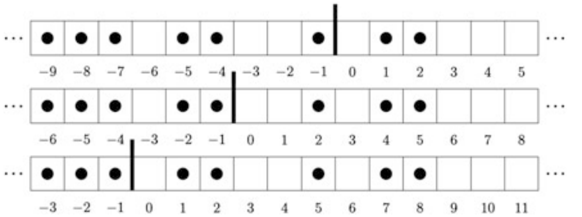


Fig. 1 Three equivalent Maya diagrams corresponding to the partition $\lambda = (4, 4, 3, 1, 1)$, together with their Frobenius representation

Exercise 9 Draw the box-and-ball representation of the Maya diagram

$$M = \{\dots, -9, -8, -7, -5, -4, -1, 1, 2\}.$$

Find the Frobenius symbol and the index of M. Find a translation k such that $M' = M + k$ is in standard form, and write the Frobenius symbol, index and box-and-ball representation of M'.

The solution to the previous exercise can be found in Fig. 1.

Observe that the third diagram is in standard form, so $k = 6$ is the necessary shift.

5.2 *Hermite Pseudo-Wronskians*

We can interpret a Maya diagram with Frobenius symbol $(s_1, \dots, s_r | t_q, \dots, t_1)$ as the multi-index that specifies a multi-step rational Darboux transformation on the harmonic oscillator, i.e. $L \mapsto L_M$, where

$$L_M = -D_{xx} + x^2 - 2\left(\mathrm{Wr}[\varphi_{-s_1}, \dots, \varphi_{-s_r}, \varphi_{t_1}, \dots, \varphi_{t_q}]\right)_{xx}$$

where φ_k are the seed functions for rational Darboux transformations of the harmonic oscillator described in (2.7). The first tuple in the Frobenius symbol specifies seed functions with conjugate Hermite polynomials in (2.7) (virtual states) while the second tuple specifies the bound states in (2.7). Getting rid of an overall exponential factor, we can associate to every Maya diagram a polynomial called a Hermite pseudo-Wronskian.

Definition 5.4 Let M be a Maya diagram and $(s_1, \dots, s_r | t_q, \dots, t_1)$ its corresponding Frobenius symbol. Define the polynomial

$$H_M = e^{-rx^2} \mathrm{Wr}[e^{x^2} \tilde{H}_{s_1}, \dots, e^{x^2} \tilde{H}_{s_r}, H_{t_q}, \dots H_{t_1}], \tag{5.2}$$

where Wr denotes the Wronskian determinant of the indicated functions, and

$$\tilde{H}_n(x) = \mathrm{i}^{-n} H_n(\mathrm{i}x) \tag{5.3}$$

is the nth degree conjugate Hermite polynomial.

It is not evident that H_M in (5.2) is a polynomial, but this becomes clear once we represent it using a slightly different determinant.

Proposition 5.5 *The Wronskian H_M admits the following alternative determinantal representation*

$$H_M = \begin{vmatrix} \tilde{H}_{s_1} & \tilde{H}_{s_1+1} & \dots & \tilde{H}_{s_1+r+q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{s_r} & \tilde{H}_{s_r+1} & \dots & \tilde{H}_{s_r+r+q-1} \\ H_{t_q} & D_x H_{t_q} & \dots & D_x^{r+q-1} H_{t_q} \\ \vdots & \vdots & \ddots & \vdots \\ H_{t_1} & D_x H_{t_1} & \dots & D_x^{r+q-1} H_{t_1} \end{vmatrix} \tag{5.4}$$

The term Hermite pseudo-Wronskian was coined in [29] because (5.4) is a mix of a Casoratian and a Wronskian determinant.

Exercise 10 Prove Proposition 5.5, i.e. prove the relation

$$H_M = e^{-rx^2}\,\mathrm{Wr}[e^{x^2}\tilde{H}_{s_1}, \dots, e^{x^2}\tilde{H}_{s_r}, H_{t_q}, \dots H_{t_1}] = \begin{vmatrix} \tilde{H}_{s_1} & \tilde{H}_{s_1+1} & \dots & \tilde{H}_{s_1+r+q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{s_r} & \tilde{H}_{s_r+1} & \dots & \tilde{H}_{s_r+r+q-1} \\ H_{t_q} & D_x H_{t_q} & \dots & D_x^{r+q-1} H_{t_q} \\ \vdots & \vdots & \ddots & \vdots \\ H_{t_1} & D_x H_{t_1} & \dots & D_x^{r+q-1} H_{t_1} \end{vmatrix}$$

The desired identity follows by the fundamental relations satisfied by Hermite polynomials

$$\begin{aligned} D_x H_n(x) &= 2n H_{n-1}(x), \quad n \ge 0, \\ D_x \tilde{H}_n(x) &= 2n \tilde{H}_{n-1}(x), \quad n \ge 0, \\ 2x H_n(x) &= H_{n+1}(x) + 2n H_{n-1}(x), \\ 2x \tilde{H}_n(x) &= \tilde{H}_{n+1}(x) - 2n \tilde{H}_{n-1}(x), \\ D_x(e^{x^2}\tilde{H}_n(x)) &= e^{x^2}\tilde{H}_{n+1}(x), \\ D_x(e^{-x^2}h_n(x)) &= -e^{-x^2}h_{n+1}(x). \end{aligned} \tag{5.5}$$

together with the Wronskian identity

$$\mathrm{Wr}[gf_1, \ldots, gf_s] = g^s\, \mathrm{Wr}[f_1, \ldots, f_s], \tag{5.6}$$

One remarkable property satisfied by all Maya diagrams in the same equivalence class, is that their associated Hermite pseudo-Wronskians enjoy a very simple relation: with an appropriate scaling, the Hermite pseudo-Wronskian of a given Maya diagram is invariant under translations.

Proposition 5.6 *Let $\hat{H}_M$ be the normalized pseudo-Wronskian*

$$\hat{H}_M = \frac{(-1)^{rq} H_M}{\prod_{1\le i<j\le r}(2s_j - 2s_i)\prod_{1\le i<j\le q}(2t_i - 2t_j)}. \tag{5.7}$$

Then for any Maya diagram M and $k \in \mathbb{Z}$ we have

$$\hat{H}_M = \hat{H}_{M+k}. \tag{5.8}$$

The proof of this Proposition is not too hard and proceeds by induction: it is enough to prove the equality by a shift of $k = 1$. We leave it as an exercise for the interested reader. The proof can be seen in [29]. At least, to gain some practice and convince ourselves of this result, we propose the following exercise.

Exercise 11 Let M be the Maya diagram with Frobenius symbol $(3, 2|2, 4)$. Write down $\hat{H}_M$, $\hat{H}_{M+3}$ and $\hat{H}_{M-4}$. Compute the determinants and check that (5.8) is verified.

The remarkable aspect of Eq. (5.8) is that the identity involves determinants of different sizes. As mentioned above, every unlabelled Maya diagram contains a Maya diagram in standard form, and its associated Hermite pseudo-Wronskian (5.2) is just an ordinary Wronskian determinant whose entries are Hermite polynomials.An interesting problem is to determine the smallest determinant in a given equivalence class, i.e. the minimum number of Darboux transformations to reach a given potential. The details on how to solve this problem are given in [29].

Due to Proposition 5.6, we could restrict the analysis without loss of generality to Maya diagrams in standard form and Wronskians of Hermite polynomials, but we will employ the general notation as it brings conceptual clarity to the description of Maya cycles.

We will now introduce and study a class of potentials for Schrödinger operators that will be used as building blocks for cyclic dressing chains: the set of rational extensions of the harmonic oscillator, which, as we will see, amounts to the set of potentials that one can obtain from $U(x) = x^2$ by applying rational Darboux-Crum transformations.

5.3 Rational Extensions of the Harmonic Oscillator

Definition 5.7 A rational extension of the harmonic oscillator is a potential of the form

$$U(x) = x^2 + \frac{a(x)}{b(x)}, \qquad a, b \text{ polynomials}, \qquad \deg a \leq \deg b,$$

that is *exactly solvable by polynomials*, in the sense of Definition 2.1.

If $b(x)$ has no real zeros, then L is a Sturm-Liouville operator on $\mathbb{R}$ with quasi-polynomial eigenfunctions. The next Proposition proved in [26] states that rational extensions of the harmonic oscillator can be put in one to one correspondence with Maya diagrams. The details of this result are based on the theory of trivial monodromy potentials and they exceed the scope of these lecture notes. The interested reader is referred to [26] and [48] for further details.

Proposition 5.8 *Let $M \subset \mathbb{Z}$ be a Maya diagram. Define*

$$U_M(x) = x^2 - 2D_x^2 \log H_M + 2s_M, \tag{5.9}$$

where H_M is the corresponding pseudo-Wronskian (5.2)–(5.4)*, and $s_M \in \mathbb{Z}$ is the index of M. Up to an additive constant, every rational extension of the harmonic oscillator takes the form* (5.9)*.*

The class of Schrödinger operators with potentials that are rational extensions of the harmonic oscillator is invariant under a rational Darboux transformations. Otherwise speaking, if we perform a rational Darboux transformation on a rational extension of the harmonic oscillator, indexed by a Maya diagram M, we will obtain another potential in the same class, indexed by M'. Both M and M' differ only in one element, as we show next.

Definition 5.9 We define the flip at position $m \in \mathbb{Z}$ to be the involution $\phi_m : \mathcal{M} \to \mathcal{M}$ defined by

$$\phi_m : M \mapsto \begin{cases} M \cup \{m\} & \text{if } m \notin M \\ M \setminus \{m\} & \text{if } m \in M \end{cases}, \qquad M \in \mathcal{M}. \tag{5.10}$$

In the first case, we say that ϕ_m acts on M by a state-deleting transformation ($\square \to \boxdot$). In the second case, we say that ϕ_m acts by a state-adding transformation ($\boxdot \to \square$).

Using Crum's formula for iterated Darboux transformations (2.8), and the seed functions for rational DTs of the harmonic oscillator (2.7), it can be shown that every quasi-rational eigenfunction of $L = -D_x^2 + U_M(x)$ has the form

$$\psi_{M,m} = e^{\epsilon x^2/2} \frac{H_{\phi_m(M)}}{H_M}, \qquad m \in \mathbb{Z}, \tag{5.11}$$

with

$$\epsilon = \begin{cases} -1 & \text{if } m \notin M \\ +1 & \text{if } m \in M \end{cases}.$$

Explicitly, we have

$$L\psi_{M,m} = (2m+1)\psi_{M,m}, \quad m \in \mathbb{Z}. \tag{5.12}$$

Remark 5.10 The seed eigenfunctions (5.11) include the true eigenfunctions of L plus other set of formal non square-integrable eigenfunctions, sometimes known in the physics literature as *virtual states*,[51, 53]. For a correct spectral theoretic interpretation one needs to ensure that the potential U_M is regular, i.e. that H_M has no zeros in $\mathbb{R}$. The set of Maya diagrams for which H_M has no real zeros was characterized (in a more general setting) independently by Krein [36] and Adler [1], while the number of real zeros for H_M was given in [19]. However, for the purpose of this paper it is convenient stay within a purely formal setting and keep the whole class of potentials U_M, regardless of whether they have real poles or not.

The relation between dressing chains of Darboux transformations for the class of operators (5.9) and flip operations on Maya diagrams is made explicit by the following proposition.

Proposition 5.11 *Two Maya diagrams M, M' are related by a flip* (5.10) *if and only if their associated rational extensions $U_M, U_{M'}$ are connected by a Darboux transformation* (4.21)*.*

Proof Suppose that $m \notin M$ and that $M' = M \cup \{m\}$ is a state-deleting flip transformation of M. The seed function for the factorization is $\psi_{M,m}$ defined in (5.11). Set

$$w_{M,m} = \frac{\psi'_{M,m}}{\psi_{M,m}} = -x + \frac{H'_{M'}}{H_{M'}} - \frac{H'_M}{H_M}. \tag{5.13}$$

Since

$$s_{M'} = s_M + 1,$$

by (5.9), we have

$$\frac{1}{2}(U_{M'} - U_M) = 1 + D_x\left(\frac{H'_M}{H_M} - \frac{H'_{M'}}{H_{M'}}\right) = -w'_{M,m}, \tag{5.14}$$

so that (4.21) holds. Conversely, suppose that M and M' are such that (5.14) holds for some $w = w(x)$. If we define

$$w = \frac{\psi'}{\psi}, \qquad \psi = e^{-x^2/2}\frac{H_{M'}}{H_M},$$

then ψ must be a quasi-rational seed function for U_M and it follows by (5.11) of Proposition 5.8 that that $M' = M \cup \{m\}$ for some $m \notin M$. The corresponding result for state-adding Darboux transformations is done in a similar way. □

We see thus that the class of rational extensions of the harmonic oscillator is indexed by Maya diagrams, and that the Darboux transformations that preserve this class can be described by flip operations on Maya diagrams. Now we are ready to introduce the concept of cyclic Maya diagrams, and use them later to build Darboux dressing chains on these potentials, and solutions to $A_N^{(1)}$-Painlevé.

5.4 Cyclic Maya Diagrams

Cyclic Maya diagrams are just the ones such that we can perform a number of flip operations on them, and recover the same Maya diagram up to a shift, [30]. We introduce the necessary notation and precise definitions below.

Definition 5.12 For $p \in \mathbb{N}$ let $\mathcal{Z}_p$ denote the set of all subsets of $\mathbb{Z}$ having cardinality p. For $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_p\} \in \mathcal{Z}_p$ we now define $\phi_{\boldsymbol{\mu}}$ to be the multi-flip

$$\phi_{\boldsymbol{\mu}} = \phi_{\mu_1} \circ \cdots \circ \phi_{\mu_p}. \tag{5.15}$$

Definition 5.13 We say that M is p-cyclic with shift k, or (p, k) cyclic, if there exists a $\boldsymbol{\mu} \in \mathcal{Z}_p$ such that

$$\phi_{\boldsymbol{\mu}}(M) = M + k. \tag{5.16}$$

We will say that M is p-cyclic if it is (p, k) cyclic for some $k \in \mathbb{Z}$.

Proposition 5.14 *For Maya diagrams $M, M' \in \mathcal{M}$, define the set*

$$\Upsilon(M, M') = (M \setminus M') \cup (M' \setminus M) \tag{5.17}$$

Then the multi-flip $\phi_{\boldsymbol{\mu}}$ where $\boldsymbol{\mu} = \Upsilon(M, M')$ is the unique multi-flip such that $M' = \phi_{\boldsymbol{\mu}}(M)$ and $\Upsilon(M, M')$ $M = \phi_{\boldsymbol{\mu}}(M')$.

Intuitively, is the set of sites at which M and M' differ, so it is evident that a multi-flip on these sites will turn M into M' and viceversa. As an immediate corollary, we have the following.

Proposition 5.15 *Let k be a non-zero integer. Every Maya diagram $M \in \mathcal{M}$ is (p, k) cyclic where p is the cardinality of $\boldsymbol{\mu} = \Upsilon(M, M + k)$.*

Exercise 12 For the following Maya diagrams, find the sequence of flip transformations $\boldsymbol{\mu} = \{\mu_0, \mu_1, \mu_2\}$ such that $M' = \phi_{\boldsymbol{\mu}}(M) = M + k$

$$k = 1 \qquad M = (\emptyset|3, 4, 5, 6) = (-\infty, -1] \cup \{3, 4, 5, 6\}$$
$$k = 3 \qquad M = (3|1, 2, 4, 5, 8) = (-\infty, -4] \cup \{-2, -1\} \cup \{1, 2, 4, 5, 8\}$$

In the first case, we see that $M' = M + 1 = (\emptyset|0, 4, 5, 6, 7)$, so $\boldsymbol{\mu} = \Upsilon(M, M + 1) = (0, 3, 7)$. The first and third flips correspond to state-deleting transformations ($\square \to \boxed{\bullet}$), while the second is a state-adding transformation ($\boxed{\bullet} \to \square$). In the second case, we have

$$M' = M + 3 = (\emptyset|1, 2, 4, 5, 7, 8, 11) = (-\infty, -1] \cup \{1, 2, 4, 5, 7, 8, 11\}$$

so $\boldsymbol{\mu} = \Upsilon(M, M + 3) = (-3, 7, 11)$. In this case, all three transformations are state-deleting ($\square \to \boxed{\bullet}$).

Now we are able to establish the link between Maya cycles and cyclic dressing chains composed of rational extensions of the harmonic oscillator.

Theorem 5.16 *Let $M \in \mathcal{M}$ be a Maya diagram, k a non-zero integer, and p the cardinality of $\boldsymbol{\mu} = \Upsilon(M, M + k)$. Let $\boldsymbol{\mu} = \{\mu_0, \ldots, \mu_{p-1}\}$ be an arbitrary enumeration of $\boldsymbol{\mu}$ and set*

$$M_0 = M, \quad M_{i+1} = \phi_{\mu_i}(M_i), \qquad i = 0, 1, \ldots, p - 1 \tag{5.18}$$

so that $M_p = M_0 + k$ by construction. Set

$$w_i = s_i\, x + \frac{H'_{M_{i+1}}}{H_{M_{i+1}}} - \frac{H'_{M_i}}{H_{M_i}}, \qquad i = 0, \ldots, p - 1. \tag{5.19}$$

$$\alpha_i = 2(\mu_i - \mu_{i+1}), \tag{5.20}$$

where

$$s_i = \begin{cases} -1 & \text{if } \mu_i \notin M \\ +1 & \text{if } \mu_i \in M \end{cases}, \tag{5.21}$$

and

$$\mu_p = \mu_0 + k.$$

Then, $(w_0, \ldots, w_{p-1}; \alpha_0, \ldots, \alpha_{p-1})$ constitutes a rational solution to the p-cyclic dressing chain (4.27) with shift $\Delta = 2k$.

Proof The result follows from the structure of the seed eigenfunctions (5.11) with eigenvalues given by (5.12), after applying (4.20) and (4.23). The sign of s_i indicates whether the $(i+1)$-th step of the chain that takes L_i to L_{i+1} is a state-adding $(+1)$ or state-deleting (-1) transformation. □

So now we know that given a Maya n-cycle, we can build an n-cyclic dressing chain and a rational solution to the Noumi-Yamada system. But we would like to go further and classify cyclic Maya diagrams for any given (odd) period, which we tackle next.

Remark 5.17 Under the correspondence described by Proposition 5.16, the reversal symmetry (4.28) manifests as the transformation

$$(M_0, \dots, M_p) \mapsto (M_p, \dots, M_0), \quad (\mu_1, \dots, \mu_p) \mapsto (\mu_p, \dots, \mu_1), \quad k \mapsto -k.$$

In light of the above remark, there is no loss of generality if we restrict our attention to cyclic Maya diagrams with a positive shift $k > 0$.

6 Classification of Cyclic Maya Diagrams

In this section we introduce two new concepts on Maya diagrams: *genus* and *interlacing*, which become a key ingredient in the characterization of cyclic Maya diagrams. But before we do so, let us introduce another way to specify a Maya diagram, which becomes more convenient for the task that we now face.

For $\boldsymbol{\beta} \in \mathcal{Z}_{2g+1}$ define the Maya diagram

$$\Xi(\boldsymbol{\beta}) = (-\infty, \beta_0) \cup [\beta_1, \beta_2) \cup \dots \cup [\beta_{2g-1}, \beta_{2g}) \tag{6.1}$$

where

$$[m, n) = \{j \in \mathbb{Z} \colon m \le j < n\}$$

and where $\beta_0 < \beta_1 < \dots < \beta_{2g}$ is the strictly increasing enumeration of $\boldsymbol{\beta}$.

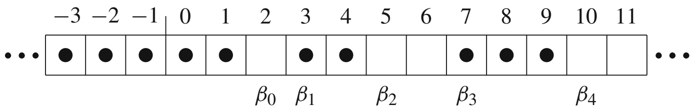
Proposition 6.1 *Every Maya diagram $M \in \mathcal{M}$ has a unique representation of the form $M = \Xi(\boldsymbol{\beta})$ where $\boldsymbol{\beta}$ is a set of integers of odd cardinality* $2g+1$.

Definition 6.2 We call the integer $g \ge 0$ the genus of $M = \Xi(\boldsymbol{\beta})$ and $(\beta_0, \beta_1, \dots, \beta_{2g})$ the block coordinates of M.

Remark 6.3 To motivate Definition 6.2, it is perhaps more illustrative to understand the visual meaning of the genus of M, see the Maya diagram below. After removal of the initial infinite ▣ segment and the trailing infinite □ segment, a Maya diagram consists of alternating empty □ and filled ▣ segments of variable length. The genus g counts the number of such pairs. The even block coordinates β_{2i} indicate the starting positions of the empty segments, and the odd block coordinates β_{2i+1} indicated the starting positions of the filled segments. Also, note that M is in standard form if and only if $\beta_0 = 0$.

Exercise 13 Draw the box-ball diagram corresponding to the genus-2 Maya diagram with block coordinates $(\beta_0, \ldots, \beta_4) = (2, 3, 5, 7, 10)$ and give its Frobenius symbol.

Since $\beta_0 = 2$, to the left of 2 all sites are filled and site 2 is empty. Next we have filled block $[\beta_1, \beta_2) = [3, 5)$ of size 2 and another filled block at $[\beta_3, \beta_3) = [7, 10)$ of size 3. All sites are empty to the right of $\beta_4 = 10$.



$$M = (-\infty, \beta_0) \cup [\beta_1, \beta_2) \cup [\beta_3, \beta_4)$$

Note that the genus is both the number of finite-size empty blocks and the number of finite-size filled blocks.

Exercise 14 Let $M = \Xi(\boldsymbol{\beta})$ be a Maya diagram specified by its block coordinates . Prove that

$$\boldsymbol{\beta} = \Upsilon(M, M + 1).$$

Proof Observe that

$$M + 1 = (-\infty, \beta_0] \cup (\beta_1, \beta_2] \cup \cdots \cup (\beta_{2g-1}, \beta_{2g}],$$

where

$$(m, n] = \{j \in \mathbb{Z} : m < j \leq n\}.$$

It follows that

$$(M + 1) \setminus M = \{\beta_0, \ldots, \beta_{2g}\}$$
$$M \setminus (M + 1) = \{\beta_1, \ldots, \beta_{2g-1}\}.$$

The desired conclusion follows immediately. $\square$

Let $\mathcal{M}_g$ denote the set of Maya diagrams of genus g. The above discussion may be summarized by saying that the mapping (14) defines a bijection $\Xi : \mathcal{Z}_{2g+1} \to \mathcal{M}_g$, and that the block coordinates are precisely the flip sites required for a translation $M \mapsto M+1$.

The next concept we need to introduce is the interlacing and modular decomposition.

Definition 6.4 Fix a $k \in \mathbb{N}$ and let $M^{(0)}, M^{(1)}, \dots M^{(k-1)} \subset \mathbb{Z}$ be sets of integers. We define the interlacing of these to be the set

$$\Theta\left(M^{(0)}, M^{(1)}, \dots M^{(k-1)}\right) = \bigcup_{i=0}^{k-1} (kM^{(i)} + i), \tag{6.2}$$

where

$$kM + j = \{km + j : m \in M\}, \quad M \subset \mathbb{Z}.$$

Dually, given a set of integers $M \subset \mathbb{Z}$ and a $k \in \mathbb{N}$ define the sets

$$M^{(i)} = \{m \in \mathbb{Z} : km + i \in M\}, \quad i = 0, 1, \dots, k-1.$$

We will call the k-tuple of sets $\left(M^{(0)}, M^{(1)}, \dots M^{(k-1)}\right)$ the k-modular decomposition of M.

The following result follows directly from the above definitions.

Proposition 6.5 *We have* $M = \Theta\left(M^{(0)}, M^{(1)}, \dots M^{(k-1)}\right)$ *if and only if* $\left(M^{(0)}, M^{(1)}, \dots M^{(k-1)}\right)$ *is the k-modular decomposition of M.*

Even though the above operations of interlacing and modular decomposition apply to general sets, they have a well defined restriction to Maya diagrams. Indeed, it is not hard to check that if $M = \Theta\left(M^{(0)}, M^{(1)}, \dots M^{(k-1)}\right)$ and M is a Maya diagram, then $M^{(0)}, M^{(1)}, \dots M^{(k-1)}$ are also Maya diagrams. Conversely, if the latter are all Maya diagrams, then so is M. Another important case concerns the interlacing of finite sets. The definition (6.2) implies directly that if $\boldsymbol{\mu}^{(i)} \in \mathcal{Z}_{p_i}$, $i = 0, 1, \dots, k-1$ then

$$\boldsymbol{\mu} = \Theta\left(\boldsymbol{\mu}^{(0)}, \dots, \boldsymbol{\mu}^{(k-1)}\right)$$

is a finite set of cardinality $p = p_0 + \cdots + p_{k-1}$.

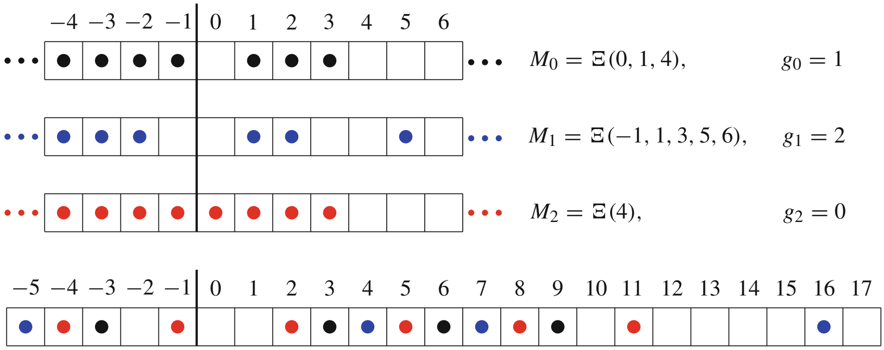
Visually, each of the k Maya diagrams is dilated by a factor of k, shifted by one unit with respect to the previous one and superimposed, so the interlaced Maya diagram incorporates the information from $M^{(0)}, \dots M^{(k-1)}$ in k different modular classes. In other words, the interlaced Maya diagram is built by copying sequentially a filled or empty box as determined by each of the k Maya diagrams.

Exercise 15 For the following three Maya diagrams, given by their block coordinates:

$$M_0 = \Xi(0, 1, 4), \quad M_1 = \Xi(-1, 1, 3, 5, 6), \quad M_2 = \Xi(4)$$

Draw the box-square diagram of the interlaced diagram $M = \Theta(M_0, M_1, M_2)$ and give the block coordinates and the 3-block coordinates of M.

Equipped with these notions of genus and interlacing, we are now ready to state the main result for the classification of cyclic Maya diagrams.



$M = \Theta(M_0, M_1, M_2) = \Xi_3(0, 1, 4| - 1, 1, 3, 5, 6|4) = \Xi(-2, -1, 0, 2, 10, 11, 12, 16, 17)$

Theorem 6.6 *Let* $M = \Theta\left(M^{(0)}, M^{(1)}, \ldots M^{(k-1)}\right)$ *be the* k*-modular decomposition of a given Maya diagram* M. *Let* g_i *be the genus of* $M^{(i)}$, $i = 0, 1, \ldots, k-1$. *Then,* M *is* (p, k)*-cyclic where*

$$p = p_0 + p_1 + \cdots + p_{k-1}, \qquad p_i = 2g_i + 1. \tag{6.3}$$

Proof Let $\boldsymbol{\beta}^{(i)} = \Upsilon\left(M^{(i)}, M^{(i+1)}\right) \in \mathcal{Z}_{p_i}$ be the block coordinates of $M^{(i)}$, $i = 0, 1, \ldots, k-1$. Consider the interlacing $\boldsymbol{\mu} = \Theta\left(\boldsymbol{\beta}^{(0)}, \ldots, \boldsymbol{\beta}^{(k-1)}\right)$. From Exercise 14 we have that,

$$\phi_{\boldsymbol{\beta}^{(i)}}\left(M^{(i)}\right) = M^{(i)} + 1.$$

so it follows that

$$\begin{aligned}\phi_{\boldsymbol{\mu}}(M) &= \phi_{\Theta\left(\boldsymbol{\beta}^{(0)}, \ldots, \boldsymbol{\beta}^{(k-1)}\right)} \Theta\left(M^{(0)}, \ldots, M^{(k-1)}\right) \\ &= \Theta\left(\phi_{\boldsymbol{\beta}^{(0)}}(M^{(0)}), \ldots, \phi_{\boldsymbol{\beta}^{(k-1)}}(M^{(k-1)})\right)\end{aligned}$$

$$
\begin{aligned}
&= \Theta\left(M^{(0)}+1,\ldots,M^{(k-1)}+1\right)\\
&= \Theta\left(M^{(0)},\ldots,M^{(k-1)}\right)+k\\
&= M+k.
\end{aligned}
$$

Therefore, M is (p,k) cyclic where the value of p agrees with (6.3). □

Normally we will have a bound on the period p, and the classification problem is to describe all possible (p,k)-cyclic Maya diagrams for all values of $k > 0$. Theorem 6.6 sets the way to do this.

Corollary 6.7 *For a fixed period $p \in \mathbb{N}$, there exist p-cyclic Maya diagrams with shifts $k = p, p-2, \ldots, \lfloor p/2 \rfloor$, and no other positive shifts are possible.*

Remark 6.8 The highest shift $k = p$ corresponds to the interlacing of p trivial (genus 0) Maya diagrams.

6.1 Indexing Maya p-Cycles

We now introduce a combinatorial system for describing rational solutions of p-cyclic factorization chains. First, we require a generalized notion of block coordinates suitable for describing p-cyclic Maya diagrams.

Definition 6.9 Let $M = \Theta\left(M^{(0)},\ldots M^{(k-1)}\right)$ be a k-modular decomposition of a (p,k) cyclic Maya diagram. For $i = 0,1,\ldots,k-1$ let $\boldsymbol{\beta}^{(i)} = \left(\beta_0^{(i)},\ldots,\beta_{p_i-1}^{(i)}\right)$ be the block coordinates of $M^{(i)}$ enumerated in increasing order. In light of the fact that

$$
M = \Theta\left(\Xi(\boldsymbol{\beta}^{(0)}),\ldots,\Xi(\boldsymbol{\beta}^{(k-1)})\right),
$$

we will refer to the concatenated sequence

$$
\begin{aligned}
(\beta_0,\beta_1,\ldots,\beta_{p-1}) &= (\boldsymbol{\beta}^{(0)}|\boldsymbol{\beta}^{(1)}|\ldots|\boldsymbol{\beta}^{(k-1)})\\
&= \left(\beta_0^{(0)},\ldots,\beta_{p_0-1}^{(0)}|\beta_0^{(1)},\ldots,\beta_{p_1-1}^{(1)}|\ldots|\beta_0^{(k-1)},\ldots,\beta_{p_{k-1}-1}^{(k-1)}\right)
\end{aligned}
$$

as the k-block coordinates of M. Formally, the correspondence between k-block coordinates and Maya diagram is described by the mapping

$$
\Xi_k : \mathcal{Z}_{2g_0+1}\times\cdots\times\mathcal{Z}_{2g_{k-1}+1} \to \mathcal{M}
$$

with action

$$\Xi_k\colon (\boldsymbol{\beta}^{(0)}|\boldsymbol{\beta}^{(1)}|\ldots|\boldsymbol{\beta}^{(k-1)}) \mapsto \Theta\left(\Xi(\boldsymbol{\beta}^{(0)}),\ldots,\Xi(\boldsymbol{\beta}^{(k-1)})\right)$$

Definition 6.10 Fix a $k\in\mathbb{N}$. For $m\in\mathbb{Z}$ let $[m]_k\in\{0,1,\ldots,k-1\}$ denote the residue class of m modulo division by k. For $m,n\in\mathbb{Z}$ say that $m\preccurlyeq_k n$ if and only if

$$[m]_k<[n]_k,\qquad\text{or}\qquad [m]_k=[n]_k\ \text{ and } m\le n.$$

In this way, the transitive, reflexive relation $\preccurlyeq_k$ forms a total order on $\mathbb{Z}$.

Proposition 6.11 *Let M be a (p,k) cyclic Maya diagram. There exists a unique p-tuple of integers $(\mu_0,\ldots,\mu_{p-1})$ strictly ordered relative to $\preccurlyeq_k$ such that*

$$\phi_{\boldsymbol{\mu}}(M)=M+k \tag{6.4}$$

Proof Let $(\beta_0,\ldots,\beta_{p-1})=(\boldsymbol{\beta}^{(0)}|\boldsymbol{\beta}^{(1)}|\ldots|\boldsymbol{\beta}^{(k-1)})$ be the k-block coordinates of M. Set

$$\boldsymbol{\mu}=\Theta\left(\boldsymbol{\beta}^{(0)},\ldots,\boldsymbol{\beta}^{(k-1)}\right)$$

so that (6.4) holds by the proof to Theorem 6.6. The desired enumeration of $\boldsymbol{\mu}$ is given by

$$(k\beta_0,\ldots,k\beta_{p-1})+(0^{p_0},1^{p_1},\ldots,(k-1)^{p_{k-1}})$$

where the exponents indicate repetition. Explicitly, $(\mu_0,\ldots,\mu_{p-1})$ is given by

$$\left(k\beta_0^{(0)},\ldots,k\beta_{p_0-1}^{(0)},k\beta_0^{(1)}+1,\ldots,k\beta_{p_1-1}^{(1)}+1,\ldots,k\beta_0^{(k-1)}+k-1,\ldots,k\beta_{p_{k-1}-1}^{(k-1)}+k-1\right).$$

□

Definition 6.12 In light of (6.4) we will refer to the just defined tuple $(\mu_0,\mu_1,\ldots,\mu_{p-1})$ as the k-canonical flip sequence of M and refer to the tuple $(p_0,p_1,\ldots,p_{k-1})$ as the k-signature of M.

By Proposition 5.16 a rational solution of the p-cyclic dressing chain requires a (p,k) cyclic Maya diagram, and an additional item data, namely a fixed ordering of the canonical flip sequence. We will specify such ordering as

$$\boldsymbol{\mu}_{\boldsymbol{\pi}}=(\mu_{\pi_0},\ldots,\mu_{\pi_{p-1}})$$

where $\boldsymbol{\pi}=(\pi_0,\ldots,\pi_{p-1})$ is a permutation of $(0,1,\ldots,p-1)$. With this notation, the chain of Maya diagrams described in Proposition 5.16 is generated as

$$M_0=M,\qquad M_{i+1}=\phi_{\mu_{\pi_i}}(M_i),\qquad i=0,1,\ldots,p-1. \tag{6.5}$$

Remark 6.13 Using a translation it is possible to normalize M so that $\mu_0 = 0$. Using a cyclic permutation and it is possible to normalize $\boldsymbol{\pi}$ so that $\pi_p = 0$. The net effect of these two normalizations is to ensure that $M_0, M_1, \dots, M_{p-1}$ have standard form.

Remark 6.14 In the discussion so far we have imposed the hypothesis that the sequence of flips that produces a translation $M \mapsto M + k$ does not contain any repetitions. However, in order to obtain a full classification of rational solutions, it will be necessary to account for degenerate chains which include multiple flips at the same site.

To that end it is necessary to modify Definition 5.12 to allow $\boldsymbol{\mu}$ to be a multi-set,[3] and to allow $\mu_0, \mu_1, \dots, \mu_{p-1}$ in (14) to be merely a non-decreasing sequence. This has the effect of permitting $\square$ and $\boxdot$ segments of zero length wherever $\mu_{i+1} = \mu_i$. The Ξ-image of such a non-decreasing sequence is not necessarily a Maya diagram of genus g, but rather a Maya diagram whose genus is bounded above by g.

It is no longer possible to assert that there is a unique $\boldsymbol{\mu}$ such that $\phi_{\boldsymbol{\mu}}(M) = M + k$, because it is possible to augment the non-degenerate $\boldsymbol{\mu} = \Upsilon(M, M+k)$ with an arbitrary number of pairs of flips at the same site to arrive at a degenerate $\boldsymbol{\mu}'$ such that $\phi_{\boldsymbol{\mu}'}(M) = M + k$ also. The rest of the theory remains unchanged.

6.2 *Rational Solutions of A_4-Painlevé*

In this section we will put together all the results derived above in order to describe an effective way of labelling and constructing all the rational solutions to the A_{2k}-Painlevé system based on cyclic dressing chains of rational extensions of the harmonic oscillator [10]. We conjecture that the construction described below covers all rational solutions to such systems. As an illustrative example, we describe all rational solutions to the A_4-Painlevé system, and we furnish examples in each signature class.

For odd p, in order to specify a Maya p-cycle, or equivalently a rational solution of a p-cyclic dressing chain, we need to specify three items of data:

1. a signature sequence $(p_0, \dots, p_{k-1})$ consisting of odd positive integers that sum to p. This sequence determines the genus of the k interlaced Maya diagrams that give rise to a (p, k)-cyclic Maya diagram M. The possible values of k are given by Corollary 6.7.
2. Once the signature is fixed, we need to specify the k-block coordinates

$$(\beta_0, \dots, \beta_{p-1}) = (\boldsymbol{\beta}^{(0)}| \dots |\boldsymbol{\beta}^{(k-1)})$$

[3] A multi-set is generalization of the concept of a set that allows for multiple instances for each of its elements.

where $\boldsymbol{\beta}^{(i)} = (\beta_0^{(i)}, \ldots, \beta_{p_i}^{(i)})$ are the block coordinates that define each of the interlaced Maya diagrams $M^{(i)}$. These two items of data specify uniquely a (p, k)-cyclic Maya diagram M, and a canonical flip sequence $\boldsymbol{\mu} = (\beta_0, \ldots, \beta_{p-1})$. The next item specifies a given p-cycle that contains M.
3. Once the k-block coordinates and canonical flip sequence $\boldsymbol{\mu}$ are fixed, we still have the freedom to choose a permutation $\boldsymbol{\pi} \in S_p$ of $(0, 1, \ldots, p-1)$ that specifies the actual flip sequence $\boldsymbol{\mu}_{\boldsymbol{\pi}}$, i.e. the order in which the flips in the canonical flip sequence are applied to build the Maya p-cycle.

For any signature of a Maya p-cycle, we need to specify the p integers in the canonical flip sequence, but following Remark 6.13, we can get rid of translation invariance by setting $\mu_0 = \beta_0^{(0)} = 0$, leaving only $p-1$ free integers. Moreover, we can restrict ourselves to permutations such that $\boldsymbol{\pi}_p = 0$ in order to remove the invariance under cyclic permutations. The remaining number of degrees of freedom is $p-1$, which (perhaps not surprisingly) coincides with the number of generators of the symmetry group $A_{p-1}^{(1)}$. This is a strong indication that the class described above captures a generic orbit of a seed solution under the action of the symmetry group.

We now illustrate the general theory by describing the rational solutions of the $A_4^{(1)}$- Painlevé system [17, 44], whose equations are given by

$$\begin{aligned} f_0' + f_0(f_1 - f_2 + f_3 - f_4) &= \alpha_0, \\ f_1' + f_1(f_2 - f_3 + f_4 - f_0) &= \alpha_1, \\ f_2' + f_2(f_3 - f_4 + f_0 - f_1) &= \alpha_2, \\ f_3' + f_3(f_4 - f_0 + f_1 - f_2) &= \alpha_3, \\ f_4' + f_4(f_0 - f_1 + f_2 - f_3) &= \alpha_4, \end{aligned} \tag{6.6}$$

with normalization

$$f_0 + f_1 + f_2 + f_3 + f_4 = z.$$

Theorem 6.15 *Rational solutions of the $A_4^{(1)}$-Painlevé system (6.6) correspond to chains of 5-cyclic Maya diagrams belonging to one of the following signature classes:*

$$(5), (3, 1, 1), (1, 3, 1), (1, 1, 3), (1, 1, 1, 1, 1).$$

With the normalization $\boldsymbol{\pi}_4 = 0$ and $\mu_0 = 0$, each rational solution may be uniquely labeled by one of the above signatures, a 4-tuple of arbitrary non-negative integers (n_1, n_2, n_3, n_4), and a permutation $(\pi_0, \pi_1, \pi_2, \pi_3)$ of $(1, 2, 3, 4)$. For each of the *above signatures, the corresponding k-block coordinates of the initial 5-cyclic Maya diagram are then given by*

$$\begin{array}{lll} k=1 & (5) & (0, n_1, n_1+n_2, n_1+n_2+n_3, n_1+n_2+n_3+n_4) \\ k=3 & (3,1,1) & (0, n_1, n_1+n_2|n_3|n_4) \\ k=3 & (1,3,1) & (0|n_1, n_1+n_2, n_1+n_2+n_3|n_4) \\ k=3 & (1,1,3) & (0|n_1|n_2, n_2+n_3, n_2+n_3+n_4) \\ k=5 & (1,1,1,1,1) & (0|n_1|n_2|n_3|n_4) \end{array}$$

We show specific examples with shifts $k = 1, 3$ and 5 and signatures (5), $(1, 1, 3)$ and $(1, 1, 1, 1, 1)$.

Exercise 16 Construct a $(5, 1)$-cyclic Maya diagram in the signature class (5) with $(n_1, n_2, n_3, n_4) = (2, 3, 1, 1)$ and permutation (34210). Build the corresponding set of rational solutions to A_4-Painlevé.

The first Maya diagram in the cycle is $M_0 = \Xi(0, 2, 5, 6, 7)$, depicted in the first row of Fig. 2. The canonical flip sequence is $\boldsymbol{\mu} = (0, 2, 5, 6, 7)$. The permutation (34210) gives the chain of Maya diagrams shown in Fig. 2. Note that the permutation specifies the sequence of block coordinates that get shifted by one at each step of the cycle. This type of solutions with signature (5) were already studied in [17], and they are based on a genus 2 generalization of the generalized Hermite polynomials that appear in the solution of $\mathrm{P_{IV}}$ (A_2-Painlevé).

We shall now provide the explicit construction of the rational solution to the A_4-Painlevé system (6.6), by using Propositions 5.16 and 4.4. The permutation $\boldsymbol{\pi} = (34210)$ on the canonical sequence $\boldsymbol{\mu} = (0, 2, 5, 6, 7)$ produces the flip sequence $\boldsymbol{\mu}_{\boldsymbol{\pi}} = (6, 7, 5, 2, 0)$, so that the values of the α_i parameters given by (5.20) become $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = (-2, 4, 6, 4, -14)$. The pseudo-Wronskians corresponding

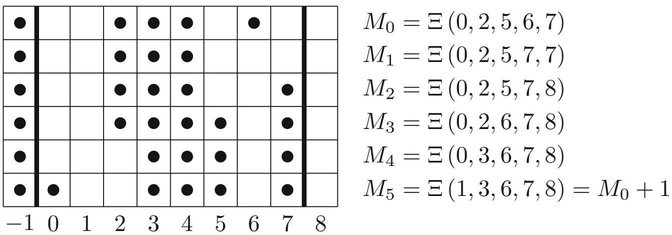


Fig. 2 A Maya 5-cycle with shift $k = 1$ for the choice $(n_1, n_2, n_3, n_4) = (2, 3, 1, 1)$ and permutation $\boldsymbol{\pi} = (34210)$

to each Maya diagram in the cycle are ordinary Wronskians, which will always be the case with the normalization imposed in Remark 6.13. They read (see Fig. 2):

$$\begin{aligned} H_{M_0}(z) &= \mathrm{Wr}(H_2, H_3, H_4, H_6) \\ H_{M_1}(z) &= \mathrm{Wr}(H_2, H_3, H_4) \\ H_{M_2}(z) &= \mathrm{Wr}(H_2, H_3, H_4, H_7) \\ H_{M_3}(z) &= \mathrm{Wr}(H_2, H_3, H_4, H_5, H_7) \\ H_{M_4}(z) &= \mathrm{Wr}(H_3, H_4, H_5, H_7) \end{aligned}$$

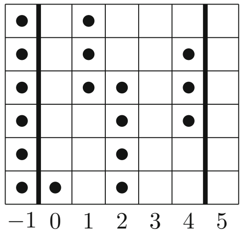
where $H_n = H_n(z)$ is the n-th Hermite polynomial. The rational solution to the dressing chain is given by the tuple $(w_0, w_1, w_2, w_3, w_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where α_i and w_i are given by (5.19) and (5.20) as:

$$\begin{aligned} w_0(z) &= z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(z) - \log H_{M_0}(z)\Big], && a_0 = -2, \\ w_1(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(z) - \log H_{M_1}(z)\Big], && a_1 = 4, \\ w_2(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(z) - \log H_{M_2}(z)\Big], && a_2 = 6, \\ w_3(z) &= z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(z) - \log H_{M_3}(z)\Big], && a_3 = 4, \\ w_4(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(z) - \log H_{M_4}(z)\Big], && a_4 = -14. \end{aligned}$$

Finally, Proposition 5.16 implies that the corresponding rational solution to the A_4-Painlevé system (4.15) is given by the tuple $(f_0, f_1, f_2, f_3, f_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where

$$\begin{aligned} f_0(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(c_1 z) - \log H_{M_0}(c_1 z)\Big], && \alpha_0 = 1, \\ f_1(z) &= z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(c_1 z) - \log H_{M_1}(c_1 z)\Big], && \alpha_1 = -2, \\ f_2(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(c_1 z) - \log H_{M_2}(c_1 z)\Big], && \alpha_2 = -3, \\ f_3(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(c_1 z) - \log H_{M_3}(c_1 z)\Big], && \alpha_3 = -2, \\ f_4(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(c_1 z) - \log H_{M_4}(c_1 z)\Big], && \alpha_4 = 7. \end{aligned}$$

with $c_1^2 = -\frac{1}{2}$.



$$
\begin{aligned}
M_0 &= \Xi\,(0,1,2,4,4)\\
M_1 &= \Xi\,(0,1,2,4,5)\\
M_2 &= \Xi\,(0,1,3,4,5)\\
M_3 &= \Xi\,(0,2,3,4,5)\\
M_4 &= \Xi\,(0,2,3,5,5)\\
M_5 &= \Xi\,(1,2,3,5,5) = M_0 + 1
\end{aligned}
$$

Fig. 3 A degenerate Maya 5-cycle with $k = 1$ for the choice $(n_1, n_2, n_3, n_4) = (1, 1, 2, 0)$ and permutation $\boldsymbol{\pi} = (42130)$

Example 6.16 To illustrate the existence of degenerate Maya cycles, we construct one such degenerate example belonging to the (5) signature class, by choosing $(n_1, n_2, n_3, n_4) = (1, 1, 2, 0)$. The presence of $n_4 = 0$ means that the first Maya diagram has genus 1 instead of the generic genus 2, with block coordinates given by $M_0 = \Xi\,(0, 1, 2, 4, 4)$. The canonical flip sequence $\boldsymbol{\mu} = (0, 1, 2, 4, 4)$ contains two flips at the same site, so it is not unique. Choosing the permutation (42130) produces the chain of Maya diagrams shown in Fig. 3. The explicit construction of the rational solutions follows the same steps as in the previous example, and we shall omit it here. It is worth noting, however, that due to the degenerate character of the chain, three linear combinations of $f_0, \dots, f_4$ will provide a solution to the lower rank A_2-Painlevé. If the two flips at the same site are performed consecutively in the cycle, the embedding of $A_2^{(1)}$ into $A_4^{(1)}$ is trivial and corresponds to setting two consecutive f_i to zero. This is not the case in this example, as the flip sequence is $\boldsymbol{\mu}_{\boldsymbol{\pi}} = (4, 2, 1, 4, 0)$, which produces a non-trivial embedding.

Exercise 17 Construct a $(5, 3)$-cyclic Maya diagram in the signature class $(1, 1, 3)$ with $(n_1, n_2, n_3, n_4) = (3, 1, 1, 2)$ and permutation (41230). Show the explicit form of the rational solutions to the dressing chain and A_4-Painlevé.

The first Maya diagram has 3-block coordinates $(0|3|1, 2, 4)$ and the canonical flip sequence is given by $\boldsymbol{\mu} = \Theta\,(0|3|1, 2, 4) = (0, 10, 5, 8, 14)$. The permutation (41230) gives the chain of Maya diagrams shown in Fig. 4. Note that, as in Example 16, the permutation specifies the order in which the 3-block coordinates are shifted by +1 in the subsequent steps of the cycle. This type of solutions in the signature class $(1, 1, 3)$ were not given in [17], and they are new to the best of our knowledge.

We proceed to build the explicit rational solution to the A_4-Painlevé system (6.6). In this case, the permutation $\boldsymbol{\pi} = (41230)$ on the canonical sequence $\boldsymbol{\mu} = (0, 10, 5, 8, 14)$ produces the flip sequence $\boldsymbol{\mu}_{\boldsymbol{\pi}} = (14, 10, 5, 8, 0)$, so that the values of the α_i parameters given by (5.20) become $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = (8, 10, -6, 16, -34)$. The pseudo-Wronskians corresponding to each Maya diagram in the cycle are ordinary Wronskians, which will always be the case with the

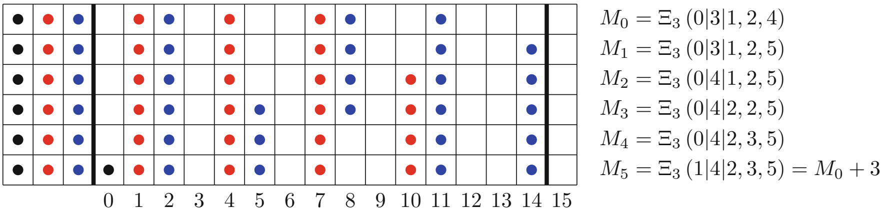


Fig. 4 A Maya 5-cycle with shift $k = 3$ for the choice $(n_1, n_2, n_3, n_4) = (3, 1, 1, 2)$ and permutation $\boldsymbol{\pi} = (41230)$

normalization imposed in Remark 6.13. They read (see Fig. 4):

$$\begin{aligned}
H_{M_0}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_7, H_8, H_{11})\\
H_{M_1}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_7, H_8, H_{11}, H_{14})\\
H_{M_2}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_7, H_8, H_{10}, H_{11}, H_{14})\\
H_{M_3}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_5, H_7, H_8, H_{10}, H_{11}, H_{14})\\
H_{M_4}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_5, H_7, H_{10}, H_{11}, H_{14})
\end{aligned}$$

where $H_n = H_n(z)$ is the n-th Hermite polynomial. The rational solution to the dressing chain is given by the tuple $(w_0, w_1, w_2, w_3, w_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where α_i and w_i are given by (5.19) and (5.20) as:

$$\begin{aligned}
w_0(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(z) - \log H_{M_0}(z)\Big], && a_0 = 8,\\
w_1(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(z) - \log H_{M_1}(z)\Big], && a_1 = 10,\\
w_2(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(z) - \log H_{M_2}(z)\Big], && a_2 = -6,\\
w_3(z) &= z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(z) - \log H_{M_3}(z)\Big], && a_3 = 16,\\
w_4(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(z) - \log H_{M_4}(z)\Big], && a_4 = -34.
\end{aligned}$$

Finally, Proposition 4.4 implies that the corresponding rational solution to the A_4-Painlevé system (4.15) is given by the tuple $(f_0, f_1, f_2, f_3, f_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where

$$\begin{aligned}
f_0(z) &= \tfrac{1}{3}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(c_2 z) - \log H_{M_0}(c_2 z)\Big], && \alpha_0 = -\tfrac{4}{3},\\
f_1(z) &= \tfrac{1}{3}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(c_2 z) - \log H_{M_1}(c_2 z)\Big], && \alpha_1 = -\tfrac{5}{3},
\end{aligned}$$

$$f_2(z) = \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(c_2 z) - \log H_{M_2}(c_2 z)\Big], \qquad \alpha_2 = 1,$$

$$f_3(z) = \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(c_2 z) - \log H_{M_3}(c_2 z)\Big], \qquad \alpha_3 = -\tfrac{8}{3},$$

$$f_4(z) = \tfrac{1}{3}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(c_2 z) - \log H_{M_4}(c_2 z)\Big], \qquad \alpha_4 = \tfrac{17}{3}.$$

with $c_2^2 = -\frac{1}{6}$.

Exercise 18 Construct a $(5,5)$-cyclic Maya diagram in the signature class $(1,1,1,1,1)$ with $(n_1, n_2, n_3, n_4) = (2,3,0,1)$ and permutation (32410). Show the explicit form of the rational solutions to the dressing chain and A_4-Painlevé.

With the above choice, the first Maya diagram o the cycle has 5-block coordinates $(0|2|3|0|1)$, and the canonical flip sequence is given by $\boldsymbol{\mu} = \Theta\,(0|2|3|0|1) = (0, 11, 17, 3, 9)$. The permutation (32410) gives the chain of Maya diagrams shown in Fig. 5. Note that, as it happens in the previous examples, the permutation specifies the order in which the 5-block coordinates are shifted by $+1$ in the subsequent steps of the cycle. This type of solutions with signature $(1,1,1,1,1)$ were already studied in [17], and they are based on a generalization of the Okamoto polynomials that appear in the solution of $\mathrm{P_{IV}}$ (A_2-Painlevé).

We proceed to build the explicit rational solution to the A_4-Painlevé system (6.6). In this case, the permutation $\boldsymbol{\pi} = (32410)$ on the canonical sequence $\boldsymbol{\mu} = (0, 11, 17, 3, 9)$ produces the flip sequence $\boldsymbol{\mu}_{\boldsymbol{\pi}} = (3, 17, 9, 11, 0)$, so that the values of the α_i parameters given by (5.20) become $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = (-28, 16, -4, 22, -16)$. The pseudo-Wronskians corresponding to each Maya diagram in the cycle are ordinary Wronskians, which will always be the case with

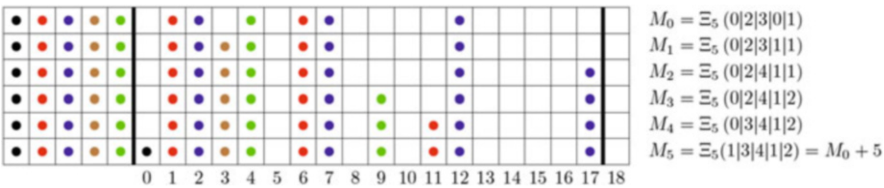


Fig. 5 A Maya 5-cycle with shift $k = 5$ for the choice $(n_1, n_2, n_3, n_4) = (2,3,0,1)$ and permutation $\boldsymbol{\pi} = (32410)$

the normalization imposed in Remark 6.13. They read:

$$
\begin{aligned}
H_{M_0}(z) &= \mathrm{Wr}(H_1, H_2, H_4, H_6, H_7, H_{12})\\
H_{M_1}(z) &= \mathrm{Wr}(H_1, H_2, H_3, H_4, H_6, H_7, H_{12})\\
H_{M_2}(z) &= \mathrm{Wr}(H_1, H_2, H_3, H_4, H_6, H_7, H_{12}, H_{17})\\
H_{M_3}(z) &= \mathrm{Wr}(H_1, H_2, H_3, H_4, H_6, H_7, H_9, H_{12}, H_{17})\\
H_{M_4}(z) &= \mathrm{Wr}(H_1, H_2, H_3, H_4, H_6, H_7, H_9, H_{11}, H_{12}, H_{17})
\end{aligned}
$$

where $H_n = H_n(z)$ is the n-th Hermite polynomial. The rational solution to the dressing chain is given by the tuple $(w_0, w_1, w_2, w_3, w_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where α_i and w_i are given by (5.19) and (5.20) as:

$$
\begin{aligned}
w_0(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(z) - \log H_{M_0}(z)\Big], && a_0 = -28\\
w_1(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(z) - \log H_{M_1}(z)\Big], && a_1 = 16,\\
w_2(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(z) - \log H_{M_2}(z)\Big], && a_2 = -4,\\
w_3(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(z) - \log H_{M_3}(z)\Big], && a_3 = 22\\
w_4(z) &= -z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(z) - \log H_{M_4}(z)\Big], && a_4 = -16.
\end{aligned}
$$

Finally, Proposition 5.16 implies that the corresponding rational solution to the A_4-Painlevé system (4.15) is given by the tuple $(f_0, f_1, f_2, f_3, f_4|\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where

$$
\begin{aligned}
f_0(z) &= \tfrac{1}{5}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_2}(c_3 z) - \log H_{M_0}(c_3 z)\Big], && \alpha_0 = \tfrac{14}{5},\\
f_1(z) &= \tfrac{1}{5}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_3}(c_3 z) - \log H_{M_1}(c_3 z)\Big], && \alpha_1 = -\tfrac{8}{5},\\
f_2(z) &= \tfrac{1}{5}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_4}(c_3 z) - \log H_{M_2}(c_3 z)\Big], && \alpha_2 = \tfrac{2}{5},\\
f_3(z) &= \tfrac{1}{5}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_0}(c_3 z) - \log H_{M_3}(c_3 z)\Big], && \alpha_3 = -\tfrac{11}{5},\\
f_4(z) &= \tfrac{1}{5}z + \frac{\mathrm{d}}{\mathrm{d}z}\Big[\log H_{M_1}(c_3 z) - \log H_{M_4}(c_3 z)\Big], && \alpha_4 = \tfrac{8}{5}.
\end{aligned}
$$

with $c_3^2 = -\frac{1}{10}$.

Acknowledgements The research of DGU has been supported in part by Spanish MINECO-FEDER Grants MTM2015-65888-C4-3 and MTM2015-72907-EXP, and by the ICMAT-Severo Ochoa project SEV-2015-0554. The research of RM was supported in part by NSERC grant RGPIN-228057-2009. DGU would like to thank the Volkswagen Stiftung and the African Institute of Mathematical Sciences for their hospitality during the Workshop on Introduction to Orthogonal Polynomials and Applications, Duala (Cameroon), where these lectures were first taught.

References

1. V. È. Adler, A modification of Crum's method. Theor. Math. Phys. **101**(3), 1381–1386 (1994)
2. V. È. Adler, Nonlinear chains and Painlevé equations. Phys. D **73**(4), 335–351 (1994)
3. G.E. Andrews, *The Theory of Partitions* (Cambridge University Press, Cambridge, 1998). MR 1634067
4. G.E. Andrews, K. Eriksson, *Integer Partitions* (Cambridge University Press, Cambridge, 2004). MR 2122332
5. D. Bermúdez, Complex SUSY transformations and the Painlevé IV equation. SIGMA **8**, 069 (2012)
6. D. Bermúdez, D.J. Fernández, Complex solutions to the Painlevé IV equation through supersymmetric quantum mechanics, in *AIP Conference Proceedings*, vol. 1420 (AIP, College Park, 2012), pp. 47–51
7. N. Bonneux, A.B.J. Kuijlaars, Exceptional Laguerre polynomials. Stud. Appl. Math. (2018). https://doi.org/10.1111/sapm.12204
8. P.A. Clarkson, Painlevé equations – nonlinear special functions. J. Comput. Appl. Math. **153**(1–2), 127–140 (2003)
9. P.A. Clarkson, The fourth Painlevé equation and associated special polynomials. J. Math. Phys. **44**(11), 5350–5374 (2003)
10. P.A. Clarkson, D. Gómez-Ullate, Y. Grandati, R. Milson, Rational solutions of higher order Painlevé systems I (2018). Preprint. arXiv: 1811.09274
11. S.Yu. Dubov, V.M. Eleonskii, N.E. Kulagin, Equidistant spectra of anharmonic oscillators. Chaos **4**(1), 47–53 (1994)
12. A.J. Durán, Exceptional Meixner and Laguerre orthogonal polynomials. J. Approx. Theory **184**, 176–208 (2014)
13. A.J. Durán, Exceptional Charlier and Hermite orthogonal polynomials. J. Approx. Theory **182**, 29–58 (2014)
14. A.J. Durán, Higher order recurrence relation for exceptional Charlier, Meixner, Hermite and Laguerre orthogonal polynomials. Integral Transforms Spec. Funct. **26**(5), 357–376 (2015)
15. A.J. Durán, Exceptional Hahn and Jacobi orthogonal polynomials. J. Approx. Theory **214**, 9–48 (2017)
16. A.J. Durán, M. Pérez, Admissibility condition for exceptional Laguerre polynomials. J. Math. Anal. Appl. **424**(2), 1042–1053 (2015)
17. G. Filipuk, P.A. Clarkson, The symmetric fourth Painlevé hierarchy and associated special polynomials. Stud. Appl. Math. **121**(2), 157–188 (2008)
18. P.J. Forrester, N.S. Witte, Application of the τ-function theory of Painlevé equations to random matrices: PIV, PII and the GUE. Commun. Math. Phys. **219**(2), 357–398 (2001). MR 1833807
19. M. García-Ferrero, D. Gómez-Ullate, Oscillation theorems for the Wronskian of an arbitrary sequence of eigenfunctions of Schrödinger's equation. Lett. Math. Phys. **105**(4), 551–573 (2015)
20. M. García-Ferrero, D. Gómez-Ullate, R. Milson, A Bochner type characterization theorem for exceptional orthogonal polynomials. J. Math. Anal. Appl. **472**, 584–626 (2019)
21. D. Gómez-Ullate, N. Kamran, R. Milson, Supersymmetry and algebraic Darboux transformations. J. Phys. A **37**(43), 10065 (2004)

22. D. Gómez-Ullate, N. Kamran, R. Milson, The Darboux transformation and algebraic deformations of shape-invariant potentials. J. Phys. A **37**(5), 1789 (2004)
23. D. Gómez-Ullate, N. Kamran, R. Milson, An extended class of orthogonal polynomials defined by a Sturm–Liouville problem. J. Math. Anal. Appl. **359**(1), 352–367 (2009)
24. D. Gómez-Ullate, N. Kamran, R. Milson, An extension of Bochner's problem: exceptional invariant subspaces. J. Approx. Theory **162**(5), 987–1006 (2010)
25. D. Gómez-Ullate, N. Kamran, R. Milson, Two-step Darboux transformations and exceptional Laguerre polynomials. J. Math. Anal. Appl. **387**(1), 410–418 (2012)
26. D. Gómez-Ullate, Y. Grandati, R. Milson, Rational extensions of the quantum harmonic oscillator and exceptional Hermite polynomials. J. Phys. A **47**(1), 015203 (2013)
27. D. Gómez-Ullate, N. Kamran, R. Milson, A conjecture on exceptional orthogonal polynomials. Found. Comput. Math. **13**(4), 615–666 (2013)
28. D. Gómez-Ullate, Y. Grandati, R. Milson, Shape invariance and equivalence relations for pseudo-Wronskians of Laguerre and Jacobi polynomials. J. Phys. A **51**(34), 345201 (2018)
29. D. Gómez-Ullate, Y. Grandati, R. Milson, Durfee rectangles and pseudo-Wronskian equivalences for Hermite polynomials. Stud. Appl. Math. **141**(4), 596–625 (2018)
30. D. Gómez-Ullate, Y. Grandati, S. Lombardo, R. Milson, Rational solutions of dressing chains and higher order Painleve equations (2018). Preprint. arXiv:1811.10186
31. Y. Grandati, Solvable rational extensions of the isotonic oscillator. Ann. Phys. **326**(8), 2074–2090 (2011)
32. Y. Grandati, Multistep DBT and regular rational extensions of the isotonic oscillator. Ann. Phys. **327**(10), 2411–2431 (2012)
33. V.I. Gromak, I. Laine, S. Shimomura, *Painlevé Differential Equations in the Complex Plane*, vol. 28 (Walter de Gruyter, Berlin, 2008)
34. K. Kajiwara, Y. Ohta, Determinant structure of the rational solutions for the Painlevé II equation. J. Math. Phys. **37**(9), 4693–4704 (1996)
35. K. Kajiwara, Y. Ohta, Determinant structure of the rational solutions for the Painlevé IV equation. J. Phys. A **31**(10), 2431 (1998)
36. M.G. Krein, On a continual analogue of a Christoffel formula from the theory of orthogonal polynomials. Dokl. Akad. Nauk SSSR (N.S.) **113**, 970–973 (1957). MR 0091396
37. A.B.J. Kuijlaars, R. Milson, Zeros of exceptional Hermite polynomials. J. Approx. Theory **200**, 28–39 (2015)
38. I. Marquette, C. Quesne, New ladder operators for a rational extension of the harmonic oscillator and superintegrability of some two-dimensional systems. J. Math. Phys. **54**(10), 12 pp., 102102 (2013). MR 3134580
39. I. Marquette, C. Quesne, Two-step rational extensions of the harmonic oscillator: exceptional orthogonal polynomials and ladder operators. J. Phys. A **46**(15), 155201 (2013)
40. I. Marquette, C. Quesne, Connection between quantum systems involving the fourth Painlevé transcendent and k-step rational extensions of the harmonic oscillator related to Hermite exceptional orthogonal polynomial. J. Math. Phys. **57**(5), 15, 052101 (2016). MR 3501792
41. D. Masoero, P. Roffelsen, Poles of Painlevé IV rationals and their distribution. SIGMA **14** (2018), 49, Paper No. 002. MR 3742702
42. T. Masuda, Y. Ohta, K. Kajiwara, A determinant formula for a class of rational solutions of Painlevé V equation. Nagoya Math. J. **168**, 1–25 (2002)
43. J. Mateo, J. Negro, Third-order differential ladder operators and supersymmetric quantum mechanics. J. Phys. A **41**(4), 28, 045204 (2008). MR 2451071
44. K. Matsuda, Rational solutions of the Noumi and Yamada system of type $A_4^{(1)}$. J. Math. Phys. **53**(2), 023504 (2012)
45. Monty Python, And now for something completely different. https://www.imdb.com/title/tt0066765/
46. M. Noumi, *Painlevé Equations through Symmetry*, vol. 223 (Springer Science & Business, New York, 2004)
47. M. Noumi, Y. Yamada, Symmetries in the fourth Painlevé equation and Okamoto polynomials. Nagoya Math. J. **153**, 53–86 (1999)

48. A.A. Oblomkov, Monodromy-free Schrödinger operators with quadratically increasing potentials. Theor. Math. Phys. **121**(3), 1574–1584 (1999)
49. S. Odake, R. Sasaki, Infinitely many shape invariant potentials and new orthogonal polynomials. Phys. Lett. B **679**(4), 414–417 (2009)
50. S. Odake, R. Sasaki, Another set of infinitely many exceptional X_ℓ Laguerre polynomials. Phys. Lett. B **684**, 173–176 (2010)
51. S. Odake, R. Sasaki, Exactly solvable quantum mechanics and infinite families of multi-indexed orthogonal polynomials. Phys. Lett. B **702**(2–3), 164–170 (2011)
52. S. Odake, R. Sasaki, Extensions of solvable potentials with finitely many discrete eigenstates. J. Phys. A **46**(23), 235205 (2013)
53. S. Odake, R. Sasaki, Krein–Adler transformations for shape-invariant potentials and pseudo virtual states. J. Phys. A **46**(24), 245201 (2013)
54. K. Okamoto, Studies on the Painlevé equations. III. Second and fourth Painlevé equations, P_{II} and P_{IV}. Math. Ann. **275**(2), 221–255 (1986). MR 854008
55. J.B. Olsson, *Combinatorics and Representations of Finite Groups*. Fachbereich Mathematik [Lecture Notes in Mathematics], vol. 20 (Universität Essen, Essen, 1994)
56. C. Quesne, Exceptional orthogonal polynomials, exactly solvable potentials and supersymmetry. J. Phys. A Math. Theor. **41**(39), 392001 (2008)
57. A. Sen, A.N.W. Hone, P.A. Clarkson, Darboux transformations and the symmetric fourth Painlevé equation. J. Phys. A **38**(45), 9751–9764 (2005)
58. K. Takasaki, Spectral curve, Darboux coordinates and Hamiltonian structure of periodic dressing chains. Commun. Math. Phys. **241**(1), 111–142 (2003)
59. T. Tsuda, Universal characters, integrable chains and the Painlevé equations. Adv. Math. **197**(2), 587–606 (2005)
60. H. Umemura, Painlevé equations in the past 100 years. Am. Math. Soc. Transl. **204**, 81–110 (2001)
61. W. Van Assche, *Orthogonal Polynomials and Painlevé Equations*. Australian Mathematical Society Lecture Series, vol. 27 (Cambridge University Press, Cambridge, 2018). MR 3729446
62. A.P. Veselov, A.B. Shabat, Dressing chains and the spectral theory of the Schrödinger operator. Funct. Anal. Appl. **27**(2), 81–96 (1993)
63. R. Willox, J. Hietarinta, Painlevé equations from Darboux chains. I. P_{III}–P_{V}. J. Phys. A **36**(42), 10615–10635 (2003). MR 2024916

$(\mathcal{R}, p, q)$-Rogers–Szegö and Hermite Polynomials, and Induced Deformed Quantum Algebras

Mahouton Norbert Hounkonnou

Abstract Deformed quantum algebras, namely the q-deformed algebras and their extensions to (p, q)-deformed algebras, continue to attract much attention. One of the main reasons is that these topics represent a meeting point of nowadays fast developing areas in mathematics and physics like the theory of quantum orthogonal polynomials and special functions, quantum groups, integrable systems, quantum and conformal field theories and statistics.

This contribution paper aims at characterizing the $(\mathcal{R}, p, q)$-Rogers–Szegö polynomials, and the $(\mathcal{R}, p, q)$-deformed difference equation giving rise to raising and lowering operators. These polynomials induce some realizations of generalized deformed quantum algebras, (called $(\mathcal{R}, p, q)$-deformed quantum algebras), which are here explicitly constructed. The study of continuous $(\mathcal{R}, p, q)$-Hermite polynomials is also performed. Known particular cases are recovered.

Keywords Quantum algebras · $(\mathcal{R}, p, q)$-deformed quantum algebras · Orthogonal polynomials · Rogers–Szegö polynomials · Hermite polynomials

Mathematics Subject Classification (2000) Primary 33C45; Secondary 20G42

1 Introduction

Rogers–Szegö polynomials and their generalizations have been attracted a great attention since the end of the nineteenth century, motivated by their importance in the description of physical phenomena. Indeed, they appear as possible solutions of

M. N. Hounkonnou (✉)
International Chair in Mathematical Physics and Applications (ICMPA-UNESCO Chair), University of Abomey-Calavi, Cotonou, Republic of Benin
e-mail: norbert.hounkonnou@cipma.uac.bj

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_16

the q-harmonic oscillator, and can be used as different bases states for describing physical systems for which the free parameter (the deformation parameter of the algebra), ranging from 0 to 1, can account, for instance, for squeezing effects. Furthermore, these polynomials are orthonormalized on the circle, and can therefore be used in connection with angular representations of the harmonic oscillator. Moreover, the q-oscillator algebra plays a central role in the physical applications of quantum groups. Deformed quantum algebras, namely the q-deformed algebras [21, 24, 25] and their extensions to (p, q)-deformed algebras [5, 6], also continue to attract much attention. One of the main reasons is that these topics represent a meeting point of nowadays fast developing areas in mathematics and physics like the theory of quantum orthogonal polynomials and special functions, quantum groups, integrable systems and quantum and conformal field theories and statistics. Indeed, since the work of Jimbo [21], these fields have known profound interesting developments which can be partially found, for instance, in the books by Chari and Pressley [7], Klimyk and Schumudgen [22], Ismail Mourad [16] and in references therein.

The two-parameter quantum algebra, $Up, q(gl(2))$, was first introduced in Ref. [6] in view to generalize or/and unify a series of q-oscillator algebra variants, known in the earlier physics and mathematics literature on the representation theory of single-parameter quantum algebras. Then flourish investigations in the same direction, among which the work of Burban and Klimyk [5] on representations of two-parameter quantum groups and models of two parameter quantum algebra $U_{p,q}(su_{1,1})$ and (p, q)-deformed oscillator algebra. Almost simultaneously, Gelfand et al.[12] introduced the (r, s)-hypergeometric series satisfying two parameter difference equation, including r-and s-shift operators. These new series reproduces the Burban and Klimyk's P, Q-hypergeometric functions. The (p, q)-deformation rapidly found applications in physics and mathematical physics as described for instance in [9, 13, 15].

Upon recalling a technique of constructing explicit realizations of raising and lowering operators that satisfy an algebra akin to the usual harmonic oscillator algebra, through the use of the three-term recursion relation and the differentiation expression of Hermite polynomials, Galetti [10] has shown that a similar procedure can be carried out in the case of the three-term recursion relation for Rogers–Szegö and Stieltjes–Wigert polynomials and the Jackson q-derivative [19, 20]. This technique furnished new realizations of the q-deformed algebra associated with the q-deformed harmonic oscillator, which obey, well known and spread in the literature, commutation relations.

In the same vein, after recalling the connection between the Rogers–Szegö polynomials and the q-oscillator, Jagannathan and Sridhar [17] have defined a (p, q)-Rogers–Szegö polynomials, shown that they are connected with the (p, q)-deformed oscillator associated with the Jagannathan-Srinavasa (p, q)-numbers [18] and proposed a new realization of this algebra. In a previous paper [14], we have proposed a theoretical framework for the (p, q)-deformed states' generalization and provided a generalized deformed quantum algebra, based on a work by Odzijewicz [24] on a generalization of q-deformed states in which the realizations of creation and annihilation operators are given by multiplication by z and the action of the deformed derivative $\partial_{\mathcal{R},p,q}$ on the space of analytic functions defined on the disc.

The present contribution, after recalling the known main results, published in the literature, on classical Rogers–Szegö and Hermite polynomials, resumes the works on their generalizations performed in our group during the recent years [4]. Specifically, we develop a realization of our generalized deformed quantum algebras, and give an explicit definition of the $(\mathcal{R}, p, q)$-Rogers–Szegö and Hermite polynomials, together with their three-term recursion relation and the deformed difference equation giving rise to the creation and annihilation operators.

The paper is organized as follows. In Sect. 2, we present a brief review of known results on deformed numbers, deformed binomial coefficients, Hermite polynomials and Rogers-Szeg*ö*. The calculus pertaining to the definition and the computation of the three-term recursion relation of polynomials are exposed. In Sect. 3, after recalling the procedure of constructing realizations of the harmonic and q-deformed harmonic oscillators, we proceed to realizations of $(\mathcal{R}, p, q)$-deformed quantum algebras through the $(\mathcal{R}, p, q)$-difference equation and the three-term recursion relation between $(\mathcal{R}, p, q)$-Rogers–Szegö polynomials. The key result of this section is Theorem 3.3 giving the method of computation of relevant quantities. Section 4 is devoted to the study of the particular cases. Section 5 is dedicated to the study of the continuous $(\mathcal{R}, p, q)$-Hermite polynomials. We then give their definition and recursion relation. Finally, Sect. 6 ends with the concluding remarks.

2 Preliminaries

Let us recall some concepts related to the theory of q-series.

Definition 2.1 Let $q \in [0, 1)$. Then, for a non negative integer n, the q-deformed number is defined as follows:

$$[n]_q := \frac{1-q^n}{1-q} = 1 + q + \ldots + q^{n-1}, \text{ with } [0]_q := 0,$$

which is called a q-number, or basic number.

Proposition 2.2 *Let n and m be two nonnegative integers. Then, the following relations hold:*

$$[-n]_q = -q^{-n}[n]_q$$

and

$$[n+m]_q = [n]_q + q^n[m]_q$$

Proof In fact

$$[-n]_q := \frac{1-q^{-n}}{1-q} = \frac{\frac{q^n-1}{q^n}}{1-q} = -q^n[n]_q,$$

$$\begin{aligned}[n+m]_q &:= \frac{1-q^{n+m}}{1-q} \\ &= \frac{1-q^n+q^n-q^{n+m}}{1-q} \\ &= \frac{1-q^n}{1-q} + q^n\left(\frac{1-q^m}{1-q}\right) \\ &= [n]_q + q^n[m]_q.\end{aligned}$$

□

Definition 2.3 The q-deformed factorial number is defined as

$$[n]_q! := \prod_{\kappa=1}^{n} [n]_q,$$

where n is an integer.

Definition 2.4 The q-Pochhammer symbol, also called q-shifted factorial, is defined by

$$(z;q)_n := \prod_{j=0}^{n-1}\left(1-zq^j\right), \ n = 1, 2, \ldots \tag{2.1}$$

and

$$(z;q)_\infty := \prod_{j=0}^{\infty}\left(1-zq^j\right).$$

with

$$(z;q)_0 := 1.$$

Furthermore,

$$(z;q)_n = \frac{(z;q)_\infty}{(zq^n;q)_\infty}.$$

The multiple q-shifted factorial is expressed as:

$$(z_1, z_2, \ldots, z_m; q)_n := \prod_{j=1}^{m} \left(z_j; q\right)_n, \; n = 1, 2, \ldots$$

and

$$(z_1, z_2, \ldots, z_m; q)_\infty := \prod_{j=1}^{m} \left(z_j; q\right)_\infty .$$

Definition 2.5 The q-binomial coefficients are defined by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q = \frac{(q; q)_n}{(q; q)_k (q; q)_{n-k}} = \frac{\prod_{j=0}^{k-1}\left(1 - q^{n-j}\right)}{\prod_{j=0}^{k-1}\left(1 - q^{k-j}\right)}.$$

The q-Pochhammer symbol can also be defined as follows:

$$(z, q)_n := \sum_{j=0}^{n} (-1)^j \begin{bmatrix} n \\ j \end{bmatrix}_q q^{\frac{j(j-1)}{2}} z^j. \tag{2.2}$$

It is worth mentioning some useful formulas for the q-shifted factorials, given in the following relations. The reader can refer to [11, 23] for more details.

$$\begin{aligned}
(z; q)_{n+k} &= (z; q)_n (zq^n; q)_k, \\
\frac{(zq^n; q)_k}{(zq^k; q)_n} &= \frac{(z; q)_k}{(z; q)_n}, \\
(zq^k; q)_{n-k} &= \frac{(z; q)_n}{(z; q)_k}, \quad k = 0, 1, 2, \cdots, n, \\
(z; q)_n &= (z^{-1} q^{1-n}; q)_n (-z)^n q^{\binom{n}{2}}, \quad q \neq 0, \\
(zq^{-n}; q)_n &= (z^{-1} q; q)_n (-z)^n q^{-n-\binom{n}{2}}, \quad z \neq 0, \\
(z; q)_{n-k} &= \frac{(z; q)_n}{(z^{-1} q^{1-n}; q)_k} \left(-\frac{q}{z}\right)^k q^{\binom{k}{2} - nk}, \quad z \neq 0, \; k \in \mathbb{N}, \\
(q^{-n}; q)_k &= \frac{(q; q)_n}{(q; q)_{n-k}} (-1)^k q^{\binom{k}{2} - nk}, \quad k \in \mathbb{N} \\
(z; q)_{2n} &= (z, q^2)_n (zq; q^2)_n, \\
(z^2; q^2)_{2n} &= (z, q)_n (-z; q)_n,
\end{aligned}$$

$$(z;q)_\infty = (z,q^2)_\infty (zq;q^2)_\infty, \quad 0<|q|<1,$$
$$(z^2;q^2)_\infty = (z,q)_\infty(-z;q)_\infty, \quad 0<|q|<1.$$

Proposition 2.6 *We have*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{1-q^n}{1-q^k}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q .$$

Proof Indeed,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q![n-k]_q!} = \frac{[n]_q[n-1]_q!}{[k]_q[k-1]_q![n-k]_q!} \Leftrightarrow [k]_q\begin{bmatrix} n \\ k \end{bmatrix}_q = [n]_q\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q$$
$$\Leftrightarrow (1-q^k)\begin{bmatrix} n \\ k \end{bmatrix}_q = (1-q^n)\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \Leftrightarrow \begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{1-q^n}{1-q^k}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q .$$

□

Definition 2.7 Let n and q be real numbers such that $0 < q < 1$. Then, the q-factorial of n of order k is defined by

$$[n]_{k,q} = [n]_q[n-1]_q\cdots[n-k+1]_q, \quad k\in\mathbb{N}\backslash\{0\}. \tag{2.3}$$

Proposition 2.8 *The q-binomial coefficients satisfy the following two recursion relations:*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \tag{2.4}$$

and

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^k\begin{bmatrix} n-1 \\ k \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \tag{2.5}$$

Proof From the relation (2.3), we get

$$[n]_{k,q} = [n]_q[n-1]_{k-1,q}$$

and using $[n]_q = [n-k]_q + q^{n-k}[k]$, we obtain the recurrence relation

$$[n]_{k,q} = [n-1]_{k,q} + q^{n-k}[k]_q[n-1]_{k-1,q}$$

with initial condition $[n]_{0,q} = 1$. Dividing both members of the above equation by $[k]_q!$ and using

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_{k,q}}{[k]_q!}$$

the relation (2.4) is deduced. Similarly, using

$$[n]_q = [k]_q + q^k\,[n-k]_q$$

we obtain Eq. (2.5). □

Proposition 2.9 ([26]) *If A and B are two q-commuting linear operators on $\mathbb{C}[x, s]$, i.e. satisfying $BA = qAB$, then*

$$(A + B)^n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q A^k\, B^{n-k}. \tag{2.6}$$

Proof By induction on n. For $n = 1$, we get

$$\begin{aligned} (A + B)^1 &= A + B \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}_q A^0\, B^1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix}_q A^1\, B^0 \\ &= \sum_{k=0}^{1} \begin{bmatrix} 1 \\ k \end{bmatrix}_q A^k\, B^{1-k}. \end{aligned}$$

Thus, the relation (2.6) holds for $n = 1$.

We assume that the relation (2.6) holds for $n \leq m$, and let us prove it for $n = m + 1$. We have

$$\begin{aligned} (A + B)^{m+1} &= (A + B)\,(A + B)^m \\ &= (A + B) \sum_k \begin{bmatrix} m \\ k \end{bmatrix}_q A^k B^{m-k} \\ &= \sum_k \left(\begin{bmatrix} m \\ k-1 \end{bmatrix}_q + q^k \begin{bmatrix} m \\ k \end{bmatrix}_q \right) A^k B^{m-k} \\ &= \sum_k \begin{bmatrix} m+1 \\ k \end{bmatrix}_q A^k B^{m+1-k}. \end{aligned}$$

□

The q-exponential series can be expressed as

$$\exp_q(z) := \sum_{n\geq 0} \frac{z^n}{[n]_q!}.$$

Definition 2.10 The q-Jackson derivative ∂_q is defined by:

$$(\partial_q f)(z) := \frac{f(z) - f(qz)}{(1-q)z}, \quad z \neq 0$$

and

$$(\partial_q f)(0) = f'(0)$$

provided that f is differentiable at 0.

Similarly, we define the q^{-1}-derivative as

$$(\partial_{q^{-1}} f)(z) := \frac{f(z) - f(q^{-1}z)}{(1-q^{-1})z}.$$

The limit as q approaches 1 is the usual derivative

$$\lim_{q\to 1}(\partial_q f)(z) = \frac{df(z)}{dz}$$

if f is differentiable at z.

For $f(z) = z^n$, we have

$$\partial_q z^n = \frac{z^n - q^n z^n}{(1-q)z} = [n]_q \, z^{n-1}.$$

The formulae for the q-difference of a sum, a product and a quotient of functions f and g are given by

$$\partial_q(f(z) + g(z)) = \partial_q f(z) + \partial_q g(z),$$

$$\partial_q(f(z)g(z)) = g(z)\partial_q f(z) + f(qz)\partial_q g(z),$$

and

$$\partial_q\left(\frac{f(z)}{g(z)}\right) = \frac{g(z)\partial_q f(z) - f(z)\partial_q g(z)}{g(z)g(qz)}, \quad g(z)g(qz) \neq 0.$$

2.1 Hermite Polynomials

Let us briefly recall the properties of the Hermite polynomials. From the theory of classical polynomials, the Hermite polynomials $\mathbb{H}_n$ obey a three-term recursion relation:

$$\mathbb{H}_{n+1}(z) = 2z\,\mathbb{H}_n(z) - 2n\,\mathbb{H}_{n-1}(z),\ \ \mathbb{H}_0(z) := 1, \tag{2.7}$$

with a differentiation relation

$$\partial_z \mathbb{H}_n(z) = 2n\,\mathbb{H}_{n-1}(z). \tag{2.8}$$

Introducing the relation (2.8) into (2.7), we get

$$\mathbb{H}_{n+1}(z) = \Big(2z - \partial_z\Big)\mathbb{H}_n(z)$$

from which we deduce the raising operator

$$\mathfrak{R} = 2z - \partial_z$$

such that the set of Hermite polynomials can be generated by successive application of this operator to the first polynomial, viz., $\mathbb{H}_0(z) = 1$,

$$\mathfrak{R}^n\,\mathbb{H}_0(z) = \mathbb{H}_n(z),$$

which can immediately be checked by induction. Indeed, we consider

$$\mathfrak{R}\,\mathbb{H}_p(z) = \mathbb{H}_{p+1}(z).$$

For $p = 0$,

$$\mathfrak{R}\,\mathbb{H}_0(z) = \mathbb{H}_1(z)$$

For $p = 1$,

$$\mathfrak{R}\,\mathbb{H}_1(z) = \mathbb{H}_2(z).$$

For $p = 2$,

$$\mathfrak{R}\,\mathbb{H}_2(z) = \mathbb{H}_3(z).$$

For $p = n - 1$,

$$\mathfrak{R}\,\mathbb{H}_{n-1}(z) = \mathbb{H}_n(z).$$

Multiplying member by member, we get

$$\mathfrak{R}^n\, \mathbb{H}_0(z) = \mathbb{H}_n(z).$$

From the relation (2.8), we derive the lowering operator $\mathfrak{L}$ acting as

$$\frac{1}{2}\partial_z\, \mathbb{H}_n(z) := \mathfrak{L}\, \mathbb{H}_n(z) = n\, \mathbb{H}_{n-1}(z).$$

These operators satisfy the canonical commutation relation

$$[\mathfrak{L}, \mathfrak{R}] = 1.$$

In fact,

$$\begin{aligned}[\mathfrak{L}, \mathfrak{R}]\mathbb{H}_n(z) &= \mathfrak{L}\mathfrak{R}\mathbb{H}_n(z) - \mathfrak{R}\mathfrak{L}\mathbb{H}_n(z) = \frac{1}{2}\partial_z\mathbb{H}_{n+1}(z) - \mathfrak{R}n\mathbb{H}_{n-1}(z) \\ &= (n+1)\mathbb{H}_n(z) - n\mathbb{H}_n(z) = \mathbb{H}_n(z).\end{aligned}$$

The number operator is expressed under the form

$$N = \mathfrak{R}\, \mathfrak{L}$$

such that

$$N\, \mathbb{H}_n(z) = \mathfrak{R}\, \mathfrak{L}\, \mathbb{H}_n(z) = n\mathbb{H}_n(z).$$

From the above equation, we deduce

$$\Big[(2z - \partial_z)(\frac{1}{2}\partial_z) - n\Big]\mathbb{H}_n(z) = 0,$$

or, equivalently,

$$\big[\partial_z^2 - 2z\, \partial_z + 2n\big]\mathbb{H}_n(z) = 0,$$

which is the form of the second order differential equation for the Hermite polynomials. The operators N, $\mathfrak{R}$ and $\mathfrak{L}$ obey the standard commutation relations

$$[N, \mathfrak{R}] = \mathfrak{R}$$

and

$$[N, \mathfrak{L}] = -\mathfrak{L}.$$

In fact,

$$[N, \mathfrak{R}] = [\mathfrak{R}\mathfrak{L}, \mathfrak{R}] = \mathfrak{R}[\mathfrak{L}, \mathfrak{R}] = \mathfrak{R}$$

and

$$[N, \mathfrak{L}] = [\mathfrak{R}\mathfrak{L}, \mathfrak{L}] = -[\mathfrak{R}, \mathfrak{L}]\mathfrak{L} = -\mathfrak{L}.$$

Thus, we obtained the raising, lowering and number operators from the two basic relations, (namely, the three-term recursion relation and the differentiation relation), satisfied by the Hermite polynomials, such that these operators satisfy the above commutation relations.

On the other hand, if one considers the usual Fock Hilbert space spanned by the vectors $|n\rangle$, generated from the vacuum state $|0\rangle$ by the raising operator $\mathfrak{R}$, then together with the lowering operator $\mathfrak{L}$, the following relations hold:

$$\mathfrak{L}\,\mathfrak{R} - \mathfrak{R}\,\mathfrak{L} = 1, \quad \langle 0|0\rangle = 1, \quad |n\rangle = \mathfrak{R}^n|0\rangle, \quad \mathfrak{L}|0\rangle = 0.$$

In particular, the next expressions, established using the previous equations, are in order:

$$\mathfrak{R}|n\rangle = |n+1\rangle, \quad \mathfrak{L}|n\rangle = |n-1\rangle, \quad \langle m|n\rangle = n!\delta_{mn}.$$

Let us now mention another relevant construction of the above operators for the harmonic oscillator, very popular in the literature [17], which is also useful for our development in the sequel. We consider the sequence of polynomials $\{\varphi_n(z)\ |n = 0, 1, 2, \cdots\}$,

$$\varphi_n(z) = \frac{1}{\sqrt{n!}}\mathbf{h}_n(z),$$

where

$$\mathbf{h}_n(z) = (1+z)^n = \sum_{k=0}^{n} \binom{n}{k} z^k,$$

obeying the relations

$$\frac{d}{dz}\varphi_n(z) = \sqrt{n}\varphi_{n-1}(z),$$

$$(1+z)\varphi_n(z) = \sqrt{n+1}\varphi_{n+1}(z), \tag{2.9}$$

$$(1+z)\frac{d}{dz}\varphi_n(z) = n\varphi_n(z), \tag{2.10}$$

$$\frac{d}{dz}\left((1+z)\varphi_n(z)\right) = (n+1)\varphi_n(z).$$

Here Eqs. (2.9) and (2.10) are the recursion relation and the differential equation for polynomials $\varphi_n(z)$, respectively. Thus, the set $\{\varphi_n(z)\ |n = 0, 1, 2, \cdots\}$ forms a basis for the following Bargman–Fock realization of the harmonic oscillator:

$$\mathfrak{R} = (1 + z), \qquad \mathfrak{L} = \frac{d}{dz}, \qquad N = (1 + z)\frac{d}{dz}.$$

2.2 The Rogers–Szegö Polynomials

The Rogers–Szegö polynomials are defined as [1, 27]

$$H_n(z; q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q z^k \qquad n = 0, 1, 2, \ldots \tag{2.11}$$

Mutltiplying this by $t^n/(q; q)_n$, and summing over n lead to the generating function

$$\sum_{n=1}^{\infty} H_n(z; q)\frac{t^n}{(q; q)_n} = \frac{1}{(t, tz; q)_\infty}.$$

The Rogers–Szegö polynomials satisfy a three-term recursion relation

$$H_{n+1}(z; q) = (1 + z)H_n(z; q) - z(1 - q^n)H_{n-1}(z; q) \tag{2.12}$$

as well as the q-difference equation

$$\partial_q H_n(z; q) = \frac{H_n(z; q) - H_n(qz; q)}{(1 - q)z} = \frac{\Delta H_n(z; q)}{(1 - q)z} = [n]_q H_{n-1}(z; q), \tag{2.13}$$

since the operator ∂_q acts nicely on the $H_n's$. This shows that ∂_q acts as a lowering operator.

In the limit case $q \to 1$, the Rogers–Szegö polynomial of degree n ($n = 0, 1, 2, \ldots$) well converges to

$$h_n(z) = \sum_{k=0}^{n} \binom{n}{k} z^n$$

as required.

Another important property of the Rogers–Szegö polynomials is their orthogonality on the circle [10], when the Jacobi $\omega(y; q)$ function is taken as the measure function [28]. In order to explicitly verify this, we should perform a proper choice for the variable y, $y = -q^{-\frac{1}{2}}e^{i\theta}$ such that

$$H_n(y;q) = H_n\left(-q^{-\frac{1}{2}}e^{i\theta};q\right).$$

In this form, the orthonormalization integral is written as

$$I_{mn}(q) = \int_{-\pi}^{\pi} H_m\left(-q^{-\frac{1}{2}}e^{i\theta};q\right) H_n\left(-q^{-\frac{1}{2}}e^{-i\theta};q\right)\omega(\theta;q)\frac{d\theta}{2\pi}$$

with

$$\omega(\theta;q) = \sum_{-\infty}^{\infty} q^{\frac{m^2}{2}}e^{im\theta} = \sum_{-\infty}^{\infty} e^{-\mu m^2 + im\theta},$$

where $\mu = -\ln(q)/2$, which is the measure function. Using the definition of the Rogers–Szegö polynomials, we see that

$$I_{mn}(q) = \sum_{r=0}^{m}\sum_{s=0}^{n}(-1)^{r+s}\begin{bmatrix} m \\ r \end{bmatrix}_q \begin{bmatrix} n \\ s \end{bmatrix}_q q^{\frac{r(r+1)}{2}} q^{\frac{s(s+1)}{2}} q^{-rs}.$$

Exploiting (2.1) and (2.2), this becomes

$$I_{mn}(q) = \sum_{r=0}^{m}(-1)^r\begin{bmatrix} m \\ r \end{bmatrix}_q q^{\frac{r(r+1)}{2}}\prod_{s=0}^{n-1}\left(1 - q^{s-r}\right). \tag{2.14}$$

Now, without any loss of generality, we can assume that $m \le n$ (the inverse could also be considered). There are two situations to be discussed:

1. For $m < n$, it is evident that the product on the rhs of (2.14) will vanish for all r (the rhs is constituted of a sum of products. Each summand has a product of terms where one of them will give $(1 - q^{r-r}) = 0$, since, as $m < n$, s will necessarily assume the value r). Therefore, the sum only has vanishing summands, since there will always be a zero factor in the products.

   Thus

$$I_{mn} = 0 \quad for \quad m < n. \tag{2.15}$$

2. For $m = n$, there will be only one term to be considered, namely $r = m$, that will give

$$I_{nn}(q) = (-1)^n q^{\frac{n(n+1)}{2}}\prod_{s=0}^{n-1}\left(1 - q^{s-n}\right).$$

To calculate this expression, let us explicitly write the product

$$I_{nn}(q) = (-1)^n q^{\frac{n(n+1)}{2}}\left(1-\frac{1}{q^n}\right)\left(1-\frac{1}{q^{n-1}}\right)\dots\left(1-\frac{1}{q}\right)$$
$$= \frac{(1-q^n)(1-q^{n-1})\dots(1-q)}{q^n}$$

which, upon identifying the numerator, gives

$$I_{nn}(q) = \frac{(q;q)_n}{q^n}.$$

This contribution together with (2.15) gives the final result

$$I_{mn} = q^{-n}(q;q)_n\delta_{nm}.$$

2.3 Second Order q-Differential Equation for Rogers–Szegö Polynomials

Let us derived the second order q-differential equation obeyed by the Rogers–Szegö polynomials [10].

The q-generalization of the harmonic oscillator algebra is introduced in [2, 3]. Let us consider the Fock-Hilbert space $\mathcal{H}_q$, where q-is a parameter. This space is spanned by the vectors $|n\rangle$, which are generated from the vacuum $|0\rangle$ by the action of the raising operator $\mathfrak{R}$.

We have the relation

$$\mathfrak{L}\,\mathfrak{R} - q\,\mathfrak{R}\,\mathfrak{L} = 1. \tag{2.16}$$

Feinsilver [8] gave an equivalent of the relation (2.16),

$$\mathfrak{L}\,\mathfrak{R} - \mathfrak{R}\,\mathfrak{L} = q^N,$$

where N is a usual number operator. The following relations hold:

$$\langle 0|0\rangle = 1, \quad |n\rangle = \mathfrak{R}^n|0\rangle, \quad \mathfrak{L}|0\rangle.$$

$$\mathfrak{R}|n\rangle = |n+1\rangle, \quad \mathfrak{L}|n\rangle = [n]|n-1\rangle, \quad \langle m|n\rangle = [n]\delta_{mn}.$$

The vectors $\frac{1}{[n]!}|n-1\rangle$ form an orthonormal basis set, and the Fock-Hilbert space $\mathcal{H}_q$ consists of all vectors $|n\rangle = \sum_{n=0}^{\infty} u_n |n\rangle$, where u_n is complex such that

$$\langle m|n\rangle = \sum_{n=0}^{\infty} |u_n|^2 [n]!$$

is finite.

Let us now obtain the corresponding raising, lowering and q-number operators, and look for the corresponding commutation relations. Define the difference operator by

$$\Delta H_n(z,q) = H_n(z,q) - H_n(qz,q).$$

From the direct use of the relation (2.11), we obtain

$$\Delta H_n(z,q) = (1-q^n)\, z\, \Delta H_{n-1}(z,q). \tag{2.17}$$

In fact, since the q-derivative is given by

$$\partial_q H_n(z,q) = [n]_q\, H_{n-1}(z,q),$$

or, equivalently, by

$$\frac{\Delta H_n(z,q)}{z(1-q)} = [n]_q\, H_{n-1}(z,q),$$

hence, one can deduce

$$\Delta H_n(z,q) = (1-q^n)\, z\, \Delta H_{n-1}(z,q).$$

If we substitute the above expression for $H_{n-1}(z;q)$ in (2.12), we end up with

$$H_{n+1}(z,q) = \big[(1+z) - \Delta\big] H_n(z,q). \tag{2.18}$$

Indeed

$$\begin{aligned} H_{n+1}(z;q) &= (1+z)H_n(z;q) - z(1-q^n)H_{n-1}(z;q) \\ &= (1+z)H_n(z;q) - \Delta H_n(z,q) \\ &= \big[(1+z) - \Delta\big] H_n(z,q). \end{aligned}$$

The above relation can be understood as the action of a raising operator $\mathfrak{R}$ for the Rogers–Szegö polynomials. Setting

$$A_+ := (1+z) - \Delta,$$

the relation (2.18) becomes

$$H_{n+1}(z,q) = A_+ \, H_n(z,q).$$

From (2.17), we get

$$\frac{1}{z}\Delta H_n(z,q) = (1-q^n)\Delta H_{n-1}(z,q).$$

Setting the lowering operator

$$A_- := \frac{1}{z},$$

we get

$$A_- \Delta H_n(z,q) = (1-q^n)\Delta H_{n-1}(z,q).$$

The operator defined as $A_+ \, A_-$ gives

$$A_+ \, A_- H_n(z,q) = (1-q^n) H_n(z,q).$$

In fact,

$$\begin{aligned} A_+ \, A_- H_n(z,q) &= A_+ \Big(\frac{z(1-q^n)}{z} \, H_{n-1}(z,q) \Big) \\ &= (1-q^n) A_+ H_{n-1}(z,q) \\ &= (1-q^n) H_n(z,q). \end{aligned}$$

The raising and lowering operators may be written as

$$\mathfrak{S}_+ = (1+xz) - (1-q) z \, \partial_q$$

and

$$\mathfrak{S}_- = \partial_q$$

such that

$$\begin{aligned}
\mathfrak{S}_- H_n(z,q) &:= \partial_q H_n(z,q) = \sum_{k=0}^{n} \binom{n}{k} \partial_q z^k = \sum_{k=1}^{n} \binom{n}{k} [k]_q z^{k-1} \\
&= \sum_{t+1=0}^{n} \binom{n}{t+1} [t+1]_q z^t = \sum_{t=0}^{n-1} \binom{n}{t+1} [t+1]_q z^t \\
&= \sum_{t=0}^{n-1} \frac{[n]_q![t+1]_q}{[t+1]_q[t]_q![n-(t+1)]_q!} z^t \\
&= \sum_{t=0}^{n-1} \frac{[n]!}{[t]_q![n-(t+1)]_q!} z^t \\
&= [n]_q H_{n-1}(z,q),
\end{aligned}$$

and, using the same method, we also get

$$\mathfrak{S}_+ H_n(z,q) = H_{n+1}(z,q).$$

Indeed,

$$\begin{aligned}
\mathfrak{S}_+ H_n(z,q) &:= \left[(1+z) - (1-q)z\partial_q\right] H_n(z,q) \\
&= (1+z)H_n(z,q) - (1-q)z[n]_q H_{n-1}(z,q) \\
&= (1+z)H_n(z,q) - (1-q^n)z\, H_{n-1}(z,q) \\
&= H_{n+1}(z,q).
\end{aligned}$$

Defining $N_q := \mathfrak{S}_+\mathfrak{S}_-$, we see that

$$\begin{aligned}
N_q H_n(y;q) &= \mathfrak{S}_+ \left([n] H_{n-1}(y;q)\right) \\
&= [n]_q H_n(z;q)
\end{aligned}$$

which plays the role of the q-number operator. In fact,

$$\begin{aligned}
N_q\; H_n(z;q) &:= \mathfrak{S}_+\mathfrak{S}_-\; H_n(z;q) = \mathfrak{S}_+\, [n]_q\; H_{n-1}(z,q) \\
&= [n]_q\; \mathfrak{S}_+\; H_{n-1}(z,q) = [n]_q\; H_n(z,q) \\
&= \frac{1-q^n}{1-q} H_n(z,q).
\end{aligned}$$

The commutation relation between these operators can be obtained by direct computation:

$$\begin{aligned}[\mathfrak{S}_-, \mathfrak{S}_+]H_n(z, q) &= [\mathfrak{S}_-\mathfrak{S}_+ - \mathfrak{S}_+\mathfrak{S}_-]H_n(z, q)\\ &= \mathfrak{S}_-\mathfrak{S}_+H_n(z, q) - \mathfrak{S}_+\mathfrak{S}_-H_n(z, q)\\ &= \mathfrak{S}_-H_{n+1}(z, q) - N_q\, H_n(z, q)\\ &= [n+1]_q H_n(z, q) - N_q\, H_n(z, q)\\ &= q\, N_q\, H_n(z, q) + H_n(z, q) - N_q\, H_n(z, q)\\ &= \Big(1 - (1-q)N_q\Big)H_n(z, q),\end{aligned}$$

yielding

$$[\mathfrak{S}_+, \mathfrak{S}_-] = 1 - (1-q)N_q.$$

On the other hand, since we know (2.13), we further observe that, making use of the standard number operator N

$$NH_n(y; q) = nH_n(y; q),$$

we may write

$$N_q = \frac{1-q^N}{1-q}$$

such that

$$[\mathfrak{S}_-, \mathfrak{S}_+] = q^N$$

which is the particular case of the commutation relation for the q-deformed harmonic oscillator we have presented above. In the same form we can also obtain

$$[N_q, \mathfrak{S}_-] = [\mathfrak{S}_+\mathfrak{S}_-, \mathfrak{S}_-] = -[\mathfrak{S}_+, \mathfrak{S}_-]\mathfrak{S}_- = -q^N\mathfrak{S}_-,$$

$$[N_q, \mathfrak{S}_+] = [\mathfrak{S}_+\mathfrak{S}_-, \mathfrak{S}_+] = \mathfrak{S}_+[\mathfrak{S}_+, \mathfrak{S}_-] = \mathfrak{S}_+q^N$$

while

$$[N, \mathfrak{S}_-] = -\mathfrak{S}_-, \quad [N, \mathfrak{S}_+] = \mathfrak{S}_+.$$

It is also immediate to see that the q-commutation relation of these operators is given by

$$\mathfrak{S}_-\mathfrak{S}_+ - q\mathfrak{S}_+\mathfrak{S}_- = 1 + qN_q - qN_q = [N+1] - q[N] = 1,$$

which is the equivalent form of the commutation relation of the q-deformed harmonic oscillator. Now, from (2.13), and using the explicit realization of the raising and lowering operators, we can write

$$\begin{aligned} N_q H_n(z,q) &= \mathfrak{S}_+\mathfrak{S}_- H_n(z,q) = [1-(1-q)z\partial_q]\partial_q \\ &= [(1+z)\partial_q - (1-q)z\partial_q^2]H_n(z,q) = [n]H_n(z,q) \end{aligned}$$

from which we get the second order q-differential equation obeyed by the Rogers–Szegö polynomials

$$\Big[z\partial_q^2 + \frac{1+z}{1-q}\partial_q + \frac{[n]_q}{1-q}\Big]H_n(z,q) = 0.$$

3 ($\mathcal{R}, p, q$)-Generalized Rogers–Szegö Polynomials and Quantum Algebras

We start this section by some notions concerning the ($\mathcal{R}$, p, q)-calculus. Let us now consider two positive real numbers p and q, such that $0 < q < p$, and a meromorphic function $\mathcal{R}$, defined on $\mathbb{C} \times \mathbb{C}$ by:

$$\mathcal{R}(x,y) = \sum_{k,l=-L}^{\infty} r_{kl}x^k y^l$$

with an eventual isolated singularity at the zero, where r_{kl} are complex numbers, $L \in \mathbb{N} \cup \{0\}$, $\mathcal{R}(p^n, q^n) > 0\ \forall n \in \mathbb{N}$, and $\mathcal{R}(1, 1) = 0$. Denote by $\mathbb{D}_R = \{z \in \mathbb{C} : |z| < R\}$ a complex disc and by $\mathcal{O}(\mathbb{D}_R)$ the set of holomorphic functions defined on $\mathbb{D}_R$.

Definition 3.1 ([14]) The ($\mathcal{R}$, p, q)-deformed numbers is given by:

$$[n]_{\mathcal{R},p,q} := \mathcal{R}(p^n, q^n), \qquad n = 0, 1, 2, \cdots$$

leading to define ($\mathcal{R}$, p, q)-deformed factorials as

$$[n]!_{\mathcal{R},p,q} := \begin{cases} 1 \quad \text{for} \quad n = 0 \\ \mathcal{R}(p,q)\cdots\mathcal{R}(p^n,q^n) \quad \text{for} \quad n \geq 1, \end{cases}$$

and the $(\mathcal{R}, p, q)$-deformed binomial coefficients

$$\begin{bmatrix} m \\ n \end{bmatrix}_{\mathcal{R},p,q} := \frac{[m]!_{\mathcal{R},p,q}}{[n]!_{\mathcal{R},p,q}[m-n]!_{\mathcal{R},p,q}}, \quad m, n = 0, 1, 2, \cdots; \quad m \geq n.$$

The relation

$$\begin{bmatrix} m \\ n \end{bmatrix}_{\mathcal{R},p,q} = \begin{bmatrix} m \\ m-n \end{bmatrix}_{\mathcal{R},p,q}, \quad m, n = 0, 1, 2, \cdots; \quad m \geq n$$

holds.

Indeed, for $m, n = 0, 1, 2, \cdots$, and $m \geq n$, we get

$$\begin{aligned} \begin{bmatrix} m \\ n \end{bmatrix}_{\mathcal{R},p,q} &= \frac{[m]!_{\mathcal{R},p,q}}{[n]!_{\mathcal{R},p,q}\,[m-n]!_{\mathcal{R},p,q}} \\ &= \frac{[m]!_{\mathcal{R},p,q}}{[m-n]!_{\mathcal{R},p,q}\,[m-m+n]!_{\mathcal{R},p,q}} \\ &= \begin{bmatrix} m \\ m-n \end{bmatrix}_{\mathcal{R},p,q}. \end{aligned}$$

We recall also the following linear operators defined on $\mathcal{O}(\mathbb{D}_R)$ by (see [14] and references therein for more details):

$$\begin{aligned} Q : \varphi &\longmapsto Q\varphi(z) := \varphi(qz) \\ P : \varphi &\longmapsto P\varphi(z) := \varphi(pz) \\ \partial_{p,q} : \varphi &\longmapsto \partial_{p,q}\varphi(z) := \frac{\varphi(pz) - \varphi(qz)}{z(p-q)}, \end{aligned}$$

and the $(\mathcal{R}, p, q)$-deformed derivative given as follows:

$$\partial_{\mathcal{R},p,q} := \partial_{p,q}\frac{p-q}{P-Q}\mathcal{R}(P, Q) = \frac{p-q}{pP-qQ}\mathcal{R}(pP, qQ)\partial_{p,q}.$$

Proposition 3.2 *If there exist two functions* Φ_1 *and* $\Phi_2 : \mathbb{C} \times \mathbb{C}$ *such that*

$$\Phi_s(p, q) > 0, \quad \textit{for} \quad s = 1, 2$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{\mathcal{R},p,q} = \Phi_1^k(p, q)\begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} + \Phi_2^{n+1-k}(p, q)\begin{bmatrix} n \\ k-1 \end{bmatrix}_{\mathcal{R},p,q}$$

and

$$ba = \Phi_1(p, q)ab,\, xy = \Phi_2(p, q)yx,\, \text{and} \quad [s, t] = 0 \quad \text{for}\, s \in \{a, b\}, t \in \{x, y\}$$

for some algebra elements a, b, x, y, *then*

$$(a\,x + b\,y)^n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{n-k}\, b^k\, y^k\, x^{n-k}. \tag{3.1}$$

Proof By induction on n. In fact, for $n = 1$, we get

$$\begin{aligned}(a\,x + b\,y)^1 &= a\,x + b\,y \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{R},p,q} a^1\, b^0\, y^0\, x^1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{R},p,q} a^0\, b^1\, y^1\, x^0 \\ &= \sum_{k=0}^{1} \begin{bmatrix} 1 \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{1-k}\, b^k\, y^k\, x^{1-k}.\end{aligned}$$

Thus, the relation (3.1) holds for $n = 1$.

We assume that the relation (3.1) holds for $n \leq m$, this means in particular for $n = m$,

$$(a\,x + b\,y)^m = \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m-k}\, b^k\, y^k\, x^{m-k}.$$

Let us prove that it is also valid for $n = m + 1$. In fact,

$$\begin{aligned}(a\,x + b\,y)^{m+1} &= (a\,x + b\,y)^m (a\,x + b\,y) \\ &= \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m-k}\, b^k\, y^k\, x^{m-k} (a\,x + b\,y) \\ &= \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m-k}\, b^k\, y^k\, x^{m-k} a\,x \\ &+ \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m-k}\, b^k\, y^k\, x^{m-k} b\,y \\ &= \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} \Phi_1^k(p, q) a^{m+1-k}\, b^k\, y^k\, x^{m+1-k} \\ &+ \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} \Phi_2^{m-k}(p, q) a^{m-k}\, b^{k+1}\, y^{k+1}\, x^{m-k}\end{aligned}$$

$$
\begin{aligned}
&= a^{m+1}x^{m+1} + b^{m+1}y^{m+1} \\
&+ \sum_{k=1}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} \Phi_1^k(p,q)a^{m+1-k}b^k y^k x^{m+1-k} \\
&+ \sum_{k=1}^{m} \begin{bmatrix} m \\ k-1 \end{bmatrix}_{\mathcal{R},p,q} \Phi_2^{m+1-k}(p,q)a^{m+1-k} \\
&\times b^k\, y^k\, x^{m+1-k} \\
&= a^{m+1}x^{m+1} + b^{m+1}y^{m+1} \\
&+ \sum_{k=1}^{m} \begin{bmatrix} m \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m+1-k}\, b^k\, y^k\, x^{m+1-k} \\
&= \sum_{k=0}^{m+1} \begin{bmatrix} m+1 \\ k \end{bmatrix}_{\mathcal{R},p,q} a^{m+1-k}\, b^k\, y^k\, x^{m+1-k}.
\end{aligned}
$$

□

In [14], we introduced the quantum algebra associated with the $(\mathcal{R}, p, q)$-deformation. It is a quantum algebra, $\mathcal{A}_{\mathcal{R},p,q}$, generated by the set of operators $\{1, A, A^\dagger, N\}$ satisfying the following commutation relations:

$$
\begin{aligned}
AA^\dagger &= [N+1]_{\mathcal{R},p,q}, \qquad & A^\dagger A &= [N]_{\mathcal{R},p,q}. \\
[N,\ A] &= -A, & \left[N,\ A^\dagger\right] &= A^\dagger
\end{aligned}
\tag{3.2}
$$

with its realization on $\mathcal{O}(\mathbb{D}_R)$ given by:

$$
A^\dagger := z, \qquad A := \partial_{\mathcal{R},p,q}, \qquad N := z\partial_z,
$$

where $\partial_z := \frac{\partial}{\partial z}$ is the usual derivative on $\mathbb{C}$.

The second part of this section is dedicated to the general procedure for constructing the recursion relation for the $(\mathcal{R}, p, q)$-Rogers–Szegö polynomials and the related $(\mathcal{R}, p, q)$-difference equation that allow to define the creation, annihilation and number operators for a given $(\mathcal{R}, p, q)$-deformed quantum algebra. This is summarized as follows.

Theorem 3.3 *If $\phi_i(x, y)$ $(i = 1, 2, 3)$ are functions satisfying the following:*

$$
\phi_i(p,q) \neq 0 \qquad \textit{for } i = 1, 2, 3, \tag{3.3}
$$

$$
\phi_i(P,Q)z^k = \phi_i^k(p,q)z^k \quad \textit{for}\ \ z \in \mathbb{C},\ k = 0, 1, 2, \cdots \quad i = 1, 2 \tag{3.4}
$$

and if moreover the following relation between $(\mathcal{R}, p, q)$*-binomial coefficients holds*

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{\mathcal{R},p,q} = \phi_1^k(p,q) \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} + \phi_2^{n+1-k}(p,q) \begin{bmatrix} n \\ k-1 \end{bmatrix}_{\mathcal{R},p,q} \\ -\phi_3(p,q)[n]_{\mathcal{R},p,q} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{\mathcal{R},p,q} \tag{3.5}$$

for $1 \leq k \leq n$*, then the* $(\mathcal{R}, p, q)$*-Rogers–Szegö polynomials defined as*

$$H_n(z; \mathcal{R}, p, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} z^k, \quad n = 0, 1, 2 \cdots$$

satisfy the three-term recursion relation

$$H_{n+1}(z; \mathcal{R}, p, q) = H_n\left(\phi_1(p,q)z : \mathcal{R}, p, q\right) \\ +z\phi_2^n(p,q)H_n\left(z\phi_2^{-1}(p,q); \mathcal{R}, p, q\right) \\ -z\phi_3(p,q)[n]_{\mathcal{R},p,q}H_{n-1}\left(z; \mathcal{R}, p, q\right) \tag{3.6}$$

and $(\mathcal{R}, p, q)$*-difference equation*

$$\partial_{\mathcal{R},p,q} H_n(z; \mathcal{R}, p, q) = [n]_{\mathcal{R},p,q} H_{n-1}(z; \mathcal{R}, p, q). \tag{3.7}$$

Proof Multiplying the two sides of the relation (3.5) by z^k, and adding for $k = 1$ to n, we get

$$\sum_{k=1}^{n} \begin{bmatrix} n+1 \\ k \end{bmatrix}_{\mathcal{R},p,q} z^k = \sum_{k=1}^{n} \phi_1^k(p,q) \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} z^k \\ + \sum_{k=1}^{n} \phi_2^{n+1-k}(p,q) \begin{bmatrix} n \\ k-1 \end{bmatrix}_{\mathcal{R},p,q} z^k \\ -\phi_3(p,q)[n]_{\mathcal{R},p,q} \sum_{k=1}^{n} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{\mathcal{R},p,q} z^k.$$

After a short computation, and using the condition (3.5), we get the relation (3.6). Moreover,

$$\partial_{\mathcal{R},p,q} H_n(z; \mathcal{R}, p, q) = \partial_{\mathcal{R},p,q} \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} z^k \\ = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} \partial_{\mathcal{R},p,q} z^k$$

$$= \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{\mathcal{R},p,q} [n]_{\mathcal{R},p,q}\, z^{k-1}$$

$$= [n]_{\mathcal{R},p,q} \sum_{k=0}^{n-1} \begin{bmatrix} n-1 \\ k \end{bmatrix}_{\mathcal{R},p,q} z^{k}.$$

□

Setting

$$\psi_n(z;\mathcal{R},p,q) = \frac{1}{\sqrt{[n]!_{\mathcal{R},p,q}}} H_n(z;\mathcal{R},p,q), \tag{3.8}$$

and using Eqs. (3.6) and (3.7) yield the three-term recursion relation

$$\left(\phi_1(P,Q) + z\phi_2^n(p,q)\phi_2^{-1}(P,Q) - z\phi_3(p,q)\partial_{\mathcal{R},p,q}\right)\psi_n(z;\mathcal{R},p,q) = \sqrt{[n+1]_{\mathcal{R},p,q}}\psi_{n+1}(z;\mathcal{R},p,q)$$

and the $(\mathcal{R},p,q)$-difference equation

$$\partial_{\mathcal{R},p,q}\psi_n(z;\mathcal{R},p,q) = \sqrt{[n]_{\mathcal{R},p,q}}\psi_{n-1}(z;\mathcal{R},p,q)$$

for the polynomials $\psi_n(z;\mathcal{R},p,q)$ with the virtue that for $n = 0, 1, 2, \cdots$

$$\partial^{n+1}_{\mathcal{R},p,q}\psi_n(z;\mathcal{R},p,q) = 0 \quad \text{and} \quad \partial^{m}_{\mathcal{R},p,q}\psi_n(z;\mathcal{R},p,q) \neq 0 \text{ for any } m < n+1.$$

Indeed, from the relations (3.6) and (3.7), we get

$$\begin{aligned} H_{n+1}(z;\mathcal{R},p,q) &= H_n\left(\phi_1(p,q)z : \mathcal{R},p,q\right) \\ &+ z\phi_2^n(p,q)H_n\left(z\phi_2^{-1}(p,q);\mathcal{R},p,q\right) \\ &- z\phi_3(p,q)\partial_{\mathcal{R},p,q}H_n(z;\mathcal{R},p,q), \end{aligned}$$

and using the relation (3.8) yields

$$\begin{aligned} \sqrt{[n+1]_{\mathcal{R},p,q}}\,\psi_{n+1}(z;\mathcal{R},p,q) &= \psi_n\left(\phi_1(p,q)z : \mathcal{R},p,q\right) \\ &+ z\phi_2^n(p,q)\psi_n\left(z\phi_2^{-1}(p,q);\mathcal{R},p,q\right) \\ &- z\phi_3(p,q)\partial_{\mathcal{R},p,q}\psi_n\left(z;\mathcal{R},p,q\right). \end{aligned}$$

From the assumption (3.4), we get

$$\Big(\phi_1(P, Q) + z\phi_2^n(p, q)\phi_2^{-1}(P, Q) - z\phi_3(p, q)\partial_{\mathcal{R}, p, q}\Big)\,\psi_n(z; \mathcal{R}, p, q) = \sqrt{[n+1]_{\mathcal{R}, p, q}}\,\psi_{n+1}(z; \mathcal{R}, p, q).$$

Furthermore,

$$\begin{aligned}
\partial_{\mathcal{R}, p, q}\psi_n(z; \mathcal{R}, p, q) &:= \partial_{\mathcal{R}, p, q}\frac{1}{\sqrt{[n]!_{\mathcal{R}, p, q}}}H_n(z; \mathcal{R}, p, q) \\
&= \frac{1}{\sqrt{[n]!_{\mathcal{R}, p, q}}}\partial_{\mathcal{R}, p, q}H_n(z; \mathcal{R}, p, q) \\
&= \frac{1}{\sqrt{[n]!_{\mathcal{R}, p, q}}}[n]_{\mathcal{R}, p, q}H_{n-1}(z; \mathcal{R}, p, q) \\
&= \sqrt{[n]_{\mathcal{R}, p, q}}\,\psi_{n-1}(z; \mathcal{R}, p, q).
\end{aligned}$$

Now, formally defining the number operator N as

$$N\psi_n(z; \mathcal{R}, p, q) := n\psi_n(z; \mathcal{R}, p, q),$$

and the raising and lowering operators by

$$A^\dagger := \Big(\phi_1(P, Q) + z\phi_2^N(p, q)\phi_2^{-1}(P, Q) - z\phi_3(p, q)\partial_{\mathcal{R}, p, q}\Big) \quad \text{and} \quad A := \partial_{\mathcal{R}, p, q}$$

respectively, the set of polynomials $\{\psi_n(z; \mathcal{R}, p, q) \mid n = 0, 1, 2, \cdots\}$ provides a basis for a realization of $(\mathcal{R}, p, q)$-deformed quantum algebra $\mathcal{A}_{\mathcal{R}, p, q}$, satisfying the commutation relations (3.2). Provided the above formulated theorem, we can recover realizations in terms of Rogers–Szegö polynomials for different known deformations simply by determining the functions ϕ_i $(i = 1, 2, 3)$ that satisfy the relations (3.3)–(3.5).

4 Particulars Cases

4.1 (p, q)-Rogers–Szegö Polynomials and (p, q)-Oscillator from Jagannathan-Srinivasa Deformation [17]

Let us recall the definitions and notations used in the sequel.

Taking $\mathcal{R}(x, y) = \frac{x-y}{p-q}$, we obtain the Jagannathan-Srinivasa (p, q)-numbers and (p, q)-factorials

$$[n]_{p,q} = \frac{p^n - q^n}{p - q},$$

and

$$[n]!_{p,q} = \begin{cases} 1 \quad \text{for} \quad n = 0 \\ \frac{((p,q);(p,q))_n}{(p-q)^n} \quad \text{for} \quad n \geq 1, \end{cases}$$

respectively.

There result the following relevant properties.

Proposition 4.1 *If n, m are nonnegative integers, then*

$$[n]_{p,q} = \sum_{k=0}^{n-1} p^{n-1-k} q^k,$$
$$[n+m]_{p,q} = q^m [n]_{p,q} + p^n [m]_{p,q} = p^m [n]_{p,q} + q^n [m]_{p,q},$$

$$[-m]_{p,q} = -q^{-m} p^{-m} [m]_{p,q},$$
$$[n-m]_{p,q} = q^{-m} [n]_{p,q} - q^{-m} p^{n-m} [m]_{p,q} = p^{-m} [n]_{p,q} - q^{n-m} p^{-m} [m]_{p,q},$$

$$[n]_{p,q} = [2]_{p,q} [n-1]_{p,q} - pq [n-2]_{p,q}.$$

Proposition 4.2 *The (p, q)-binomial coefficients*

$$\begin{bmatrix} n \\ k \end{bmatrix}_{p,q} = \frac{((p, q); (p, q))_n}{((p, q); (p, q))_k ((p, q); (p, q))_{n-k}}, \quad 0 \leq k \leq n; \quad n \in \mathbb{N},$$

where $((p, q); (p, q))_m = (p - q)(p^2 - q^2) \cdots (p^m - q^m)$, $m \in \mathbb{N}$, *satisfy the following identities:*

$$\begin{bmatrix} n \\ k \end{bmatrix}_{p,q} = \begin{bmatrix} n \\ n-k \end{bmatrix}_{p,q} = p^{k(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_{q/p} = p^{k(n-k)} \begin{bmatrix} n \\ n-k \end{bmatrix}_{q/p},$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q} = p^k \begin{bmatrix} n \\ k \end{bmatrix}_{p,q} + q^{n+1-k} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q},$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q} = p^k \begin{bmatrix} n \\ k \end{bmatrix}_{p,q} + p^{n+1-k} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q} - (p^n - q^n) \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q} \tag{4.1}$$

with

$$\begin{bmatrix} n \\ k \end{bmatrix}_{q/p} = \frac{(q/p; q/p)_n}{(q/p; q/p)_k (q/p; q/p)_{n-k}},$$

where $(q/p; q/p)_n = (1-q/p)(1-q^2/p^2)\cdots(1-q^n/p^n)$*; and the* (p, q)*-shifted factorial*

$$((a, b); (p, q))_n := (a-b)(ap-bq)\cdots(ap^{n-1}-bq^{n-1})$$

or

$$((a, b); (p, q))_n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q} (-1)^k p^{(n-k)(n-k-1)/2} q^{k(k-1)/2} a^{n-k} b^k.$$

Finally, the algebra $\mathcal{A}_{p,q}$, generated by $\{1, \; A, \; A^\dagger, \; N\}$, associated with (p, q)-Janagathan-Srinivasa deformation, satisfies the following commutation relations:

$$\begin{aligned} &A\,A^\dagger - pA^\dagger A = q^N, \qquad &&A\,A^\dagger - qA^\dagger A = p^N \\ &[N, \; A^\dagger] = A^\dagger, &&[N, \; A] = -A. \end{aligned} \tag{4.2}$$

The (p, q)-Rogers–Szegö polynomials studied in [17] appear as a particular case obtained by choosing $\phi_1(x, y) = \phi_2(x, y) = \phi(x, y) = x$ and $\phi_3(x, y) = x - y$. Indeed, $\phi(p, q) = p \neq 0$, $\phi_3(p, q) = p - q \neq 0$, $\phi(P, Q)z^k = \phi_1^k(p, q)z^k$ and Eq. (4.1) shows that

$$\begin{aligned} \begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q} = p^k \begin{bmatrix} n \\ k \end{bmatrix}_{p,q} + p^{n+1-k} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q} \\ -(p-q)[n]_{p,q} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q}. \end{aligned}$$

Hence, the hypotheses of the above theorem are satisfied and, therefore, the (p, q)-Rogers–Szegö polynomials

$$H_n(z; p, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q} z^k \quad n = 0, 1, 2, \cdots$$

satisfy the three-term recursion relation

$$H_{n+1}(z; p, q) = H_n(pz; p, q) + zp^n H_n(p^{-1}z; p, q) - z(p^n - q^n)H_{n-1}(z; p, q) \tag{4.3}$$

and the (p,q)-difference equation

$$\partial_{p,q} H_n(z;\,p,q) = [n]_{p,q} H_{n-1}(z;\,p,q).$$

Finally, the set of polynomials

$$\psi_n(z;\,p,q) = \frac{1}{\sqrt{[n]!_{p,q}}} H_n(z;\,p,q), \quad n = 0,1,2,\cdots$$

forms a basis for a realization of the (p,q)-deformed harmonic oscillator and quantum algebra $\mathcal{A}_{p,q}$ satisfying the commutation relations (4.2) with the number operator N defined as

$$N\psi_n(z;\,p,q) := n\psi_n(z;\,p,q),$$

relating the annihilation and creation operators given by

$$A = \partial_{p,q} \quad \text{and} \quad A^\dagger = P + zp^N P^{-1} - z(p-q)\partial_{p,q}$$

respectively.

Let us compute the first three (p,q)-Rogers–Szegö polynomials using the three-term recursion relation (4.3) with the initial one $H_0(z;\,p,q) = 1$ and $H_{-1}(z;\,p,q) = 0$. We get,

- $$H_1(z;\,p,q) = H_0(pz;\,p,q) + zp^0 H_0(p^{-1}z;\,p,q) - z(p^0 - q^0)H_{-1}(z;\,p,q)$$

  which gives

  $$H_1(z;\,p,q) = z + 1.$$

- $$\begin{aligned} H_2(z;\,p,q) &= H_1(pz;\,p,q) + zpH_1(p^{-1}z;\,p,q) - z(p-q)H_0(z;\,p,q) \\ &= 1 + pz + zp(1 + p^{-1}z) - z(p-q) \\ &= 1 + pz + zp + zpp^{-1}z - zp + zq \\ &= z^2 + (p+q)z + 1 \end{aligned}$$

- $$H_3(z;\,p,q) = H_2(pz;\,p,q) + zp^2 H_2(p^{-1}z;\,p,q) - z(p^2 - q^2)H_1(z;\,p,q).$$

Since

$$\begin{aligned} H_2(pz;\, p, q) &= (pz)^2 + (p+q)pz + 1 \\ &= p^2\, z^2 + p^2\, z + q\, p\, z + 1 \\ &= p^2\, z^2 + (p^2 + q\, p)z + 1 \end{aligned}$$

$$\begin{aligned} zp^2 H_2(p^{-1}z;\, p, q) &= zp^2\big((p^{-1}z)^2 + (p+q)p^{-1}z + 1\big) \\ &= zp^2\big(p^{-2}z^2 + z + qp^{-1}z + 1\big) \\ &= z^3 + z^2p^2 + qpz^2 + zp^2 \\ &= z^3 + (p^2 + qp)z^2 + zp^2 \end{aligned}$$

and

$$\begin{aligned} z(p^2 - q^2)H_1(z;\, p, q) &= z(p^2 - q^2)(1+z) \\ &= z(p^2 - q^2) + z^2(p^2 - q^2), \end{aligned}$$

therefore,

$$H_3(z;\, p, q) = z^3 + (p^2 + qp + q^2)z^2 + (p^2 + qp + q^2)z + 1.$$

4.2 Rogers–Szegö Polynomial Associated to the Chakrabarty-Jagannathan Deformation [6]

We start this section by recalling some notions concerning the Chakrabarty - Jagannathan deformation. We recover the deformation of Chakrabarty and Jagannathan [6] by taking $\mathcal{R}(x, y) = \frac{1-xy}{(p^{-1}-q)x}$. Indeed, the $(\mathcal{R}, p, q)$-numbers and $(\mathcal{R}, p, q)$-factorials are reduced to (p^{-1}, q)-numbers and (p^{-1}, q)-factorials, namely,

$$[n]_{p^{-1},q} = \frac{p^{-n} - q^n}{p^{-1} - q},$$

and

$$[n]!_{p^{-1},q} = \begin{cases} 1 \quad \text{for} \quad n = 0 \\ \frac{((p^{-1},q);(p^{-1},q))_n}{(p^{-1}-q)^n} \quad \text{for} \quad n \geq 1, \end{cases}$$

respectively. The deformation properties can be recovered in the following propositions.

Proposition 4.3 *If n, m are nonnegative integers, then*

$$[n]_{p^{-1},q} = \sum_{k=0}^{n-1} p^{-n+1+k} q^k,$$

$$[n+m]_{p^{-1},q} = q^m[n]_{p^{-1},q} + p^{-n}[m]_{p^{-1},q} = p^{-m}[n]_{p^{-1},q} + q^n[m]_{p^{-1},q},$$

$$[-m]_{p^{-1},q} = -q^{-m}p^m[m]_{p^{-1},q},$$

$$\begin{aligned}[n-m]_{p^{-1},q} &= q^{-m}[n]_{p^{-1},q} - q^{-m}p^{-n+m}[m]_{p,q} \\ &= p^m[n]_{p^{-1},q} - q^{n-m}p^m[m]_{p^{-1},q},\end{aligned}$$

$$[n]_{p^{-1},q} = [2]_{p^{-1},q}[n-1]_{p^{-1},q} - p^{-1}q[n-2]_{p^{-1},q}.$$

Proposition 4.4 *The* (p^{-1}, q)*-binomial coefficients*

$$\begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} = \frac{((p^{-1},q);(p^{-1},q))_n}{((p^{-1},q);(p^{-1},q))_k((p^{-1},q);(p^{-1},q))_{n-k}},\ 0 \le k \le n;\ n \in \mathbb{N},$$

where $((p,q);(p,q))_m = (p^{-1}-q)(p^{-2}-q^2)\cdots(p^{-m}-q^m)$*,* $m \in \mathbb{N}$*, satisfy the following identities:*

$$\begin{aligned}\begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} &= \begin{bmatrix} n \\ n-k \end{bmatrix}_{p^{-1},q} = p^{-k(n-k)}\begin{bmatrix} n \\ k \end{bmatrix}_{q/p^{-1}}, \\ \begin{bmatrix} n+1 \\ k \end{bmatrix}_{p^{-1},q} &= p^k\begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} + q^{n+1-k}\begin{bmatrix} n \\ k-1 \end{bmatrix}_{p^{-1},q}, \\ \begin{bmatrix} n+1 \\ k \end{bmatrix}_{p^{-1},q} &= p^{-k}\begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} + p^{-n-1+k}\begin{bmatrix} n \\ k-1 \end{bmatrix}_{p^{-1},q} \\ &\quad -(p^{-n}-q^n)\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p^{-1},q}.\end{aligned}$$

with

$$\begin{bmatrix} n \\ k \end{bmatrix}_{q/p^{-1}} = \frac{(q/p^{-1};q/p^{-1})_n}{(q/p^{-1};q/p^{-1})_k(q/p^{-1};q/p^{-1})_{n-k}},$$

where $(q/p^{-1}; q/p^{-1})_n = (1 - q/p^{-1})(1 - q^2/p^{-2}) \cdots (1 - q^n/p^{-n})$*; and the* (p^{-1}, q)*-shifted factorial*

$$((a, b); (p^{-1}, q))_n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} (-1)^k p^{-(n-k)(n-k-1)/2} q^{k(k-1)/2} a^{n-k} b^k$$

or

$$((a, b); (p^{-1}, q))_n := (a - b)(ap^{-1} - bq) \cdots (ap^{-n+1} - bq^{n-1}).$$

The $(\mathcal{R}, p, q)$-derivative is also reduced to (p^{-1}, q)-derivative. Indeed,

$$\begin{aligned} \partial_{\mathcal{R},p,q} &= \partial_{p,q} \frac{p - q}{P - Q} \frac{1 - PQ}{(p^{-1} - q)P} \\ &= \frac{1}{(p^{-1} - q)z} (P^{-1} - Q) := \partial_{p^{-1},q} \end{aligned}$$

The algebra $\mathcal{A}_{p^{-1},q}$, induced by $\{1, \ A, \ A^\dagger, \ N\}$, related to the (p, q)-Chakrabarty and Jagannathan deformation satisfies the following commutation relations:

$$\begin{aligned} &A\ A^\dagger - p^{-1} A^\dagger A = q^N, &&A\ A^\dagger - q A^\dagger A = p^{-N} \\ &[N, \ A^\dagger] = A^\dagger &&[N, \ A] = -A. \end{aligned} \tag{4.4}$$

As previously shown, the Chakrabarty-Jagannathan deformation [6] can be obtained from the (p, q) deformation by replacing the parameter p by p^{-1} and the operator of dilatation P by P^{-1}. Hence, the (p^{-1}, q)-Rogers–Szegö polynomials

$$H_n(z; p^{-1}, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} z^k, \quad n = 0, 1, 2, \cdots$$

satisfy the three-term recursion relation

$$\begin{aligned} H_{n+1}(z; p^{-1}, q) = H_n(p^{-1}z; p^{-1}, q) + zp^{-n} H_n(pz; p^{-1}, q) \\ -z(p^{-n} - q^n) H_{n-1}(z; p^{-1}, q) \end{aligned} \tag{4.5}$$

and the (p^{-1}, q)-difference equation

$$\partial_{p^{-1},q} H_n(z; p, q) = [n]_{p^{-1},q} H_{n-1}(z; p, q).$$

Finally, the set of polynomials

$$\psi_n(z; p^{-1}, q) = \frac{1}{\sqrt{[n]!_{p^{-1},q}}} H_n(z; p^{-1}, q), \quad n = 0, 1, 2, \cdots$$

forms a basis for a realization of the (p^{-1}, q)-deformed harmonic oscillator and quantum algebra $\mathcal{A}_{p^{-1},q}$ satisfying the commutation relations (4.4) with the number operator N formally defined as

$$N\psi_n(z; p^{-1}, q) := n\psi_n(z; p^{-1}, q),$$

the annihilation and creation operators given by

$$A := \partial_{p^{-1},q} \quad \text{and} \quad A^{\dagger} := P^{-1} + zp^{-N}P - z(p^{-1} - q)\partial_{p^{-1},q},$$

respectively.

Now, we compute the first three (p^{-1}, q)-Rogers–Szegö polynomials using the three-term recursion relation (4.5) with the initial one $H_0(z; p^{-1}, q) = 1$ and $H_{-1}(z; p^{-1}, q) = 0$. We get,

-
$$\begin{aligned} H_1(z; p^{-1}, q) = H_0(p^{-1}z; p^{-1}, q) + zp^0 H_0(pz; p^{-1}, q) \\ -z(p^0 - q^0)H_{-1}(z; p^{-1}, q) \end{aligned}$$

  which gives

$$H_1(z; p^{-1}, q) = z + 1.$$

-
$$\begin{aligned} H_2(z; p^{-1}, q) &= H_1(p^{-1}z; p^{-1}, q) + zpH_1(pz; p^{-1}, q) \\ &\quad - z(p^{-1} - q)H_0(z; p^{-1}, q) \\ &= 1 + p^{-1}z + zp^{-1}(1 + pz) - z(p^{-1} - q) \\ &= 1 + p^{-1}z + zp^{-1} + zpp^{-1}z - zp^{-1} + zq \\ &= z^2 + (p^{-1} + q)z + 1 \end{aligned}$$

-
$$\begin{aligned} H_3(z; p^{-1}, q) = H_2(p^{-1}z; p^{-1}, q) + zp^{-2}H_2(pz; p^{-1}, q) \\ -z(p^{-2} - q^2)H_1(z; p^{-1}, q). \end{aligned}$$

Since

$$\begin{aligned} H_2(p^{-1}z;\, p^{-1}, q) &= (p^{-1}z)^2 + (p^{-1}+q)p^{-1}z + 1 \\ &= p^{-2}\, z^2 + p^{-2}\, z + q\, p^{-1}\, z + 1 \\ &= p^{-2}\, z^2 + (p^{-2} + q\, p^{-1})z + 1 \end{aligned}$$

$$\begin{aligned} zp^{-2}H_2(pz;\, p^{-1}, q) &= zp^{-2}\big((pz)^2 + (p^{-1}+q)pz + 1\big) \\ &= zp^{-2}\big(p^2z^2 + z + qpz + 1\big) \\ &= z^3 + z^2p^{-2} + qp^{-1}z^2 + zp^{-2} \\ &= z^3 + (p^{-2} + qp^{-1})z^2 + zp^{-2} \end{aligned}$$

and

$$\begin{aligned} z(p^{-2} - q^2)H_1(z;\, p^{-1}, q) &= z(p^{-2} - q^2)(1+z) \\ &= z(p^{-2} - q^2) + z^2(p^{-2} - q^2), \end{aligned}$$

therefore,

$$H_3(z;\, p^{-1}, q) = z^3 + (p^{-2} + qp^{-1} + q^2)z^2 + (p^{-2} + qp^{-1} + q^2)z + 1.$$

4.3 (p, q)-Rogers–Szegö Polynomials Associated with the Quesne's Deformed Quantum Algebra [13]

In [13], we investigated a generalization of q-Quesne algebra [25] and defined the (p, q)-numbers and a generalized (p, q)-Quesne algebra. Our results can be recovered from the previous general theory by taking $\mathcal{R}(x, y) = \frac{xy-1}{(q-p^{-1})y}$. Indeed, the (p, q)-Quesne numbers and (p, q)-Quesne factorials are given by

$$[n]^Q_{p,q} = \frac{p^n - q^{-n}}{q - p^{-1}},$$

and

$$[n]!^Q_{p,q} = \begin{cases} 1 \quad \text{for} \quad n = 0 \\ \frac{((p,q^{-1});(p,q^{-1}))_n}{(q-p^{-1})^n} \quad \text{for} \quad n \geq 1, \end{cases}$$

respectively. Then follow some remarkable properties.

Proposition 4.5 *If n, m are nonnegative integers, then*

$$[-m]^Q_{p,q} = -p^{-m}q^m[m]^Q_{p,q}, \tag{4.6}$$

$$[n+m]^Q_{p,q} = q^{-m}[n]^Q_{p,q} + p^n[m]^Q_{p,q} = p^m[n]^Q_{p,q} + q^{-n}[m]^Q_{p,q}, \tag{4.7}$$

$$[n-m]^Q_{p,q} = q^m[n]^Q_{p,q} - p^{n-m}q^m[m]^Q_{p,q} = p^{-m}[n]^Q_{p,q} + p^{-m}q^{m-n}[m]^Q_{p,q}, \tag{4.8}$$

$$[n]^Q_{p,q} = \frac{q-p^{-1}}{p-q^{-1}}[2]^Q_{p,q}[n-1]^Q_{p,q} - pq^{-1}[n-2]^Q_{p,q}. \tag{4.9}$$

Proof Equations (4.6) and (4.7) are immediate by the application of the relations $p^{-m} - q^m = -p^{-m}q^m(p^m - q^{-m})$ and $p^{n+m} - q^{-n-m} = q^{-m}(p^n - q^{-n}) + p^n(p^m - q^{-m}) = p^m(p^n - q^{-n}) + q^{-n}(p^m - q^{-m})$, respectively, while Eq. (4.8) results from the combination of Eqs. (4.6) and (4.7). Finally, the relation

$$[n]_{p,q^{-1}} = \frac{p^n - q^{-n}}{p - q^{-1}} = \frac{q-p^{-1}}{p-q^{-1}}\frac{p^n - q^{-n}}{q-p^{-1}} = \frac{q-p^{-1}}{p-q^{-1}}[n]^Q_{p,q}, \quad n = 1, 2, \cdots$$

cumulatively taken with the identity

$$[n]_{p,q-1} = [2]_{p,q-1}[n-1]_{p,q-1} - pq^{-1}[n-2]_{p,q^{-1}}$$

gives Eq. (4.9). □

Proposition 4.6 *The* (p,q)*-Quesne binomial coefficients*

$$\begin{bmatrix} n \\ k \end{bmatrix}^Q_{p,q} = \frac{((p,q^{-1});(p,q^{-1}))_n}{((p,q^{-1});(p,q^{-1}))_k((p,q^{-1});(p,q^{-1}))_{n-k}}, \quad 0 \le k \le n; \quad n \in \mathbb{N},$$

satisfy the following properties

$$\begin{aligned} \begin{bmatrix} n \\ k \end{bmatrix}^Q_{p,q} = \begin{bmatrix} n \\ n-k \end{bmatrix}^Q_{p,q} &= p^{k(n-k)}\begin{bmatrix} n \\ k \end{bmatrix}_{1/qp} \\ &= p^{k(n-k)}\begin{bmatrix} n \\ n-k \end{bmatrix}_{1/qp}, \end{aligned}$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}^Q_{p,q} = p^k\begin{bmatrix} n \\ k \end{bmatrix}^Q_{p,q} + q^{-n-1+k}\begin{bmatrix} n \\ k-1 \end{bmatrix}^Q_{p,q},$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q}^{Q} = p^k \begin{bmatrix} n \\ k \end{bmatrix}_{p,q}^{Q} + p^{n+1-k} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q}^{Q} - (p^n - q^{-n}) \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q}^{Q}. \quad (4.10)$$

Proof It is straightforward, using Proposition 4.3 and

$$\begin{bmatrix} n \\ k \end{bmatrix}_{p,q}^{Q} = \begin{bmatrix} n \\ k \end{bmatrix}_{p,q^{-1}}.$$

□

Finally, the algebra $\mathcal{A}_{p,q}^{Q}$, generated by $\{1,\ A,\ A^{\dagger},\ N\}$, associated with (p, q)-Quesne deformation satisfies the following commutation relations:

$$p^{-1}A\,A^{\dagger} - A^{\dagger}A = q^{-N-1}, \qquad qA\,A^{\dagger} - A^{\dagger}A = p^{N+1}$$
$$[N,\ A^{\dagger}] = A^{\dagger}, \qquad [N,\ A] = -A. \quad (4.11)$$

The (p, q)-Rogers–Szegö polynomials corresponding to the Quesne deformation are deduced from our generalization by choosing $\phi_1(x, y) = \phi_2(x, y) = \phi(x, y) = x$ and $\phi_3(x, y) = y - x^{-1}$. Indeed, it is worthy of attention that we get in this case $\phi(p, q) = p \neq 0$, $\phi_3(p, q) = q - p^{-1} \neq 0$, $\phi(P, Q)z^k = \phi_1^k(p, q)z^k$ and from Eq. (4.10)

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q}^{Q} = p^k \begin{bmatrix} n \\ k \end{bmatrix}_{p,q}^{Q} + p^{n+1-k} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q}^{Q} - (q - p^{-1})[n]_{p,q} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q}^{Q}.$$

Hence, the hypotheses of the theorem are satisfied and, therefore, the (p, q)-Rogers–Szegö polynomials

$$H_n^{Q}(z; p, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q}^{Q} z^k \quad n = 0, 1, 2, \cdots$$

satisfy the three-term recursion relation

$$H_{n+1}^{Q}(z; p, q) = H_n^{Q}(pz; p, q) + zp^n H_n^{Q}(p^{-1}z; p, q) - z(p^n - q^{-n})H_{n-1}^{Q}(z; p, q) \quad (4.12)$$

and the (p,q)-difference equation

$$\partial^Q_{p,q} H^Q_n(z;p,q) = [n]^Q_{p,q} H^Q_{n-1}(z;p,q).$$

Thus, the set of polynomials

$$\psi^Q_n(z;p,q) = \frac{1}{\sqrt{[n]!^Q_{p,q}}} H^Q_n(z;p,q), \quad n = 0, 1, 2, \cdots$$

forms a basis for a realization of the (p,q)-Quesne deformed harmonic oscillator and quantum algebra $\mathcal{A}^Q_{p,q}$ engendering the commutation relations (4.11) with the number operator N formally defined as

$$N\psi^Q_n(z;p,q) := n\psi^Q_n(z;p,q),$$

the annihilation and creation operators given by

$$A := \partial^Q_{p,q} \quad \text{and} \quad A^\dagger := P + zp^N P^{-1} - z(q - p^{-1})\partial_{p,q},$$

respectively. Naturally, setting $p = 1$ gives the Rogers–Szegö polynomials associated with the q-Quesne deformation [25].

Let us compute the first three (p,q)-Rogers–Szegö polynomials associated with the q-Quesne quantum algebra using the three-term recursion relation (4.12) with the initial one $H_0(z;p,q^{-1}) = 1$ and $H_{-1}(z;p,q^{-1}) = 0$. We get,

- $$\begin{aligned} H_1(z;p,q^{-1}) = H_0(pz;p,q^{-1}) + zp^0 H_0(p^{-1}z;p,q^{-1}) \\ -z(p^0 - q^0)H_{-1}(z;p,q^{-1}) \end{aligned}$$

  which gives

  $$H_1(z;p,q^{-1}) = z + 1.$$

- $$\begin{aligned} H_2(z;p,q^{-1}) &= H_1(pz;p,q^{-1}) + zpH_1(p^{-1}z;p,q^{-1}) \\ &\quad -z(p - q^{-1})H_0(z;p,q^{-1}) \\ &= 1 + pz + zp(1 + p^{-1}z) - z(p - q^{-1}) \\ &= 1 + pz + zp + zpp^{-1}z - zp + zq^{-1} \\ &= z^2 + (p + q^{-1})z + 1 \end{aligned}$$

•

$$H_3(z;\, p, q^{-1}) = H_2(pz;\, p, q^{-1}) + zp^2 H_2(p^{-1}z;\, p, q^{-1}) \\ -z(p^2 - q^{-2}) H_1(z;\, p, q^{-1}).$$

Since

$$\begin{aligned} H_2(pz;\, p, q^{-1}) &= (pz)^2 + (p + q^{-1})pz + 1 \\ &= p^2 z^2 + p^2 z + q^{-1}\, p\, z + 1 \\ &= p^2 z^2 + (p^2 + q^{-1}\, p)z + 1 \end{aligned}$$

$$\begin{aligned} zp^2 H_2(p^{-1}z;\, p, q^{-1}) &= zp^2\big((p^{-1}z)^2 + (p + q^{-1})p^{-1}z + 1\big) \\ &= zp^2\big(p^{-2}z^2 + z + q^{-1}p^{-1}z + 1\big) \\ &= z^3 + z^2p^2 + q^{-1}pz^2 + zp^2 \\ &= z^3 + (p^2 + q^{-1}p)z^2 + zp^2 \end{aligned}$$

and

$$\begin{aligned} z(p^2 - q^{-2})H_1(z;\, p, q^{-1}) &= z(p^2 - q^{-2})(1 + z) \\ &= z(p^2 - q^{-2}) + z^2(p^2 - q^{-2}), \end{aligned}$$

therefore,

$$H_3(z;\, p, q^{-1}) = z^3 + (p^2 + q^{-1}p + q^{-2})z^2 + (p^2 + q^{-1}p + q^{-2})z + 1.$$

4.4 (p, q, μ, ν, h)-Rogers–Szegö Polynomials Associated to Hounkonnou-Ngompe Quantum Algebra [15]

The $(p, q; \mu, \nu, h)$-deformation studied in [15] can be recovered by taking $\mathcal{R}(x, y) = h(p, q)y^\nu/x^\mu[\frac{xy-1}{(q-p^{-1})y}]$ with $0 < pq < 1$, $p^\mu < q^{\nu-1}$, $p > 1$, h being a well behaved real and non-negative function of deformation parameters p and q such that $h(p, q) \to 1$ as $(p, q) \to (1, 1)$. Here the $(\mathcal{R}, p, q)$-numbers become $(p, q; \mu, \nu, h)$-numbers defined by

$$[n]^{\mu,\nu}_{p,q,h} := h(p, q)\frac{q^{\nu n}}{p^{\mu n}}\frac{p^n - q^{-n}}{q - p^{-1}}.$$

Proposition 4.7 *The* $(p, q; \mu, \nu, h)$*-numbers verify the following properties, for* $m, n \in \mathbb{N}$*:*

$$
\begin{aligned}
[-m]^{\mu,\nu}_{p,q,h} &= -\frac{q^{-2\nu m+m}}{p^{-2\mu m+m}}[m]^{\mu,\nu}_{p,q,h}, \\
[n+m]^{\mu,\nu}_{p,q,h} &= \frac{q^{\nu m-m}}{p^{\mu m}}[n]^{\mu,\nu}_{p,q,h} + \frac{q^{\nu n}}{p^{\mu n-n}}[m]^{\mu,\nu}_{p,q,h} \\
&= \frac{q^{\nu m}}{p^{\mu m-m}}[n]^{\mu,\nu}_{p,q,h} + \frac{q^{\nu n-n}}{p^{\mu n}}[m]^{\mu,\nu}_{p,q,h}\ , \\
[n-m]^{\mu,\nu}_{p,q,h} &= \frac{q^{-\nu m+m}}{p^{-\mu m}}[n]^{\mu,\nu}_{p,q,h} - \frac{q^{\nu(n-2m)+m}}{p^{\mu(n-2m)-n+m}}[m]^{\mu,\nu}_{p,q,h} \\
&= \frac{q^{-\nu m}}{p^{-\mu m+m}}[n]^{\mu,\nu}_{p,q,h} - \frac{q^{\nu(n-2m)-n+m}}{p^{\mu(n-2m)+m}}[m]^{\mu,\nu}_{p,q,h}, \\
[n]^{\mu,\nu}_{p,q,h} &= \frac{q-p^{-1}}{p-q^{-1}}\frac{q^{-\nu}}{p^{-\mu}}\frac{1}{h(p,q)}[2]^{\mu,\nu}_{p,q,h}[n-1]^{\mu,\nu}_{p,q,h} \\
&\quad -\frac{q^{2\nu-1}}{p^{2\nu-1}}[n-2]^{\mu,\nu}_{p,q,h}.
\end{aligned}
$$

Proof This is direct using Proposition 4.5 and the fact that

$$
[n]^{\mu,\nu}_{p,q,h} = h(p,q)\frac{q^{\nu n}}{p^{\mu n}}[n]^{Q}_{p,q,h}. \tag{4.13}
$$

□

Proposition 4.8 *The* (p, q, μ, ν, h)*-binomial coefficients*

$$
\begin{bmatrix} n \\ k \end{bmatrix}^{\mu,\nu}_{p,q,h} := \frac{[n]!^{\mu,\nu}_{p,q,h}}{[k]!^{\mu,\nu}_{p,q,h}[n-k]!^{\mu,\nu}_{p,q,h}} = \frac{q^{\nu k(n-k)}}{p^{\mu k(n-k)}}\begin{bmatrix} n \\ k \end{bmatrix}^{Q}_{p,q},\ 0 \le k \le n;\ \ n \in \mathbb{N},
$$

satisfy the following properties

$$
\begin{aligned}
\begin{bmatrix} n \\ k \end{bmatrix}^{\mu,\nu}_{p,q,h} &= \begin{bmatrix} n \\ n-k \end{bmatrix}^{\mu,\nu}_{p,q,h}, \\
\begin{bmatrix} n+1 \\ k \end{bmatrix}^{\mu,\nu}_{p,q,h} &= \frac{q^{\nu k}}{p^{(\mu-1)k}}\begin{bmatrix} n \\ k \end{bmatrix}^{\mu,\nu}_{p,q,h} + \frac{q^{(\nu-1)(n+1-k)}}{p^{\mu(n+1-k)}}\begin{bmatrix} n \\ k-1 \end{bmatrix}^{\mu,\nu}_{p,q,h},
\end{aligned}
$$

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} = \frac{q^{\nu k}}{p^{(\mu-1)k}} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} + \frac{q^{\nu(n+1-k)}}{p^{(\mu-1)(n+1-k)}} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q,h}^{\mu,\nu} - (p^n - q^{-n}) \frac{q^{\nu n}}{p^{\mu n}} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q,h}^{\mu,\nu}. \tag{4.14}$$

Proof This follows from Proposition 4.6 and the fact that

$$[n]!_{p,q,h}^{\mu,\nu} = h^n(p,q) \frac{q^{n(n+1)/2}}{p^{n(n+1)/2}} [n]!_{p,q}^{Q},$$

where use of Eq. (4.13) has been made. □

The algebra $\mathcal{A}_{p,q,h}^{\mu,\nu}$, generated by $\{1,\ A,\ A^\dagger,\ N\}$, associated with (p, q, μ, ν, h)-deformation satisfies the following commutation relations:

$$\begin{aligned} p^{-1} A\, A^\dagger - \frac{q^\nu}{p^\mu} A^\dagger A &= h(p,q) \left(\frac{q^{\nu-1}}{p^\mu} \right)^{N+1}, \\ q A\, A^\dagger - \frac{q^\nu}{p^\mu} A^\dagger A &= h(p,q) \left(\frac{q^\nu}{p^{\mu-1}} \right)^{N+1} \\ [N,\ A^\dagger] = A^\dagger, \qquad & [N,\ A] = -A. \end{aligned} \tag{4.15}$$

The (p, q, μ, ν, h)-Rogers–Szegö polynomials are deduced from our general case by setting $\phi_1(x,y) = x^{1-\mu} y^\nu$, $\phi_2(x,y) = x^{-\mu} y^{\nu-1}$ and $\phi_3(x,y) = \frac{y - x^{-1}}{h(p,q)}$. Indeed, $\phi_i(p,q) \neq 0$ for $i = 1, 2, 3$; $\phi_i(P,Q) z^k = \phi_i(p,q)^k z^k$ for $i = 1, 2$ and the property (4.14) furnishes

$$\begin{bmatrix} n+1 \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} = \frac{q^{\nu k}}{p^{(\mu-1)k}} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} - \frac{q - p^{-1}}{h(p,q)} [n]_{p,q,h}^{\mu,\nu} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_{p,q,h}^{\mu,\nu} + \frac{q^{\nu(n+1-k)}}{p^{(\mu-1)(n+1-k)}} \begin{bmatrix} n \\ k-1 \end{bmatrix}_{p,q,h}^{\mu,\nu}.$$

Therefore, the (p, q, μ, ν, h)-Rogers–Szegö polynomials are defined as follows

$$H_n(z; p, q, \mu, \nu, h) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} z^k, \quad n = 0, 1, 2 \cdots$$

with the three-term recursion relation

$$H_{n+1}(z;p,q,\mu,\nu,h) = H_n\left(\frac{q^\nu}{p^{\mu-1}}z : p,q,\mu,\nu,h\right) + z\frac{q^{(\nu-1)n}}{p^{\mu n}}H_n\left(\frac{p^\nu}{q^{\nu-1}}z; p,q,\mu,\nu,h\right) - z\frac{q^{\nu n}}{p^{\mu n}}(p^n - q^{-n})H_{n-1}(z;p,q,\mu,\nu,h)$$

and the (p,q,μ,ν,h)-difference equation

$$\partial^{\mu,\nu}_{p,q,h}H_n(z;p,q,\mu,\nu,h) = [n]^{\mu,\nu}_{p,q,h}H_{n-1}(z;p,q,\mu,\nu,h).$$

Hence, the set of polynomials

$$\psi_n(z;p,q,\mu,\nu,h) = \frac{1}{\sqrt{[n]!^{\mu,\nu}_{p,q,h}}}H_n(z;p,q,\mu,\nu,h), \quad n = 0,1,2,\cdots$$

forms a basis for a realization of the (p,q,μ,ν,h)-deformed algebra $\mathcal{A}_{p,q,\mu,\nu,h}$ satisfying the commutation relations (4.15) with the number operator N formally defined as

$$N\psi^Q_n(z;p,q,\mu,\nu,h) := n\psi_n(z;p,q,\mu,\nu,h),$$

together with the annihilation and the creation operators given by

$$A := \partial^{\mu,\nu}_{p,q,h} \quad \text{and} \quad A^\dagger := \frac{Q^\nu}{P^{\mu-1}} + z\left(\frac{q^{\nu-1}}{p^\mu}\right)^N \frac{P^\mu}{Q^{\nu-1}} - z\frac{(q-p^{-1})}{h(p,q)}\partial^{\mu,\nu}_{p,q,h},$$

respectively.

The first three (p,q,μ,ν,h)-Rogers–Szegö polynomials are given as follows:

•

$$H_1(z;p,q,\mu,\nu,h) = H_0\left(\frac{q^\nu}{p^{\mu-1}}z : p,q,\mu,\nu,h\right) + z\frac{q^{(\nu-1)0}}{p^{\mu 0}}H_0\left(\frac{p^\nu}{q^{\nu-1}}z; p,q,\mu,\nu,h\right) - z\frac{q^{\nu n}}{p^{\mu 0}}(p^0 - q^{-0})H_{-1}(z;p,q,\mu,\nu,h).$$

which gives

$$H_1(z; p, q, \mu, \nu, h) = z + 1.$$

•

$$\begin{aligned} H_2(z; p, q, \mu, \nu, h) = H_1\left(\frac{q^\nu}{p^{\mu-1}} z : p, q, \mu, \nu, h\right) \\ + z\frac{q^{(\nu-1)}}{p^\mu} H_1\left(\frac{p^\mu}{q^{\nu-1}} z; p, q, \mu, \nu, h\right) \\ - z\frac{q^\nu}{p^\mu}(p - q^{-1}) H_0(z; p, q, \mu, \nu, h). \end{aligned}$$

Since

$$H_1\left(\frac{q^\nu}{p^{\mu-1}} z : p, q, \mu, \nu, h\right) = 1 + \frac{q^\nu}{p^{\mu-1}} z$$

$$\begin{aligned} z\frac{q^{(\nu-1)}}{p^\mu} H_1\left(\frac{p^\mu}{q^{\nu-1}} z; p, q, \mu, \nu, h\right) &= z\frac{q^{(\nu-1)}}{p^\mu}\Big(1 + \frac{p^\mu}{q^{\nu-1}}\Big) \\ &= z\frac{q^{(\nu-1)}}{p^\mu} + z^2 \end{aligned}$$

and

$$-z\frac{q^\nu}{p^\mu}(p - q^{-1}) H_0(z; p, q, \mu, \nu, h) = -z\frac{q^\nu}{p^\mu}(p - q^{-1})$$

therefore,

$$H_2(z; p, q, \mu, \nu, h) = z^2 + \Big(\frac{q^\nu}{p^{\mu-1}} + \frac{q^{(\nu-1)}}{p^\mu} - \frac{q^\nu}{p^\mu}(p - q^{-1})\Big) z + 1.$$

•

$$\begin{aligned} H_3(z; p, q, \mu, \nu, h) = H_2\left(\frac{q^\nu}{p^{\mu-1}} z : p, q, \mu, \nu, h\right) \\ + z\frac{q^{2(\nu-1)}}{p^{2\mu}} H_2\left(\frac{p^\mu}{q^{\nu-1}} z; p, q, \mu, \nu, h\right) \\ - z\frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2}) H_1(z; p, q, \mu, \nu, h). \end{aligned}$$

But

$$H_2\left(\frac{q^{\nu}}{p^{\mu-1}}z : p, q, \mu, \nu, h\right) = \frac{q^{2\nu}}{p^{2(\mu-1)}} + \Big(\frac{q^{2\nu}}{p^{2(\mu-1)}} + \frac{q^{2\nu-1}}{p^{2\mu-1}} - \frac{q^{2\nu}}{p^{2\mu-1}}(p - q^{-1})\Big)z + 1$$

$$z\frac{q^{2(\nu-1)}}{p^{2\mu}}H_2\left(\frac{p^{\mu}}{q^{\nu-1}}z; p, q, \mu, \nu, h\right) = z^3 + 2\frac{q^{2(\nu-1)}}{p^{2\mu}}z^2 + \frac{q^{2(\nu-1)}}{p^{2\mu}}z$$

and

$$-z\frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2})H_1(z; p, q, \mu, \nu, h) = -z\frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2}) - z^2\frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2}),$$

therefore,

$$\begin{aligned}H_3(z; p, q, \mu, \nu, h) = z^3 &+ \Big(\frac{q^{2\nu}}{p^{2(\mu-1)}} + 2\frac{q^{2(\nu-1)}}{p^{2\mu}} - \frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2})\Big)z^2 + \Big(\frac{q^{2\nu}}{p^{2(\mu-1)}} \\ &+ \frac{q^{2\nu-1}}{p^{2\mu-1}} - \frac{q^{2\nu}}{p^{2\mu-1}}(p - q^{-1}) \\ &- \frac{q^{2\nu}}{p^{2\mu}}(p^2 - q^{-2}) + \frac{q^{2(\nu-1)}}{p^{2\mu}}\Big)z + 1.\end{aligned}$$

5 Continuous $(\mathcal{R}, p, q)$-Hermite Polynomials

We exploit here the peculiar relation established in the theory of q-deformation between Rogers–Szegö polynomials and Hermite polynomials [16, 18, 22, 23] and given by

$$\mathbb{H}_n(\cos\theta; q) = e^{in\,\theta}H_n(e^{-2i\,\theta}; q) = \sum_{k=0}^{n}\begin{bmatrix} n \\ k \end{bmatrix}_q e^{i(n-2k)\theta}, \quad n = 0, 1, 2, \cdots,$$

where $\mathbb{H}_n$ and H_n stand for the Hermite and Rogers–Szegö polynomials, respectively. Is also of interest the property that all the q-Hermite polynomials can be explicitly recovered from the initial one $\mathbb{H}_0(\cos\theta; q) = 1$, using the three-term recursion relation

$$\mathbb{H}_{n+1}(\cos\theta; q) = 2\cos\theta\mathbb{H}_n(\cos\theta; q) - (1-q^n)\mathbb{H}_{n-1}(\cos\theta; q) \tag{5.1}$$

with $\mathbb{H}_{-1}(\cos\theta; q) = 0$.

5.1 Generalization

In the same way we define the $(\mathcal{R}, p, q)$-Hermite polynomials through the $(\mathcal{R}, p, q)$- Rogers–Szegö polynomials as

$$\mathbb{H}_n(\cos\theta; \mathcal{R}, p, q) = e^{in\,\theta} H_n(e^{-2i\,\theta}; \mathcal{R}, p, q), \quad n = 0, 1, 2, \cdots.$$

Then the next statement is true.

Proposition 5.1 *Under the hypotheses of Theorem 3.3, the continuous $(\mathcal{R}, p, q)$-Hermite polynomials satisfy the following three-term recursion relation*

$$\begin{aligned}\mathbb{H}_{n+1}(\cos\theta; \mathcal{R}, p, q) &= e^{i\,\theta}\phi_1^{\frac{n}{2}}(p,q)\phi_1(P,Q)\mathbb{H}_n(\cos\theta; \mathcal{R}, p, q)\\ &+ e^{-i\,\theta}\phi_2^{\frac{n}{2}}(p,q)\phi_2^{-1}(P,Q)\mathbb{H}_n(\cos\theta; \mathcal{R}, p, q)\\ &- \phi_3(p,q)[n]_{\mathcal{R},p,q}\mathbb{H}_{n-1}(\cos\theta; \mathcal{R}, p, q).\end{aligned}$$

Proof Multiplying the two sides of the three-term recursion relation (3.6) by $e^{i(n+1)\theta}$, we obtain, for $z = e^{-2i\theta}$,

$$\begin{aligned}e^{i(n+1)\theta}H_{n+1}(e^{-2i\theta}; \mathcal{R}, p, q) &= e^{i(n+1)\theta}H_n\big(\phi_1(p,q)e^{-2i\theta}; R, p, q\big)\\ &+ e^{i(n-1)\theta}\phi_2^n(p,q)H_n\big(\phi_2^{-1}(p,q)e^{-2i\theta}; \mathcal{R}, p, q\big)\\ &- e^{i(n-1)\theta}\phi_3(p,q)[n]_{\mathcal{R},p,q}H_{n-1}\big(e^{-2i\theta}; \mathcal{R}, p, q\big)\\ &= e^{i\theta}e^{in\theta}\phi_1(P,Q)H_n\left(e^{-2i\theta}; \mathcal{R}, p, q\right)\\ &+ e^{-i\theta}\phi_2^n(p,q)e^{in\theta}\phi_2^{-1}(P,Q)H_n\big(e^{-2i\theta}; \mathcal{R}, p, q\big)\\ &- \phi_3(p,q)[n]_{\mathcal{R},p,q}e^{i(n-1)\theta}H_{n-1}\big(e^{-2i\theta}; \mathcal{R}, p, q\big)\end{aligned}$$

The required result follows from the use of the equalities

$$e^{in\theta}\phi_1(P,Q)H_n\left(e^{-2i\theta};\mathcal{R},p,q\right)=\phi_1^{\frac{n}{2}}(p,q)\phi_1(P,Q)e^{in\theta}H_n\left(e^{-2i\theta};\mathcal{R},p,q\right),$$

$$e^{in\theta}\phi_2^{-1}(P,Q)H_n\left(e^{-2i\theta};\mathcal{R},p,q\right)=\phi_2^{-\frac{n}{2}}(P,Q)\phi_2^{-1}(P,Q)\times e^{in\theta}H_n\left(e^{-2i\theta};\mathcal{R},p,q\right)$$

with

$$\phi_j(P,Q)e^{-2ik\theta}=\phi_j^k(p,q)e^{-2ik\theta},\qquad j=1,\ 2,\ k=0,1,2,\cdots.$$

□

5.2 *Particular Cases*

5.2.1 Continuous (p,q)-Hermite Polynomials

The continuous (p,q)-Hermite polynomials have been already suggested in [18] without any further details. In the above achieved generalization, these polynomials are given by

$$\mathbb{H}_n(\cos\theta;p,q)=e^{in\theta}H_n(e^{-2i\theta};p,q)=\sum_{k=0}^{n}\begin{bmatrix}n\\k\end{bmatrix}_{p,q}e^{i(n-2k)\theta},$$

where $n=0,1,2,\cdots$ Since for the (p,q)-deformation $\phi_1(x,y)=\phi_2(x,y)=x$ and $\phi_(x,y)=x-y$, from Proposition 5.1 we deduce that the corresponding sequence of continuous (p,q)-polynomials satisfies the three-term recursion relation

$$\mathbb{H}_{n+1}(\cos\theta;p,q)=p^{\frac{n}{2}}(e^{i\theta}P+e^{-i\theta}P^{-1})\mathbb{H}_n(\cos\theta;p,q)-(p^n-q^n)\mathbb{H}_{n-1}(\cos\theta;p,q),\tag{5.2}$$

with $Pe^{i\theta}=p^{-1/2}e^{i\theta}$. This relation turns to be the well-known three-term recursion relation (5.1) of continuous q-Hermite polynomials in the limit $p\to 1$. As matter of illustration, let us explicitly compute the first three polynomials using the relation (5.2), with $\mathbb{H}_{-1}(\cos\theta;p,q)=0$ and $\mathbb{H}_0(\cos\theta;p,q)=1$:

$$\begin{aligned}\mathbb{H}_1(\cos\theta;p,q)&=p^0(e^{i\theta}P+e^{-i\theta}P^{-1})-(p^0-q^0)0\\&=e^{i\theta}+e^{-i\theta}=2\cos\theta\end{aligned}$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{p,q} e^{i\theta} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{p,q} e^{-i\theta}.$$

$$\begin{aligned}
\mathbb{H}_2(\cos\theta; p, q) &= p^{\frac{1}{2}}(e^{i\theta}P + e^{-i\theta}P^{-1})(e^{i\theta} + e^{-i\theta}) - p + q \\
&= e^{2i\theta} + e^{-2i\theta} + p + q \\
&= 2\cos 2\theta + p + q \\
&= \begin{bmatrix} 2 \\ 0 \end{bmatrix}_{p,q} e^{2i\theta} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{p,q} e^{0i\theta} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}_{p,q} e^{-i\theta}..
\end{aligned}$$

$$\begin{aligned}
\mathbb{H}_3(\cos\theta; p, q) &= p(e^{i\theta}P + e^{-i\theta}P^{-1})(e^{2i\theta} + e^{-2i\theta} + p + q) \\
&- (p^2 - q^2)(e^{i\theta} + e^{-i\theta}) \\
&= e^{3i\theta} + e^{-3i\theta} + (p^2 + pq + q^2)(e^{i\theta} + e^{-i\theta}) \\
&= 2\cos 3\theta + 2(p^2 + pq + q^2)\cos\theta \\
&= \begin{bmatrix} 3 \\ 0 \end{bmatrix}_{p,q} e^{3i\theta} + \begin{bmatrix} 3 \\ 1 \end{bmatrix}_{p,q} e^{i\theta} \\
&+ \begin{bmatrix} 3 \\ 2 \end{bmatrix}_{p,q} e^{-i\theta} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}_{p,q} e^{-3i\theta}.
\end{aligned}$$

We compute the first three (p, q)-Rogers–Szegö polynomials using the relation (5.2). We get,

- $$\begin{aligned}
H_1(e^{-2i\theta}; p, q) &= e^{-i\theta}\mathbb{H}_1(\cos\theta; p, q) \\
&= 2\,\cos\theta\, e^{-i\theta} \\
&= (e^{i\theta} + e^{-i\theta})\, e^{-i\theta} \\
&= e^{-2i\theta} + 1.
\end{aligned}$$

- $$\begin{aligned}
H_2(e^{-2i\theta}; p, q) &= e^{-2i\theta}\mathbb{H}_2(\cos\theta; p, q) \\
&= e^{-2i\theta}\big(2\cos 2\theta + p + q\big) \\
&= e^{-2i\theta}\big(e^{2i\theta} + e^{-2i\theta} + p + q\big) \\
&= e^{-4i\theta} + (p + q)e^{-2i\theta} + 1.
\end{aligned}$$

•

$$
\begin{aligned}
H_3(e^{-2i\theta}; p, q) &= e^{-3i\theta}\mathbb{H}_3(\cos\theta; p, q) \\
&= e^{-3i\theta}\big(2\cos 3\theta + 2(p^2 + pq + q^2)\cos\theta\big) \\
&= e^{-3i\theta}\big(e^{-3i\theta} + e^{3i\theta} + (p^2 + pq + q^2)(e^{-i\theta} + e^{i\theta})\big) \\
&= e^{-6i\theta} + (p^2 + pq + q^2)e^{-4i\theta} \\
&\qquad\qquad + (p^2 + pq + q^2)e^{-2i\theta} + 1.
\end{aligned}
$$

5.2.2 Continuous (p^{-1}, q)-Hermite Polynomials

The continuous (p^{-1}, q)-Hermite polynomials is obtained by putting $p = p^{-1}$ and $P = P^{-1}$ in Sect. 5.2.1. In the generalization, these polynomials are given by

$$
\mathbb{H}_n(\cos\theta; p^{-1}, q) = e^{in\theta}H_n(e^{-2i\theta}; p^{-1}, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p^{-1},q} e^{i(n-2k)\theta},
$$

where $n = 0, 1, 2, \cdots$

Since for the (p^{-1}, q)-deformation $\phi_1(x, y) = \phi_2(x, y) = x$ and $\phi_(x, y) = x^{-1} - y$, from Proposition 5.1 we deduce that the corresponding sequence of continuous (p^{-1}, q)-polynomials satisfies the three-term recursion relation

$$
\begin{aligned}
\mathbb{H}_{n+1}(\cos\theta; p^{-1}, q) &= p^{-\frac{n}{2}}(e^{i\theta}P^{-1} + e^{-i\theta}P)\mathbb{H}_n(\cos\theta; p^{-1}, q) \\
&\quad - (p^{-n} - q^n)\mathbb{H}_{n-1}(\cos\theta; p^{-1}, q),
\end{aligned} \tag{5.3}
$$

with $P^{-1}e^{i\theta} = p^{1/2}e^{i\theta}$. This relation turns to be the well-known three-term recursion relation (5.1) of continuous q-Hermite polynomials in the limit $p^{-1} \to 1$. As matter of illustration, let us explicitly compute the first three polynomials using the relation (5.3), with $\mathbb{H}_{-1}(\cos\theta; p^{-1}, q) = 0$ and $\mathbb{H}_0(\cos\theta; p^{-1}, q) = 1$:

$$
\begin{aligned}
\mathbb{H}_1(\cos\theta; p^{-1}, q) &= p^0(e^{i\theta}P^{-1} + e^{-i\theta}P - (p^0 - q^0)0 \\
&= e^{i\theta} + e^{-i\theta} = 2\cos\theta \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{p^{-1},q} e^{i\theta} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{p^{-1},q} e^{-i\theta}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{H}_2(\cos\theta; p^{-1}, q) &= p^{-\frac{1}{2}}(e^{i\theta}P^{-1} + e^{-i\theta}P)(e^{i\theta} + e^{-i\theta}) - p^{-1} + q \\
&= e^{2i\theta} + e^{-2i\theta} + p^{-1} + q \\
&= 2\cos 2\theta + p^{-1} + q
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{bmatrix} 2 \\ 0 \end{bmatrix}_{p^{-1},q} e^{2i\theta} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{p^{-1},q} e^{0i\theta} \\
&+ \begin{bmatrix} 2 \\ 2 \end{bmatrix}_{p^{-1},q} e^{-i\theta}..
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{H}_3(\cos\theta;\, p^{-1}, q) &= p^{-1}(e^{i\theta}P^{-1} + e^{-i\theta}P)(e^{2i\theta} + e^{-2i\theta} + p^{-1} + q) \\
&- (p^{-2} - q^2)(e^{i\theta} + e^{-i\theta}) \\
&= e^{3i\theta} + e^{-3i\theta} + (p^{-2} + p^{-1}q + q^2)(e^{i\theta} + e^{-i\theta}) \\
&= 2\cos 3\theta + 2(p^{-2} + p^{-1}q + q^2)\cos\theta \\
&= \begin{bmatrix} 3 \\ 0 \end{bmatrix}_{p^{-1},q} e^{3i\theta} + \begin{bmatrix} 3 \\ 1 \end{bmatrix}_{p^{-1},q} e^{i\theta} \\
&+ \begin{bmatrix} 3 \\ 2 \end{bmatrix}_{p^{-1},q} e^{-i\theta} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}_{p^{-1},q} e^{-3i\theta}.
\end{aligned}
$$

From the relation (5.2), we compute the first three (p^{-1}, q)-Rogers–Szegö polynomials. We get,

- $$
\begin{aligned}
H_1(e^{-2i\theta};\, p^{-1}, q) &= e^{-i\theta}\mathbb{H}_1(\cos\theta;\, p^{-1}, q) \\
&= 2\,\cos\theta\, e^{-i\theta} \\
&= (e^{i\theta} + e^{-i\theta})\, e^{-i\theta} \\
&= e^{-2i\theta} + 1.
\end{aligned}
$$

- $$
\begin{aligned}
H_2(e^{-2i\theta};\, p^{-1}, q) &= e^{-2i\theta}\mathbb{H}_2(\cos\theta;\, p^{-1}, q) \\
&= e^{-2i\theta}\big(2\cos 2\theta + p^{-1} + q\big) \\
&= e^{-2i\theta}\big(e^{2i\theta} + e^{-2i\theta} + p^{-1} + q\big) \\
&= e^{-4i\theta} + (p^{-1} + q)e^{-2i\theta} + 1.
\end{aligned}
$$

- $$
\begin{aligned}
H_3(e^{-2i\theta};\, p^{-1}, q) &= e^{-3i\theta}\mathbb{H}_3(\cos\theta;\, p^{-1}, q) \\
&= e^{-3i\theta}\big(2\cos 3\theta + 2(p^{-2} + p^{-1}q + q^2)\cos\theta\big) \\
&= e^{-3i\theta}\big(e^{-3i\theta} + e^{3i\theta} + (p^{-2} + p^{-1}q + q^2)(e^{-i\theta} + e^{i\theta})\big)
\end{aligned}
$$

$$= e^{-6i\theta} + (p^{-2} + p^{-1}q + q^2)e^{-4i\theta} \\ +(p^{-2} + p^{-1}q + q^2)e^{-2i\theta} + 1.$$

5.2.3 Continuous (p, q)-Hermite Polynomials Related to (p, q)-Generalization of Quesne Deformation [13]

We define the continuous (p, q)-Hermite polynomials corresponding to the (p, q)-generalization of Quesne deformation as follows:

$$\mathbb{H}_n^Q(\cos\theta; p, q) = e^{in\theta} H_n^Q(e^{-2i\theta}; p, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q}^{Q} e^{i(n-2k)\theta}$$

where $n = 0, 1, 2, \cdots$ Since for the (p, q)-generalization of Quesne deformation [13] $\phi_1(x, y) = \phi_2(x, y) = y$ and $\phi_(x, y) = y - x^{-1}$, from Proposition 5.1 we deduce that the corresponding sequence of continuous (p, q)-Hermite polynomials satisfies the three-term recursion relation

$$\mathbb{H}_{n+1}^Q(\cos\theta; p, q) = p^{\frac{n}{2}}(e^{i\theta}P + e^{-i\theta}P^{-1})\mathbb{H}_n^Q(\cos\theta; p, q) \\ - (p^n - q^{-n})\mathbb{H}_{n-1}^Q(\cos\theta; p, q) \tag{5.4}$$

with $Pe^{i\theta} = p^{-1/2}e^{i\theta}$. This relation turns to be the well-known three-term recursion relation (5.1) of continuous q-Hermite polynomials in the limit $p \to 1$. As matter of illustration, let us explicitly compute the first three polynomials using the relation (5.4), with $\mathbb{H}_{-1}^Q(\cos\theta; p, q) = 0$ and $\mathbb{H}_0^Q(\cos\theta; p, q) = 1$:

$$\begin{aligned}
\mathbb{H}_1^Q(\cos\theta; p, q) &= p^0(e^{i\theta}P + e^{-i\theta}P^{-1}) - (p^0 - q^0)0 \\
&= e^{i\theta} + e^{-i\theta} = 2\cos\theta \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{p,q}^{Q} e^{i\theta} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{p,q}^{Q} e^{-i\theta}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{H}_2^Q(\cos\theta; p, q) &= p^{\frac{1}{2}}(e^{i\theta}P + e^{-i\theta}P^{-1})(e^{i\theta} + e^{-i\theta}) - p + q^{-1} \\
&= e^{2i\theta} + e^{-2i\theta} + p + q^{-1} \\
&= 2\cos 2\theta + p + q^{-1} \\
&= \begin{bmatrix} 2 \\ 0 \end{bmatrix}_{p,q}^{Q} e^{2i\theta} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{p,q}^{Q} e^{0i\theta} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}_{p,q}^{Q} e^{-i\theta}..
\end{aligned}$$

$$\mathbb{H}_3^Q(\cos\theta; p, q) = p(e^{i\theta}P + e^{-i\theta}P^{-1})(e^{2i\theta} + e^{-2i\theta} + p + q^{-1})$$

$$
\begin{aligned}
&- (p^2 - q^{-2})(e^{i\theta} + e^{-i\theta})\\
&= e^{3i\theta} + e^{-3i\theta} + (p^2 + pq^{-1} + q^{-2})(e^{i\theta} + e^{-i\theta})\\
&= 2\cos 3\theta + 2(p^2 + pq^{-1} + q^{-2})\cos\theta\\
&= \begin{bmatrix} 3 \\ 0 \end{bmatrix}^{Q}_{p,q} e^{3i\theta} + \begin{bmatrix} 3 \\ 1 \end{bmatrix}^{Q}_{p,q} e^{i\theta}\\
&+ \begin{bmatrix} 3 \\ 2 \end{bmatrix}^{Q}_{p,q} e^{-i\theta} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}^{Q}_{p,q} e^{-3i\theta}.
\end{aligned}
$$

Let us compute the first three (p, q)-Rogers–Szegö polynomials associated with the q-Quesne quantum algebra using the relation (5.2). We get,

- $$
\begin{aligned}
H_1^Q(e^{-2i\theta}; p, q) &= e^{-i\theta}\mathbb{H}_1^Q(\cos\theta; p, q)\\
&= 2\,\cos\theta\, e^{-i\theta}\\
&= (e^{i\theta} + e^{-i\theta})\, e^{-i\theta}\\
&= e^{-2i\theta} + 1.
\end{aligned}
$$

- $$
\begin{aligned}
H_2^Q(e^{-2i\theta}; p, q) &= e^{-2i\theta}\mathbb{H}_2^Q(\cos\theta; p, q)\\
&= e^{-2i\theta}\big(2\cos 2\theta + p + q^{-1}\big)\\
&= e^{-2i\theta}\big(e^{2i\theta} + e^{-2i\theta} + p + q^{-1}\big)\\
&= e^{-4i\theta} + (p + q^{-1})e^{-2i\theta} + 1.
\end{aligned}
$$

- $$
\begin{aligned}
H_3^Q(e^{-2i\theta}; p, q) &= e^{-3i\theta}\mathbb{H}_3^Q(\cos\theta; p, q)\\
&= e^{-3i\theta}\big(2\cos 3\theta + 2(p^2 + pq^{-1} + q^{-2})\cos\theta\big)\\
&= e^{-3i\theta}\big(e^{-3i\theta} + e^{3i\theta} + e^{-i\theta}(p^2 + pq^{-1} + q^{-2})\\
&\qquad\qquad + e^{i\theta}(p^2 + pq^{-1} + q^{-2}))\big)\\
&= e^{-6i\theta} + (p^2 + pq^{-1} + q^{-2})e^{-4i\theta}\\
&\qquad\qquad + (p^2 + pq^{-1} + q^{-2})e^{-2i\theta} + 1.
\end{aligned}
$$

5.2.4 Continuous (p, q, μ, ν, h)-Hermite Polynomials Related to the Hounkonnou-Ngompe Deformation [15]

The continuous (p, q, μ, ν, h)-Hermite polynomials are defined by:

$$\begin{aligned}\mathbb{H}_n(\cos\theta; p, q, \mu, \nu, h) &= e^{in\theta} H_n(e^{-2i\theta}; p, q, \mu, \nu, h)\\ &= \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_{p,q,h}^{\mu,\nu} e^{i(n-2k)\theta}, \quad n = 0, 1, 2, \cdots .\end{aligned}$$

Since for the (p, q, μ, ν, h)-deformation $\phi_1(x, y) = x^{1-\mu} y^{\nu}$, $\phi_2(x, y) = x^{-\mu} y^{\nu-1}$ and $\phi_3(x, y) = \frac{y - x^{-1}}{h(p,q)}$, from Proposition 5.1 the corresponding sequence of continuous (p, q, μ, ν, h)-Hermite polynomials satisfies the three-term recursion relation

$$\begin{aligned}\mathbb{H}_{n+1}(\cos\theta; p, q, \mu, \nu, h) &= e^{i\theta} \frac{q^{\nu\frac{n}{2}}}{p^{(\mu-1)\frac{n}{2}}} \frac{Q^{\nu}}{P^{\mu-1}} \mathbb{H}_n(\cos\theta; p, q, \mu, \nu, h)\\ &+ e^{-i\theta} \frac{q^{(\nu-1)\frac{n}{2}}}{p^{\mu\frac{n}{2}}} \frac{Q^{-(\nu-1)}}{P^{-\mu}} \mathbb{H}_n(\cos\theta; p, q, \mu, \nu, h)\\ &- (p^n - q^{-n}) \frac{q^{\nu n}}{p^{\mu n}} \mathbb{H}_{n-1}(\cos\theta; p, q, \mu, \nu, h) \qquad (5.5)\end{aligned}$$

where $Pe^{i\theta} = p^{-1/2} e^{i\theta}$. This relation turns to be the well-known three-term recursion relation (5.1) of continuous q-Hermite polynomials in the limit $p \to 1$. As matter of illustration, let us explicitly compute the first two polynomials using the relation (5.5), with $\mathbb{H}_{-1}(\cos\theta; p, q, \mu, \nu, h) = 0$ and $\mathbb{H}_0(\cos\theta; p, q, \mu, \nu, h) = 1$:

- $$\mathbb{H}_1(e^{i\theta}; p, q, \mu, \nu, h) = e^{i\theta} \frac{Q^{\nu}}{P^{\mu-1}} + e^{-i\theta} \frac{Q^{-(\nu-1)}}{P^{-\mu}}$$

- $$\begin{aligned}\mathbb{H}_2(e^{i\theta}; p, q, \mu, \nu, h) &= \big(e^{i\theta} \frac{q^{\nu\frac{1}{2}}}{p^{(\mu-1)\frac{1}{2}}} \frac{Q^{\nu}}{P^{\mu-1}} + e^{-i\theta} \frac{q^{(\nu-1)\frac{1}{2}}}{p^{\mu\frac{1}{2}}} \frac{Q^{-(\nu-1)}}{P^{-\mu}}\big)\\ &\times \big(e^{i\theta} \frac{Q^{\nu}}{P^{\mu-1}} + e^{-i\theta} \frac{Q^{-(\nu-1)}}{P^{-\mu}}\big) - (p - q^{-1}) \frac{q^{\nu}}{p^{\mu}}\end{aligned}$$

$$= e^{2i\theta}\frac{q^{\frac{\nu}{2}}}{p^{\frac{\mu-1}{2}}}\frac{Q^{2\nu}}{P^{2(\mu-1)}} + e^{-2i\theta}\frac{Q^{-2(\nu-1)}}{P^{-2\mu}}\frac{q^{\frac{\nu-1}{2}}}{p^{\frac{\mu}{2}}}$$
$$+ Q\,P\frac{q^{\frac{\nu}{2}}}{p^{\frac{\mu}{2}}}(p^{\frac{1}{2}} + q^{\frac{-1}{2}}) - (p - q^{-1})\frac{q^{\nu}}{p^{\mu}}.$$

The first two (p, q, μ, ν, h)-Rogers–Szegö polynomials related to the Hounkonnou-Ngompe deformation. We get,

-
$$\begin{aligned} H_1(e^{-2i\theta}; p, q, \mu, \nu, h) &= e^{-i\theta}\mathbb{H}_1(\cos\theta; p, q, \mu, \nu, h) \\ &= e^{-i\theta}\Big(e^{i\theta}\frac{Q^{\nu}}{P^{\mu-1}} + e^{-i\theta}\frac{Q^{-(\nu-1)}}{P^{-\mu}}\Big) \\ &= e^{-2i\theta}\frac{Q^{-(\nu-1)}}{P^{-\mu}} + \frac{Q^{\nu}}{P^{\mu-1}}. \end{aligned}$$

-
$$\begin{aligned} H_2(e^{-2i\theta}; p, q, \mu, \nu, h) &= e^{-2i\theta}\mathbb{H}_2(\cos\theta; p, q, \mu, \nu, h) \\ &= e^{-4i\theta}\frac{Q^{-2(\nu-1)}}{P^{-2\mu}}\frac{q^{\frac{\nu-1}{2}}}{p^{\frac{\mu}{2}}} + e^{-2i\theta}\Big(- (p - q^{-1})\frac{q^{\nu}}{p^{\mu}} \\ &\quad + Q\,P\frac{q^{\frac{\nu}{2}}}{p^{\frac{\mu}{2}}}(p^{\frac{1}{2}} + q^{\frac{-1}{2}})\Big) + \frac{q^{\frac{\nu}{2}}}{p^{\frac{\mu-1}{2}}}\frac{Q^{2\nu}}{P^{2(\mu-1)}}. \end{aligned}$$

6 Concluding Remarks

In this contribution paper, we have presented the construction of Rogers–Szegö polynomials, resulting from the works existing in the literature, and recently introduced $(\mathcal{R}, p, q)$-deformed Rogers–Szegö polynomials, which generalize known usual and deformed Rogers–Szegö polynomials. Their three-term recursion relation and difference equation have been provided, together with the induced $(\mathcal{R}, p, q)$-deformed quantum algebras. Relevant particular cases and examples have been given. Finally, the associated continuous $(\mathcal{R}, p, q)$-Hermite polynomials have also been characterized.

Acknowledgements This work is supported by TWAS Research Grant RGA No. 17-542 RG/MATHS/AF/AC_G-FR3240300147. The ICMPA-UNESCO Chair is in partnership with Daniel Iagolnitzer Foundation (DIF), France, and the Association pour la Promotion Scientifique de l'Afrique (APSA), supporting the development of mathematical physics in Africa. I am grateful to my students, Fridolin Melong and Cyrille Essossolim Haliya, who devoted their Christmas day to carefully read this manuscript and check the details.

References

1. W.A. Al-Salam, L. Carlitz, A q-analog of a formula of Toscano. Boll. Unione Matem. Ital. **12**, 414–417 (1957)
2. M. Arik, D. Coon, Hilbert spaces of analytic functions and generated coherent states. J. Math. Phys. **17**, 424–427 (1976)
3. N.M. Atakishiyev, S.K. Suslov, Difference analogs of the harmonic oscillator. Theor. Math. Phys. **85**, 1055–1062 (1990)
4. J.D. Bukweli Kyemba, M.N. Hounkonnou, Characterization of $(\mathcal{R}, p, q)$-deformed Rogers-Szegö polynomials: associated quantum algebras, deformed Hermite polynomials and relevant properties. J. Phys. A Math. Theor. **45**, 225204 (2012)
5. I.M. Burban, A.U. Klimyk, P, Q-differentiation, P, Q-integration, and P, Q-hypergeometric functions related to quantum groups. Integr. Transform. Spec. Funct. **2**, 15 (1994)
6. R. Chakrabarti, R. Jagannathan, A (p, q)-oscillator realization of two-parameter quantum algebras. J. Phys. A Math. Gen. **24**, L711 (1991)
7. V. Chari, A. Pressley, *A Guide to Quantum Groups* (Cambridge University Press, Cambridge, 1994)
8. P. Feinsilver, Lie algebras and recursion relations III: q-analogs and quantized algebras. Acta Appl. Math. **19**, 207–251 (1990)
9. R. Floreanini, L. Lapointe, L. Vinet, A note on (p, q)-oscillators and bibasic hypergeometric functions. J. Phys. A Math. Gen. **26**, L611–L614 (1993)
10. D. Galetti, A realization of the q-deformed harmonic oscillator: Rogers–Szegö and Stieltjes–Wigert polynomials. Braz. J. Phys. **33**(1), 148–157 (2003)
11. G. Gasper, M. Rahman, *Basic Hypergeometric Series* (Cambridge University Press, Cambridge, 1990)
12. I.M. Gelfand, M.I. Graev, L. Vinet, (r, s)-hypergeometric functions of one variable. Russ. Acad. Sci. Dokl. Math. **48**, 591 (1994)
13. M.N. Hounkonnou, E.B. Ngompe Nkouankam, On $(p, q, \mu, \nu, \phi_1, \phi_2)$ generalized oscillator algebra and related bibasic hypergeometric functions. J. Phys. A Math. Theor. **40**, 883543 (2007)
14. M.N. Hounkonnou, J.D. Bukweli Kyemba, Generalized $(\mathcal{R}, p, q)$-deformed Heisenberg algebras: coherent states and special functions. J. Math. Phys. **51**, 063518 (2010)
15. M.N. Hounkonnou, E.B. Ngompe Nkouankam, New (p, q, μ, ν, f)-deformed states. J. Phys. A Math. Theor. **40**, 12113 (2007)
16. E.H. Ismail Mourad, *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
17. R. Jagannathan, R. Sridhar, (p, q)-Rogers–Szegö polynomials and the (p, q)-oscillator, in *The Legacy of Alladi Ramakrishnan in the Mathematical Sciences* (Springer, New York, 2010), pp. 491–501
18. R. Jagannathan, K. Srinivasa Rao, Two-parameter quantum algebras, twin-basic numbers, and associated generalized hypergeometric series (2006). arXiv:math/0602613
19. F.H. Jackson, On q-functions and a certain difference operator. Trans. R. Soc. Edin. **46**, 253–281 (1908)
20. F.H. Jackson, On q-definite integrals. Q. J. Pure Appl. Math. **41**, 193–203 (1910)
21. M. Jimbo, A q-difference analogue of $U(g)$ and the Yang-Baxter equation. Lett. Math. Phys. **10**, 63–69 (1985)
22. A. Klimyk, K. Schmudgen, *Quantum Groups and their Representation* (Springer, Berlin, 1997)

23. R. Koekoek, R.F. Swarttouw, The Askey-scheme of orthogonal polynomials and its q-analogue. TUDelft Report No. 98-17, 1998
24. A. Odzijewicz, Quantum algebras and q-special functions related to coherent states maps of the disc. Commun. Math. Phys. **192**, 183–215 (1998)
25. C. Quesne, K.A. Penson, V.M. Tkachuk, Maths-type q-deformed coherent states for $q > 1$. Phys. Lett. A **313**, 29–36 (2003)
26. M.P. Schützenberger, Une interpretation de certaines solutions de l'equation fonctionnelle: $F(x + y)F(x) + F(y)$. C. R. Acad. Sci. Paris **236**, 352–353 (1953)
27. G. Szegö, *Collected Papers*, in ed. by R. Askey, vol. 1 (Birkäuser, Basel, 1982), pp. 793–805
28. G. Szegö, in *Orthogonal Polynomials*. Colloquium Publications, vol. 23 (American Mathematical Society, Providence, 1991)

Zeros of Orthogonal Polynomials

Kerstin Jordaan

Abstract In this lecture we discuss properties of zeros of orthogonal polynomials. We review properties that have been used to derive bounds for the zeros of orthogonal polynomials. Topics to be covered include Markov's theorem on monotonicity of zeros and its generalisations, the proof of a conjecture by Askey and its extensions, interlacing properties of zeros, Sturm's comparison theorem and convexity of zeros.

Keywords Orthogonal polynomials · Zeros · Jacobi polynomials · Monotonicity of zeros · Interlacing of zeros · Stieltjes interlacing · Bounds for zeros

Mathematics Subject Classification (2000) Primary 33C45; Secondary 42C05

1 Zeros as Eigenvalues

Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of monic orthogonal polynomials satisfying

$$p_{n+1}(x) = (x - \alpha_n)p_n(x) - \beta_n\, p_{n-1}(x), \quad n = 0, 1, 2, \ldots,$$

with $p_{-1} = 0$ and $p_0 = 1$.

The recurrence coefficients may be collected in a tridiagonal matrix of the form

$$J = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \\ & & \sqrt{\beta_3} & \alpha_3 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}$$

K. Jordaan (✉)
Department of Decision Sciences, University of South Africa, Pretoria, South Africa
e-mail: jordakh@unisa.ac.za

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_17

known as the Jacobi matrix or Jacobi operator. One can write

$$p_n(x) = \det(xI_n - J_n)$$

where I_n is the identity matrix and J_n is the tridiagonal matrix

$$J_n = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & & \\ & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & & \\ & & \sqrt{\beta_3} & \alpha_3 & \ddots & \\ & & & \ddots & \ddots & \\ & & & & & \sqrt{\beta_{n-1}} \\ & & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} \end{pmatrix}$$

It follows that zeros of $p_n(x)$ are the same as the eigenvalues of J_n.

2 Monotonicity of the Zeros

The manner in which the zeros of a polynomial change as the parameter changes can be used to study interlacing properties of zeros [8, 32].

In 1886, A. Markov established an important result about the monotonicity properties of zeros of orthogonal polynomials with respect to a parameter (cf. [33], [43, Thm 6.12.1].

Markov's theorem can be used to show that the zeros of classical orthogonal polynomials like Laguerre and Jacobi polynomials are monotone functions of the parameter(s) involved by using the derivative of the weight function with respect to the parameter(s). A slightly generalised version of Markov's theorem, stated as an exercise in [16, Chap. 3, ex. 15] and proved in [23, Thm 7.1.1] (see also [7, Thm 1]) can also be applied to discrete orthogonal polynomials such as Meixner and Hahn polynomials.

Theorem 2.1 (cf. [23]) *Let* $\{p_n(x,\tau)\}_{n=0}^{\infty}$ *be orthogonal with respect to* $d\alpha(x,\tau) = w(x,\tau)d\alpha(x)$ *on the interval* $[a,b]$ *depending on a parameter* τ, *such that* $w(x,\tau)$ *is positive and continuous for* $a < x < b$, $\tau_1 < \tau < \tau_2$. *Also, suppose that the partial derivative* $w_\tau(x,\tau)$ *for* $a < x < b$, $\tau_1 < \tau < \tau_2$ *exists and is continuous, and the integrals*

$$\int_a^b x^\nu w_\tau(x,\tau)d\alpha(x),\ \nu = 0, 1, 2, \ldots, 2n-1,$$

converge uniformly in every closed interval $[\tau', \tau''] \subset (\tau_1, \tau_2)$. *If the zeros of* $p_n(x,\tau)$ *are denoted by* $b > x_1(\tau) > x_2(\tau) > \cdots > x_n(\tau) > a$, *then the* ν*th zero*

$x_\nu(\tau)$ (for a fixed value of ν) is an increasing (decreasing) function of τ provided that w_τ/w is an increasing (decreasing) function of x, $a < x < b$.

Proof The mechanical quadrature formula (cf. [23, (2.4.1)])

$$\int_a^b \rho(x)d\alpha(x,\tau) = \sum_{\nu=1}^{n} \lambda_\nu(\tau)\rho(x_\nu(\tau)), \tag{2.1}$$

holds for polynomials $\rho(x)$ of degree at most $2n-1$. Differentiating (2.1) with respect to τ, we obtain

$$\int_a^b \rho(x)w_\tau(x,\tau)d\alpha(x) = \sum_{\nu=1}^{n} \lambda_\nu(\tau)\rho^{'}(x_\nu)x_\nu^{'}(\tau) + \sum_{\nu=1}^{n} \lambda_\nu^{'}(\tau)\rho(x_\nu).$$

Now we choose

$$\rho(x) = \frac{\{p_n(x,\tau)\}^2}{x-x_\nu},$$

then, since x_ν is a removable singularity, $\rho^{'}(x_\nu) = \{p_n^{'}(x_\nu,\tau)\}^2$ while $\rho^{'}(x_\mu) = 0$ if $\mu \neq \nu$ and hence

$$\int_a^b w_\tau(x,\tau)\frac{\{p_n(x,\tau)\}^2}{x-x_\nu}d\alpha(x) = \lambda_\nu(\tau)\{p_n^{'}(x_\nu,\tau)\}^2 x_\nu^{'}(\tau). \tag{2.2}$$

In view of the orthogonality the integral

$$\int_a^b \frac{\{p_n(x,\tau)\}^2}{x-x_\nu}w(x,\tau)d\alpha(x) = 0,$$

so (2.2) can be rewritten as

$$\int_a^b \left\{w_\tau(x,\tau) - \frac{w_\tau(x_\nu,\tau)}{w(x_\nu,\tau)}w(x,\tau)\right\}\frac{\{p_n(x,\tau)\}^2}{x-x_\nu}d\alpha(x) = \lambda_\nu(\tau)\{p_n^{'}(x_\nu,\tau)\}^2 x_\nu^{'}(\tau).$$

and we obtain

$$\int_a^b \left\{\frac{w_\tau(x,\tau)}{w(x,\tau)} - \frac{w_\tau(x_\nu,\tau)}{w(x_\nu,\tau)}\right\}\frac{\{p_n(x,\tau)\}^2}{x-x_\nu}d\alpha(x,\tau) = \lambda_\nu(\tau)\{p_n'(x_\nu,\tau)\}^2 x_\nu'(\tau). \tag{2.3}$$

The integrand in (2.3) has a constant sign, so the positivity of the so-called Christoffel numbers $\lambda_\nu(\tau)$ [43, p. 48] establishes the result. □

Example (cf. [23, Thm. 7.1.2])

(i) For Jacobi polynomials $P_n(\alpha, \beta)$, the weight function is $w(x, \alpha, \beta) = (1 - x)^\alpha (1 + x)^\beta$ and $\alpha(x) = x$, hence

$$\frac{\partial \ln \omega(x, \alpha, \beta)}{\partial \beta} = \frac{\partial \ln(1 + x)^\beta}{\partial \beta} = \ln(1 + x)$$

which is an increasing function of x.

(ii) For the Hahn polynomials, α is a step function with unit jumps at $0, 1, \ldots, N$ and

$$w(x, \alpha, \beta) = \frac{\Gamma(\alpha + 1 + x)\Gamma(\beta + 1 + N - x)}{\Gamma(\alpha + 1)\Gamma(b + 1)}$$

and, using properties of the Gamma function, we deduce that

$$\frac{\partial \ln \omega(x, \alpha, \beta)}{\partial \alpha} = \sum_{n=0}^{\infty} \left[\frac{1}{\alpha + n + 1} - \frac{1}{\alpha + n + x + 1} \right]$$

which is a decreasing function of x.

The variation of the zeros of a Jacobi polynomial with the parameter can be summarised as follows.

Lemma 2.2 (cf. [43], Thm 6.21.1, p.121) *Let $\alpha > -1$ and $\beta > -1$ and let x_k, $k = 1, 2, \ldots, n$ denote the zeros of $P_n^{(\alpha,\beta)}$ in increasing order. Then $\frac{dx_k}{d\alpha} < 0$ and $\frac{dx_k}{d\beta} > 0$ for each $k = 1, \ldots, n$.*

This implies that the zeros of $P_n^{(\alpha,\beta)}(x)$ increase as β increases and decrease as α increases.

The application of Markov's Theorem to a specific orthogonal sequence requires the calculation of the derivative of the weight function. Weight functions of orthogonal polynomials are not always simple and do not necessarily satisfy the conditions of Markov's theorem and its generalisations, hence alternative tools are required. Such alternative techniques for deriving monotonicity properties of zeros include the analysis of zeros of polynomial solutions of second-order ordinary linear differential equations (cf. [34] and [23]). The monotonicity of all the zeros as well as the extreme zeros of polynomials satisfying recursion formulas, referred to as birth and death processes, were considered in [20, 26] using a finite dimensional version of the Hellman–Feynman Theorem while Ismail and Muldoon [24] used tridiagonal matrices arising from the three term recurrence relation to study monotonicity properties of various special functions and orthogonal polynomials. Dimitrov and Rodrigues [6] applied the Routh–Hurwitz stability criterion to obtain monotonicity results for the zeros of Jacobi polynomials.

3 Interlacing of Zeros from Different Sequences

The interlacing property of zeros of polynomials is important in applications to numerical quadrature. In [31], Lubinsky generalises a quadrature formula proved by Simon [39] for orthonormal polynomials P of degree $\leq n-2$ by applying Wendroff's Theorem [47] to two real polynomials R and S of consecutive degree with interlacing zeros. He observes that other analytical methods can be used to prove the quadrature results obtained, but that the proofs are substantially simplified when interlacing properties of zeros are used.

Levit [30] was the first to study interlacing properties of zeros of different orthogonal polynomials when he considered the zeros of Hahn polynomials in 1967.

3.1 Jacobi Polynomials

In 1990, Askey [2] proved that the zeros of Jacobi polynomials $P_n^{(\alpha,\beta)}$ and $P_n^{(\alpha+t,\beta)}$, $t \in (0,1]$ interlace for $\alpha, \beta > -1$ and conjectured that the zeros of $P_n^{(\alpha,\beta)}$ and $P_n^{(\alpha+2,\beta)}$ are interlacing for each $n \in \mathbb{N}$, $\alpha,\ \beta > -1$.

In [38], Segura proved that interlacing of zeros holds, under certain assumptions, within sequences of classical orthogonal polynomials even when the parameter(s) on which they depend lie outside the value(s) required to ensure orthogonality.

It was proved in [11] that the zeros of $P_n^{(\alpha,\beta)}$ interlace with the zeros of polynomials from different Jacobi sequences, including those of $P_n^{(\alpha-t,\beta+k)}$ and $P_{n-1}^{(\alpha+t,\beta+k)}$ for $0 \leq t, k \leq 2$, thereby confirming and extending the Askey conjecture [2]. Numerical examples were given to illustrate that, in general, if t or k is greater than 2, interlacing of zeros need not necessarily occur.

The proofs make use of the Markov monotonicity theorem as applied to Jacobi polynomials (see Lemma 2.2) as well as the contiguous relations for ${}_2F_1$ hypergeometric polynomials. Note that various algorithms have been developed for computing such contiguous relations (cf. [36, 37, 44]). In addition, an efficient algorithm for large shifts of the parameters is available as a Maple program (cf. [45]).

In [9], Dimitrov, Ismail and Rafaeli used a general approach to the Askey conjecture by considering interlacing properties of zeros of orthogonal Jacobi polynomials $P_n^{(\alpha,\beta)}$ of the same degree and different parameter values α and β in the context of perturbation of the weight function of orthogonality.

The first result shows that interlacing of the zeros occurs for Jacobi polynomials of adjacent degree when both the parameters α and β are increased by t and k respectively for any $t, k \in (0, 2]$.

Theorem 3.1 (cf. [11]) *Let* $\alpha,\ \beta > -1$ *and let* $t, k \in [0, 2]$. *Let*

$$-1 < x_1 < x_2 < \cdots < x_n < 1 \textit{ be the zeros of } P_n^{(\alpha,\beta)} \textit{ and}$$
$$-1 < t_1 < t_2 < \cdots < t_{n-1} < 1 \textit{ be the zeros of } P_{n-1}^{(\alpha+t,\beta+k)}.$$

Then

$$-1 < x_1 < t_1 < x_2 < \cdots < x_{n-1} < t_{n-1} < x_n < 1.$$

The result is illustrated in Fig. 1 where it is shown that, for a fixed $\alpha > -1$ and $\beta > -1$, the zeros of $P_n^{(\alpha,\beta)}$ and $P_{n-1}^{(\alpha',\beta')}$ interlace when α' and β' are any of the values in the shaded region.

Next we consider Jacobi polynomials of the same degree with different parameters, where one or both parameters α and β increase or decrease by $t \in (0, 2]$.

Theorem 3.2 (cf. [11]) *Let* $\alpha > 1$, $\beta > -1$ *and* $t,\ k \in (0, 2)$.

$$\textit{Let} -1 < x_1 < x_2 < \cdots < x_n < 1, \textit{ be the zeros of } P_n^{(\alpha,\beta)},$$
$$-1 < t_1 < t_2 < \cdots < t_n < 1, \textit{ be the zeros of } P_n^{(\alpha-k,\beta+t)},$$
$$\textit{and} \quad -1 < y_1 < y_2 < \cdots < y_n < 1, \quad \textit{be those of} \quad P_n^{(\alpha-2,\beta+2)}.$$

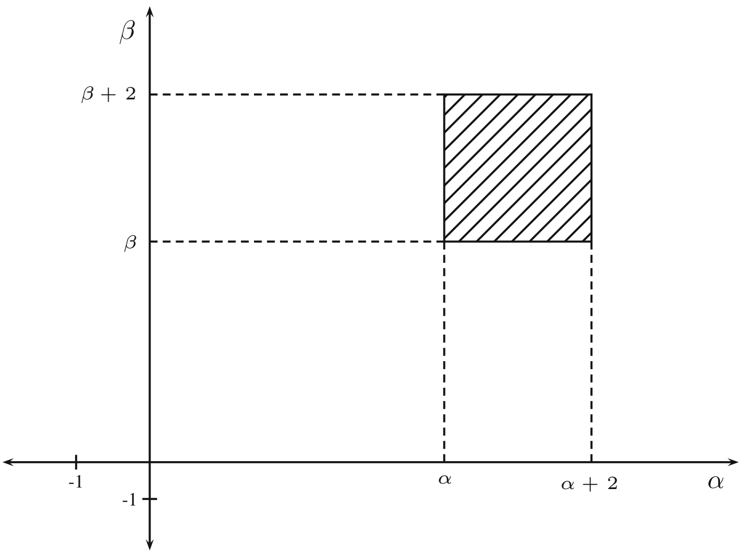


Fig. 1 Values of α' and β' for which the zeros of $P_n^{(\alpha,\beta)}$ and $P_{n-1}^{(\alpha',\beta')}$ interlace when $\alpha, \beta > -1$ are fixed

Then

$$-1 < x_1 < t_1 < y_1 < x_2 < t_2 < y_2 < \cdots < x_n < t_n < y_n < 1.$$

It follows from the symmetry property for Jacobi polynomials (cf. [23, p.82, (4.1.1)])

$$P_n^{(\alpha,\beta)}(x) = (-1)^n P_n^{(\beta,\alpha)}(-x),$$

that an analogous result holds for the zeros of $P_n^{(\alpha,\beta)}$ and $P_n^{(\alpha+k,\beta-t)}$ with $t, k \in (0, 2]$.

The interlacing property for the zeros of $P_n^{(\alpha,\beta)}$ and $P_n^{(\alpha',\beta')}$, for a fixed α and β with $(\alpha, \beta) \neq (\alpha', \beta')$; α, α', β and $\beta' > -1$ holds, is illustrated by the shaded area in the $\alpha\beta$-plane in Fig. 2.

Remark 3.3 Restrictions on the ranges of t and k are required in the theorems since the interlacing property is not retained, in general, when one or both of the parameters α, β are increased by more than 2. Examples illustrating this are provided in [11]. These restrictions on the parameters seem reasonable considering the electrostatic interpretation of the zeros as described in the next section.

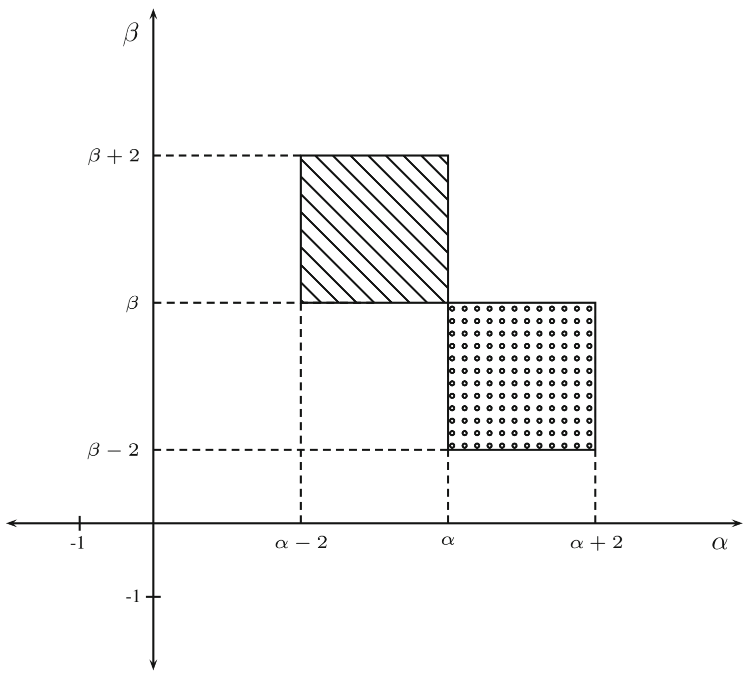


Fig. 2 Values of $\alpha', \beta' > -1$ for which the zeros of $P_n^{(\alpha,\beta)}$ and $P_n^{(\alpha',\beta')}$ interlace when $\alpha, \beta > -1$ are fixed

3.2 A One-Dimensional Electrostatic Model for Zeros of Different Jacobi Polynomials

The electrostatic interpretation of the zeros of Jacobi polynomials, originally given by Stieltjes (cf. [40]), was generalised by Ismail (cf. [21, 22]) to polynomials orthogonal with respect to a weight $w(x) = e^{-v(x)}$. For n fixed, the zeros of the Jacobi polynomial of degree n may be interpreted as the equilibrium positions of n movable positive unit charges that are able to move freely between two positive charges, placed at -1 and 1, with magnitude $\frac{\beta+1}{2}$ and $\frac{\alpha+1}{2}$ respectively where it is assumed that the interaction force between two charges e_i and e_j with a distance s apart is $\frac{2e_ie_j}{s}$. Therefore, increasing a parameter corresponds to increasing a charge at one of the endpoints of orthogonality which explains the monotonicity of the zeros. To evaluate the energy at equilibrium one needs to compute the discriminant of Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ as was done by Stieltjes in [41]. An alternative proof was provided by Hilbert in [18]. For a more detailed exposition of these results, see [23, Chapter 3] or [43, Section 6.7].

Let $n \in \mathbb{N}$, $\alpha > 1$ and $\beta > -1$ be fixed and consider the n fixed equilibrium positions of charges q_i, $i \in \{1, ..., n\}$ coinciding with the zeros of the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$. Replacing the charge at 1 by $\frac{\alpha+1}{2} - \frac{k}{2}$ and the charge at -1 by $\frac{\beta+1}{2} + \frac{t}{2}$, the positive charge at 1 decreases while the charge at -1 increases. Since positive charges repel each other, the new equilibrium positions of each of the unit charges, denoted by $q_i^{(t,k)}$, $i \in \{1, 2 \dots, n\}$, will shift to the right as both t and k increase. It also follows from Theorem 3.2 that the positions of the positive unit charges q_i and $q_i^{(t,k)}$, $i \in \{1, 2 \dots, n\}$ interlace provided that $0 < t, k \leq 2$. However, if the positive charge at -1 is increased by more than one unit charge ($k > 2$) or the charge at 1 decreased by more than one unit charge ($t > 2$), it is to be expected that the interlacing properties of the unit charges q_i and $q_i^{(t,k)}$, $i \in \{1, \dots, n\}$ will break down. Therefore the restrictions $t, k \leq 2$ in Theorem 3.2 make sense from electrostatic considerations.

4 Stieltjes Interlacing of Zeros

A classical result due to Stieltjes (cf. [43]) concerns interlacing of the zeros of polynomials of non-consecutive degree in a sequence of orthogonal polynomials; a property called Stieltjes interlacing.

Theorem 4.1 (cf. [43, Thm 3.3.3]) *Let $\{p_n\}_{n=0}^{\infty}$ be a sequence of orthogonal polynomials on the interval (a, b) with respect to $w(x) > 0$ and suppose $m < n$. Then, between any two zeros of p_m, there is at least one zero of p_n.*

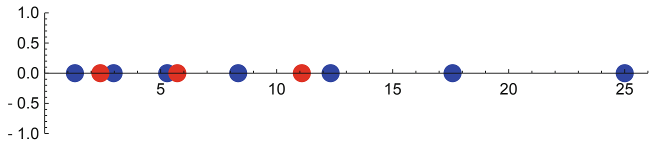


Fig. 3 of $L_7^\alpha(x)$ (blue) and $L_3^\alpha(x)$ (red) for $\alpha = 3.4$

Proof Suppose that $x_{m,k}$ and $x_{m,k+1}$ are two consecutive zeros of $p_m(x)$ and that there is no zero of $p_n(x)$ in $(x_{m,k}, x_{m,k+1})$. Consider

$$g(x) = \frac{p_m(x)}{(x - x_{m,k})(x - x_{m,k+1})}.$$

Then $g(x)p_m(x) \geq 0$ for $x \notin (x_{m,k}, x_{m,k+1})$.

If $\{x_{n,i}\}_{i=1}^n$ are the zeros of $p_n(x)$, Gauss quadrature gives

$$\int_a^b g(x)p_m(x)w(x)dx = \sum_{i=1}^n \lambda_{n,i}\, g(x_{n,i})p_m(x_{n,i}).$$

Since there are no zeros of $p_n(x)$ in $(x_{m,k}, x_{m,k+1})$ we conclude that $g(x_{n,i})p_m(x_{n,i}) \geq 0$ for all $i = 1, 2, \ldots, n$. Further we have $\lambda_{n,i} > 0$ for all $i = 1, 2, \ldots, n$ which implies that the sum on the right-hand side cannot vanish. However, the integral on the left-hand side is zero by orthogonality and we have a contradiction. □

Stieltjes interlacing of the zeros of two Laguerre polynomials is illustrated in Fig. 3.

5 Bounds for the Zeros of Orthogonal Polynomials

Classical methods to obtain bounds for zeros of orthogonal polynomials include the use of monotonicity properties (cf. [5]), known properties of zeros of polynomials in the Laguerre–Polya class (cf. [4, 15, 28, 29, 35]), Sturmian methods for differential equation (cf. [13, 14, 43]), Obrechkov's theorem (cf. [1]) as well as chain sequences and the Wall-Wetzel theorem (cf. [25, 46]).

5.1 Bounds for Zeros from Stieltjes Interlacing

Mixed recurrence relations used to prove Stieltjes interlacing of the zeros of two polynomials from different sequences provide a set of points that can be applied as inner bounds for the extreme zeros of polynomials.

Theorem 5.1 (cf. [10]) *Let $\{p_n\}_{n=0}^{\infty}$ be a sequence of polynomials orthogonal on the interval (c,d). Fix $k,n \in \mathbb{N}$ with $k < n-1$ and suppose deg(g_{n-k-1})$= n-k-1$ with*

$$f(x)g_{n-k-1}(x) = G_k(x)p_{n-1}(x) + H(x)p_n(x) \tag{5.1}$$

where $f(x) \neq 0$ for $x \in (c,d)$ and deg$(G_k) = k$. Then the $n-1$ real, simple zeros of $G_k g_{n-k-1}$ interlace with the zeros of p_n if g_{n-k-1} and p_n are co-prime.

Proof Let $w_1 < \cdots < w_n$ denote the zeros of p_n. Since p_{n-1} and p_n are always co-prime while p_n and g_{n-k-1} are co-prime by assumption, it follows from (5.1) that $G_k(w_j) \neq 0$ for every j. From (5.1), provided $p_n(x) \neq 0$, we have

$$\frac{f(x)g_{n-k-1}(x)}{p_n(x)} = H(x) + \frac{G_k(x)p_{n-1}(x)}{p_n(x)}.$$

The decomposition into partial fractions (cf. [43, Thm 3.3.5]

$$\frac{p_{n-1}(x)}{p_n(x)} = \sum_{j=1}^{n} \frac{A_j}{x-w_j},$$

where $A_j > 0$ for every $j \in \{1,\ldots,n\}$, implies that we can write

$$\frac{f(x)g_{n-k-1}(x)}{p_n(x)} = H(x) + \sum_{j=1}^{n} \frac{G_k(x)A_j}{x-w_j}, \quad x \neq w_j.$$

Suppose that G_k does not change sign in an interval (w_j, w_{j+1}) where $j \in \{1,2,\ldots,n-1\}$. Since $A_j > 0$ and the polynomial H is bounded on I_j while the right hand side takes arbitrarily large positive and negative values on (w_j, w_{j+1}), it follows that g_{n-k-1} must have an odd number of zeros in every interval in which G_k does not change sign. Since G_k is of degree k, there are at least $n-k-1$ intervals (w_j, w_{j+1}), $j \in \{1,\ldots,n-1\}$ in which G_k does not change sign and so each of these intervals must contain exactly one of the $n-k-1$ real, simple zeros of g_{n-k}. We deduce that the k zeros of G_k are real and simple and, together with the $n-k-1$ zeros of g_{n-k-1}, interlace with the n zeros of p_n. □

Corollary 5.2 (cf. [10]) *Suppose (5.1) holds for $k,n \in \mathbb{N}$ fixed and $k < n-1$. The largest (smallest) zero of G_k is a strict lower (upper) bound for the largest (smallest) zero of p_n.*

Example For Laguerre polynomials one can show that, when $n > 1$, $n \in \mathbb{N}$, (5.1) holds for $k = 1$ and $t \in (0,1)$ with

$$g_{n-2} = L_{n-2}^{\alpha+t}$$
$$G_1(x) = x - (\alpha - 1 + (2-t)n).$$

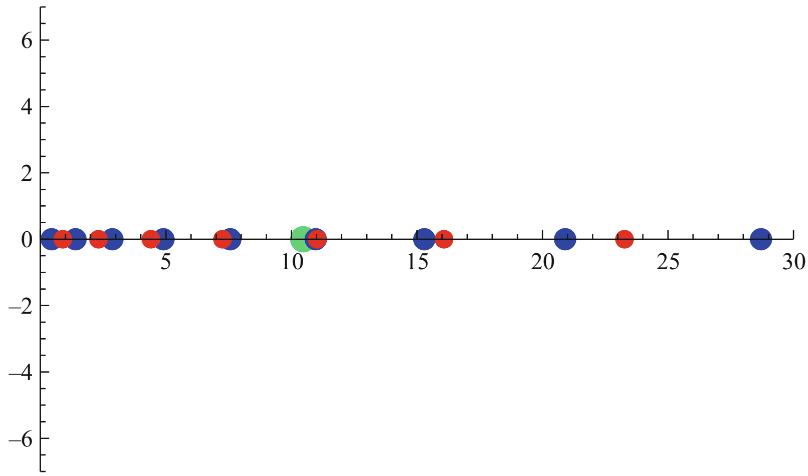


Fig. 4 Zeros of $L_n^{\alpha}(x)$ (blue), $L_{n-2}^{\alpha+t}$ (red) and $\alpha - 1 + (2-t)n$ (green) for $n = 9$, $\alpha = 1.34$ and $t = 0.9$

It follows that for all $\alpha > -1$, $n \in \mathbb{N}$, $t \in (0, 1)$, $\alpha - 1 + (2-t)n$ is an inner bound for the extreme zeros of $L_n^{\alpha}(x)$ as illustrated in Fig. 4 for $t = 0.9$.

Example For Jacobi polynomials $P_n^{(\alpha,\beta)}$, $\alpha, \beta > -1$, it was proved in [12, Thm 2.1(i)(c)] that, when $n > 1$, $n \in \mathbb{N}$, (5.1) holds for $k = 1$ with

$$g_{n-2} = P_{n-2}^{(\alpha+4,\beta)}$$

$$G_1(x) = x - \frac{2(n-1)(n+\alpha+\beta+2) + (\alpha+3)(\beta-\alpha)}{2(n-1)(n+\alpha+\beta+2) + (\alpha+3)(\alpha+\beta+2)}.$$

It follows that for all $\alpha, \beta > -1$, $n \in \mathbb{N}$,

$$\begin{aligned} w_n &> 1 - \frac{2(\alpha+1)(\alpha+3)}{2(n-1)(n+\alpha+\beta+2) + (\alpha+3)(\alpha+\beta+2)} \\ &= 1 - O(\frac{1}{n^2}). \end{aligned}$$

This bound is sharper than the lower bound for largest zero [43, (6.2.11)]

$$1 - \frac{2(\alpha+1)}{2n+\alpha+\beta} = 1 - O(\frac{1}{n}).$$

6 Distance Between the Consecutive Zeros

Sturm's convexity theorem (cf. [42]) states that for

$$y''(t) + F(t)y(t) = 0,$$

a second-order differential equation in normal form, where F is continuous in (a, b) and $y(t)$ is a nontrivial solution in (a, b), the distance between the zeros of $y(t)$ is decreasing (zeros are concave) if $F(t)$ is strictly increasing in (a, b), and the distance between the zeros of $y(t)$ is increasing (zeros are convex) if $F(t)$ is strictly decreasing in (a, b).

Problem The second-order differential equations for orthogonal polynomials and many special functions are not in normal form!

Solutions were found by Szegő [43, Thm 6.3.3] for convexity of the zeros

$$\theta_0,\ \theta_1,\ \theta_2,\ \dots,\theta_{[n/2]+1}$$

of Jacobi polynomials $P_n^{(\alpha,\,\alpha)}(\cos\theta)$ when $-1/2 < \alpha = \beta < 1/2$ and by Deano et al. (cf. [3]) for the convexity of the transformed zeros for hypergeometric functions when using the Liouville transformation.

Consider the more general differential equation

$$x'' + g(t)x' + f(t)x = 0$$

which can be changed to normal form

$$y'' + F(t)y = 0, \tag{6.1}$$

with the transformation

$$y = x \exp\left(\frac{1}{2}\int_0^t g(s)ds\right)$$

where $F(t) = f(t) - \dfrac{1}{4}g^2(t) - \dfrac{1}{2}g'(t)$.

The huge advantage of this transformation is that the zeros of x and y are the same! This transformation was used by Sturm (cf. [42]) for Bessel functions and Hille (cf. [19]) for Hermite polynomials.

Laguerre polynomials satisfy the differential equation

$$tx'' + (\alpha + 1 - t)x' + nx = 0$$

which can be transformed to normal form (6.1) where

$$F(t) = \frac{-t^2 + 2\alpha t + 2t + 4nt - \alpha^2 + 1}{4t^2}$$

and changes monotonicity at

$$t_0 := \frac{\alpha^2 - 1}{\alpha + 2n + 1}.$$

Theorem 6.1 (cf. [27]) *The zeros of* $L_n^\alpha(t)$ *on* $(0, \infty)$ *are*

(a) *all convex if* $n > 0$ *and* $-1 < \alpha \leq 3$*;*
(b) *all convex if* $\alpha > 3$ *and* $0 < n < \frac{\alpha+1}{\alpha-3}$*;*
(c) *concave for* $t < t_0$ *and convex for* $t > t_0$ *when* $\alpha > 3$ *and* $n > \frac{\alpha+1}{\alpha-3}$.

Moreover, for the distance between consecutive zeros we have

$$\Delta x_k > \frac{\pi\sqrt{2}}{\sqrt{2\alpha n + \alpha + 2n^2 + 2n + 1}}, \qquad k = 1, \ldots, n-1,$$

and also if $x_k > t_0$ *then*

$$\frac{\pi}{\sqrt{F(x_k)}} < \Delta x_k < \frac{\pi}{\sqrt{F(x_{k+1})}}, \qquad k = 1, \ldots, n-2.$$

The convexity of the zeros of Laguerre polynomials for the case in Theorem 6.1(a) is illustrated in Fig. 5.

Question Is it possible to find α and n values so that the first several zeros of the Laguerre polynomial are concave?

By Theorem 6.1(c), the answer to this question requires, at the very least, that the smallest zero, x_1, satisfies $x_1 < t_0 = \dfrac{\alpha^2 - 1}{\alpha + 2n + 1}$ and this would be so if one could find an upper bound for x_1 which lies between x_1 and t_0. The known upper bound for x_1, namely $\frac{(\alpha+1)(\alpha+2)}{\alpha+n+1}$, due to Hahn [17], satisfies $t_0 < \dfrac{(\alpha+1)(\alpha+2)}{\alpha+n+1}$ while the bound $\frac{(\alpha+1)(\alpha+3)}{\alpha+2n+1}$, due to Szegő [43], satisfies $t_0 < \dfrac{(\alpha+1)(\alpha+3)}{\alpha+2n+1}$.

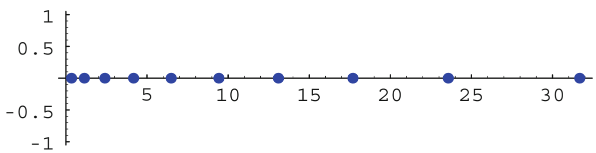


Fig. 5 Zeros of the Laguerre polynomial for $\alpha = 0.98887$ and $n = 10$

Whether a more recent and better upper bound for x_1, for example,

$$\frac{A_n - \sqrt{(\alpha+2)(\alpha+1)B_n}}{2(\alpha+n+1)(\alpha+n)}$$

where

$$A_n = (\alpha+2)(\alpha+1)(2\alpha+3n+2) \text{ and}$$
$$B_n = -4(\alpha+1)(\alpha+2) + \left(5\alpha^2+25\alpha+38\right)n^2 + 4(\alpha+1)\left(\alpha^2+4\alpha+6\right)n,$$

due to Driver and Jordaan (cf. [10]), answers this question, remains an open problem.

References

1. I. Area, D.K. Dimitrov, E. Godoy, F.R. Rafaeli, Inequalities for zeros of Jacobi polynomials via Obrechkoff's theorem. Math. Comput. **81**, 991–1912 (2012)
2. R. Askey, Graphs as an aid to understanding special functions. Asymptotic Comput. Anal. Lect. Notes Pure Appl. **124**, 3–33 (1990)
3. A. Deaño, A. Gil, J. Segura, New inequalities from classical Sturm theorems. J. Approx. Theory **131**, 208–243 (2004)
4. D.K. Dimitrov, G.P. Nikolov, Sharp bounds for the extreme zeros of classical orthogonal polynomials. J. Approx. Theory **162**, 1793–1804 (2010)
5. D.K. Dimitrov, F.R. Rafaeli, Monotonicity of zeros of Laguerre polynomials. J. Comput. Appl. Math., **223**, 699–702 (2009)
6. D.K. Dimitrov, R.O. Rodrigues, On the behaviour of zeros of Jacobi and Gegenbauer polynomials. J. Approx. Theory **116**, 224–239 (2002)
7. D.K. Dimitrov, A. Sri Ranga. Monotonicity of the zeros of orthogonal Laurent polynomials. Methods Appl. Anal. **9**, 9–12 (2002)
8. D.K. Dimitrov, M.V. Mello, F.R. Rafaeli, Monotonicity of zeros of Jacobi-Sobolev type orthogonal polynomials. Appl. Numer. Math. **60**, 263–276 (2010)
9. D.K. Dimitrov, M.E.H. Ismail, F.R. Rafaeli, Interlacing of zeros of orthogonal polynomials under modification of the measure. J. Approx. Theory **175**, 64–76 (2013)
10. K. Driver, K. Jordaan, Bounds for extreme zeros of some classical orthogonal polynomials. J. Approx. Theory **164**, 1200–1204 (2012)
11. K. Driver, K. Jordaan, N. Mbuyi, Interlacing of the zeros of Jacobi polynomials with different parameters. Numer. Algorithms **49**, 143–152 (2008)
12. K. Driver, A. Jooste, K. Jordaan, Stieltjes interlacing of zeros of Jacobi polynomials from different sequences. Electron. Trans. Numer. Anal. **38**, 317–326 (2011)
13. Á. Elbert, A. Laforgia, Upper bounds for the zeros of ultraspherical polynomials. J. Approx. Theory. **61**, 88–97 (1990)
14. Á. Elbert, A. Laforgia, L.G. Rodonó, On the zeros of Jacobi polynomials. Acta Math. Hungar. **64**(4), 351–359 (1994)
15. W.H. Foster, I. Krasikov, Inequalities for real-root polynomials and entire functions. Adv. Appl. Math. **29**, 102–114 (2002)
16. G. Freud, *Orthogonal Polynomials* (Pergamon, Oxford, 1971)

17. W. Hahn, Bericht über die Nullstellen der Laguerrschen und der Hermiteschen Polynome. Jahresber. Deutsch. Math.-Verein. **44**, 215–236 (1933)
18. D. Hilbert, Über die Diskriminante der im Endlichen abbrechenden hypergeometrischen Reihe. J. Reine. Angew. Math. **103**, 337–345 (1888)
19. E. Hille, Über die Nulstellen der Hermiteschen Polynome. Jahresber. Deutsch. Math.-Verein. **44**, 162–165 (1933)
20. M.E.H. Ismail, The variation of zeros of certain orthogonal polynomials. Adv. Appl. Math. **8**, 111–118 (1987)
21. M.E.H. Ismail, An electrostatic model for zeros of general orthogonal polynomials. Pac. J. Math. **193**, 355–369 (2000)
22. M.E.H. Ismail, More on electrostatic models for zeros of orthogonal polynomials. J. Nonlinear Funct. Anal. Optim. **21**, 43–55 (2000)
23. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
24. M.E.H. Ismail, M.E. Muldoon, A discrete approach to monotonicity of zeros of orthogonal polynomials. Trans. Am. Math. Soc. **323**, 65–78 (1991)
25. M.E.H. Ismail, X. Li, Bounds on the extreme zeros of orthogonal polynomials. Proc. Am. Math. Soc. **115**, 131–140 (1992)
26. M.E.H. Ismail, R Zhang, On the Hellmann-Feynman theorem and the variation of zeros of certain special functions. Adv. Appl. Math. **9**, 439–446 (1988)
27. K. Jordaan, F. Tookós, Convexity of the zeros of some orthogonal polynomials and related functions. J. Comp. Anal. Appl. **233**, 762–767 (2009)
28. I. Krasikov, Bounds for zeros of the Laguerre polynomials. J. Approx. Theory **121**, 287–291 (2003)
29. I. Krasikov, On zeros of polynomials and allied functions satisfying second order differential equations. East J. Approx. **9**, 41–65 (2003)
30. R.J. Levit, The zeros of the Hahn polynomials. SIAM Rev. **9**(2), 191–203 (1967)
31. D.S. Lubinsky, Quadrature identities for interlacing and orthogonal polynomials. Proc. Am. Math. Soc. **144**, 4819–4829 (2016)
32. F. Marcellán, F.R. Rafaeli, Monotonicity and Asymptotics of zeros of Laguerre-Sobolev-type orthogonal polynomials of higher order derivatives. Proc. Am. Math. Soc. **139(11)**, 3929–3936 (2011)
33. A. Markov, Sur les racines de certaines equations (Second note). Math. Ann. **27**, 177–182 (1886)
34. M.E. Muldoon, Properties of zeros of orthogonal polynomials and related functions. J. Comput. Appl. Math. **48**, 167–186 (1993)
35. G.P. Nikolov, R. Uluchev, in *Inequalities for Real-Root Polynomials. Proof of a Conjecture of Foster and Krasikov*, in ed. by D.K. Dimitrov, G.P. Nikolov, R. Uluchev. Approximation Theory: A volume dedicated to B. Bojanov (Marin Drinov Academic Publishing House, Sofia, 2004), pp. 201–216
36. P. Paule, Contiguous relations and creative telescoping, Technical report, RISC, Austria, 2001
37. R. Vidũnas, Contiguous relations of hypergeometric series. J. Comput. Appl. Math. **153**(1–2), 507–519 (2003)
38. J. Segura, Interlacing of the zeros of contiguous hypergeometric functions. Numer. Algorithms **49**, 387–407 (2008)
39. B. Simon, in *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*. American Mathematical Society Colloquium Publications, vol. 54 (American Mathematical Society, Providence, 2005)
40. T.J. Stieltjes, Sur quelques théorèmes d'algèbre. C. R. Acad. Sci. **100**, 439–440 (1885). Ouvres Complètes **1**, 440–441
41. T.J. Stieltjes, Sur les polynômes de Jacobi. C. R. Acad. Sci. **100**, 620–622 (1885). Ouvres Complètes **1**, 442–444

42. C. Sturm, Memoire sur les équations différentielles du second ordre. J. Math. Pures Appl. **1**, 106–186 (1836)
43. G. Szegő, in *Orthogonal Polynomials*. AMS Colloquium Publications, vol. 23 (American Mathematical Society, Providence, 1975)
44. N. Takayame, Gröbner basis and the problem of contiguous relations. Jpn J. Appl. Math. **6**, 147–160 (1989)
45. R. Vidũnas, T. Koornwinder, Webpage of the NWO project. Algorithmic methods for special functions by computer algebra (2000). http://www.science.uva.nl/~thk/specfun/compalg.html
46. H.S. Wall, M. Wetzel, Quadratic forms and convergence regions for continued fractions. Duke Math. J. **11**, 983–1000 (1944)
47. B. Wendroff, On orthogonal polynomials. Proc. Am. Math. Soc. **12**, 554–555 (1961)

Properties of Certain Classes of Semiclassical Orthogonal Polynomials

Kerstin Jordaan

Abstract In this lecture we discuss properties of orthogonal polynomials for weights which are semiclassical perturbations of classical orthogonality weights. We use the moments, together with the connection between orthogonal polynomials and Painlevé equations to obtain explicit expressions for the recurrence coefficients of polynomials associated with a semiclassical Laguerre and a generalized Freud weight. We analyze the asymptotic behavior of generalized Freud polynomials in two different contexts. We show that unique, positive solutions of the nonlinear difference equation satisfied by the recurrence coefficients exist for all real values of the parameter involved in the semiclassical perturbation but that these solutions are very sensitive to the initial conditions. We prove properties of the zeros of semiclassical Laguerre and generalized Freud polynomials and determine the coefficients $a_{n,n+j}$ in the differential-difference equation

$$x\frac{d}{dx}P_n(x) = \sum_{k=-1}^{0} a_{n,n+k}P_{n+k}(x),$$

where $P_n(x)$ are the generalized Freud polynomials. Finally, we show that the only monic orthogonal polynomials $\{P_n\}_{n=0}^{\infty}$ that satisfy

$$\pi(x)\mathcal{D}_q^2 P_n(x) = \sum_{j=-2}^{2} a_{n,n+j}P_{n+j}(x),\ \ x = \cos\theta,\ \ a_{n,n-2} \neq 0,\ \ n = 2, 3, \ldots,$$

where $\pi(x)$ is a polynomial of degree at most 4 and $\mathcal{D}_q$ is the Askey–Wilson operator, are Askey–Wilson polynomials and their special or limiting cases, using this relation to derive bounds for the extreme zeros of Askey–Wilson polynomials.

K. Jordaan (✉)
Department of Decision Sciences, University of South Africa, Pretoria, South Africa
e-mail: jordakh@unisa.ac.za

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_18

Keywords Orthogonal polynomials · Zeros · Laguerre polynomials · Freud polynomials · Askey–Wilson polynomials · Jacobi polynomial

Mathematics Subject Classification (2000) Primary 33C45; Secondary 42C05

1 Introduction

Let $\{p_n(x)\}_{n=0}^{\infty}$, $\deg(p_n) = n$ be the monic orthogonal polynomials with respect to the positive measure μ with support $[a, b]$ where we assume that all the moments,

$$\mu_k = \int_a^b x^k d\mu,$$

exist. By orthogonality, we have the three-term recurrence relation

$$p_{n+1} = (x - \alpha_n)p_n - \beta_n p_{n-1}, \quad n = 0, 1, 2, \ldots$$

with initial conditions

$$p_{-1} \equiv 0, \qquad p_0 \equiv 1$$

and recurrence coefficients

$$\alpha_n \in \mathbb{R},\ n = 0, 1, 2\ldots, \quad \beta_n > 0, \quad n = 1, 2, \ldots.$$

The coefficients in the three-term recurrence relation can also be expressed in terms of determinants whose entries are the moments associated with measure μ.

$$\alpha_n = \frac{\widetilde{\Delta}_{n+1}}{\Delta_{n+1}} - \frac{\widetilde{\Delta}_n}{\Delta_n}, \qquad \beta_n = \frac{\Delta_{n+1}\Delta_{n-1}}{\Delta_n^2},$$

where Δ_n is the Hankel determinant

$$\Delta_n = \det\left[\mu_{j+k}\right]_{j,k=0}^{n-1} = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_1 & \mu_2 & \cdots & \mu_n \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} \end{vmatrix}, \qquad n \geq 1,$$

with $\Delta_0 = 1$, $\Delta_{-1} = 0$, and $\widetilde{\Delta}_n$ is the determinant

$$\widetilde{\Delta}_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-2} & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n-1} & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-3} & \mu_{2n-1} \end{vmatrix}, \qquad n \geq 1,$$

with $\widetilde{\Delta}_0 = 0$ and μ_k is the kth moment.

The converse statement, known as the spectral theorem for orthogonal polynomials, is often attributed to Favard [16] but was probably discovered independently, around 1935, by both Natanson [32] and Shohat [35, 36]. The result can in fact be traced back to earlier work on continued fractions with a rudimentary form given by Stieltjes [38] in 1894, see [9], also [31, 39]. The result also appears in books by Wintner [42] and Stone [37], see [23]. A modern proof can be found in [4].

Theorem 1.1 *Consider a family of monic polynomials* $\{p_n(x)\}_{n=0}^{\infty}$ *that satisfies a three-term recurrence relation*

$$p_{n+1} = (x - \alpha_n)p_n - \beta_n p_{n-1}$$

with initial conditions $p_0 = 1$ *and* $p_{-1} = 0$ *where* $\alpha_{n-1} \in \mathbb{R}$ *and* $\beta_n > 0$, $n \in \mathbb{N}$. *Then there exists a measure* μ *on the real line such that these polynomials are monic orthogonal polynomials satisfying*

$$\int_{\mathbb{R}} p_n(x)p_m(x)\, d\mu(x) = \begin{cases} 0, & n \neq m \\ h_n \neq 0, & n = m, \end{cases} \qquad m, n = 0, 1, 2, \ldots$$

Note that the proof of this result does not give explicit information about the measure or support of the measure. Indeed, the measure need not be unique since this depends on the solution of the Hamburger moment problem.

The above results give rise to two interesting problems in the theory of orthogonal polynomials.

1. Given an orthogonality measure $\mu(x)$, what can be deduced about characterising properties of the orthogonal polynomials, such as the recurrence coefficients $\{\alpha_n, \beta_n\}$, $n \in \mathbb{N}$ or the differential equation and the differential-difference equation satisfied by the polynomials? This is known as the direct problem of orthogonal polynomials.
2. Given recurrence coefficients $\{\alpha_n, \beta_n\}$, $n \in \mathbb{N}$, what can be deduced about the uniqueness, nature and support of the orthogonality measure? This is often referred to as the inverse problem of orthogonal polynomials.

For most classical orthogonality measures, the properties satisfied by the orthogonal polynomials are known. In the case of semiclassical measures there are some interesting recent results solving the direct problem and some of these are discussed in Sect. 2 in the context of semiclassical Laguerre and generalized Freud polynomials. In Sect. 3 we consider a new characterisation of Askey–Wilson polynomials.

2 Semiclassical Orthogonal Polynomials

Semiclassical orthogonal polynomials are defined as orthogonal polynomials for which the weight function satisfies a Pearson equation

$$\frac{d}{dx}[\sigma(x)w(x)] = \tau(x)w(x),$$

with $\deg(\sigma) \geq 2$ or $\deg(\tau) \neq 1$; see Hendriksen and van Rossum [21].

2.1 Semiclassical Laguerre Polynomials

Consider monic orthogonal polynomials with respect to the semiclassical Laguerre weight

$$w(x;t) = x^{\lambda} \exp(-x^2 + tx), \quad x \in \mathbb{R}^+,\ \lambda > -1, t \in \mathbb{R} \tag{2.1}$$

which satisfy the three-term recurrence relation

$$xL_n(x;t) = L_{n+1}(x;t) + \widetilde{\alpha}_n(t)L_n(x;t) + \widetilde{\beta}_n(t)L_{n-1}(x;t), \tag{2.2}$$

Theorem 2.1 ([6]) *The coefficients $\widetilde{\alpha}_n(t)$ and $\widetilde{\beta}_n(t)$ in the recurrence relation (2.2) associated with the semiclassical Laguerre weight (2.1) satisfy the discrete system*

$$(2\widetilde{\alpha}_n - t)(2\widetilde{\alpha}_{n-1} - t) = \frac{(2\widetilde{\beta}_n - n)(2\widetilde{\beta}_n - n - \lambda)}{\widetilde{\beta}_n},$$
$$2\widetilde{\beta}_n + 2\widetilde{\beta}_{n+1} + \widetilde{\alpha}_n(2\widetilde{\alpha}_n - t) = 2n + \lambda + 1.$$

Theorem 2.2 ([17]) *The coefficients $\widetilde{\alpha}_n(t)$ in the recurrence relation (2.2) associated with the semiclassical Laguerre weight (2.1) are given by*

$$\widetilde{\alpha}_n(t) = \tfrac{1}{2}q_n(z) + \tfrac{1}{2}t,$$

with $z = \frac{1}{2}t$ where $q_n(z)$ satisfies

$$\frac{d^2q_n}{dz^2} = \frac{1}{2q_n}\left(\frac{dq_n}{dz}\right)^2 + \tfrac{3}{2}q_n^3 + 4zq_n^2 + 2(z^2 - 2n - \lambda - 1)q_n - \frac{2\lambda^2}{q_n},$$

which is $\mathrm{P_{IV}}$, *with parameters* $(A, B) = (2n + \lambda + 1, -2\lambda^2)$.

Remark 2.3 The parameters (A, B) satisfy the condition for $\mathrm{P_{IV}}$ to have solutions expressible in terms of parabolic cylinder functions.

The following lemma confirms the existence of the first moment of the semi-classical Laguerre weight in (2.1).

Lemma 2.4 *Let $x \in \mathbb{R}^+$, $\lambda > -1, t \in \mathbb{R}$, then the first moment $\mu_0(t; \lambda)$ of the semi-classical Laguerre weight $w(x; t) = x^\lambda \exp(-x^2 + tx)$ is finite.*

Proof The first moment $\mu_0(t; \lambda)$ takes the form

$$\mu_0(t; \lambda) = \int_0^\infty x^\lambda \exp(-x^2 + tx)\, dx$$

and, since for $\lambda > 0$ the integrand $w(x; t) = x^\lambda \exp(-x^2 + tx)$ is continuous on $[0, \infty)$, it is integrable on the compact set $[0, K]$ for any $K > 0$ when $\lambda > 0$. Note that $\lim_{x\to\infty} x^2 w(x; t) = 0$, hence, by definition, there exists $N > 0$ such that $x^2 w(x; t) < 1$ whenever $x > N$. Since $\displaystyle\int_N^\infty \frac{dx}{x^2} < \infty$, it follows from the Comparison Theorem that $\displaystyle\int_N^\infty w(x; t)\, dx < \infty$ for $N > 0$ and, in particular, $N = K$. Hence $\mu_0(t; \lambda) < \infty$ for $\lambda > 0$. When $-1 < \lambda < 0$, the improper integral

$$\begin{aligned}
\int_0^1 x^\lambda \exp(-x^2 + tx)\, dx &= \lim_{b\to 0^+} \int_b^1 x^\lambda \exp(-x^2 + tx)\, dx \\
&= \exp\left(\frac{t^2}{4}\right) \lim_{b\to 0^+} \int_b^1 x^\lambda \exp\left(-\left(x - \frac{t}{2}\right)^2\right) dx \\
&\le \exp\left(\frac{t^2}{4}\right) \lim_{b\to 0^+} \int_b^1 x^\lambda\, dx \\
&= \lim_{b\to 0^+} \left[\frac{x^{\lambda+1}}{\lambda + 1}\right]_b^1 < \infty
\end{aligned}$$

by the comparison theorem since $e^{-(x-\frac{t}{2})^2} < 1$ for $x \in \mathbb{R}$. Furthermore, for $-1 < \lambda < 0$, we have that

$$\int_1^\infty x^\lambda \exp(-x^2 + tx)\, dx \le \int_1^\infty \exp(-x^2 + tx) < \infty$$

since $x^\lambda < 1$ for $x > 1$ and $\lambda < 0$. Finally, for $\lambda = 0$, we have that

$$\begin{aligned}
\int_0^\infty x^\lambda \exp(-x^2 + tx)\, dx &= \lim_{b\to 0^+} \int_b^\infty x^\lambda \exp(-x^2 + tx)\, dx \\
&= \lim_{b\to 0^+} \int_b^\infty \exp(-x^2 + tx) dx
\end{aligned}$$

which is finite. □

Next we derive an explicit expression for the moment $\mu_0(t;\lambda)$ associated with the semiclassical Laguerre weight (2.1).

Theorem 2.5 ([11]) *For the weight (2.1), the moment $\mu_0(t;\lambda)$ is given by*

$$\mu_0(t;\lambda)=\begin{cases}\dfrac{\Gamma(\lambda+1)\exp\left(\frac{1}{8}t^2\right)}{2^{(\lambda+1)/2}}\,D_{-\lambda-1}\big(-\tfrac{1}{2}\sqrt{2}\,t\big), & \text{if}\quad \lambda\notin\mathbb{N},\\[2ex] \tfrac{1}{2}\sqrt{\pi}\,\dfrac{d^m}{dt^m}\left\{\exp\left(\tfrac{1}{4}t^2\right)\left[1+\operatorname{erf}(\tfrac{1}{2}t)\right]\right\}, & \text{if}\quad \lambda=m\in\mathbb{N},\end{cases}$$

with $D_\nu(\zeta)$ the parabolic cylinder function and $\operatorname{erf}(z)$ *the error function. Further $\mu_0(t;\lambda)$ satisfies the equation*

$$\frac{d^2\mu_0}{dt^2}-\tfrac{1}{2}t\frac{d\mu_0}{dt}-\tfrac{1}{2}(\lambda+1)\mu_0=0.$$

Note that the semiclassical Laguerre weight has the form

$$w(x;t)=\omega_0(x)\exp(xt),\qquad x\in[a,b],\tag{2.3}$$

where $\omega_0(x)=x^\lambda\ \exp(-x^2)$ and hence

$$\mu_k=\int_a^b x^k\omega_0(x)\exp(xt)\,dx=\int_a^b\frac{\partial^k}{\partial t^k}\left(\omega_0(x)\exp(xt)\right)\,dx.$$

Since the function $\frac{\partial^{k-1}}{\partial t^{k-1}}\left(\omega_0(x)\exp(xt)\right)$, $k>0$, satisfies the necessary consitions for reversing the order of integration and differentiation for functions of two variables (cf. [25, Thm 16.11]), we obtain

$$\int_a^b\frac{\partial^k}{\partial t^k}\left(\omega_0(x)\exp(xt)\right)\,dx=\frac{d}{dt}\int_a^b\frac{\partial^{k-1}}{\partial t^{k-1}}\left(\omega_0(x)\exp(xt)\right)\,dx$$

Iterating this, we can write the n-th moment as a derivative of the first moment as follows

$$\mu_k=\int_a^b x^k\omega_0(x)\exp(xt)\,dx=\frac{d^k}{dt^k}\left(\int_a^b\omega_0(x)\exp(xt)\,dx\right)=\frac{d^k\mu_0}{dt^k}.$$

It follows that, if the weight has the form (2.3), then the Hankel determinant is given by $\Delta_n(t)=\mathcal{W}\left(\mu_0,\dfrac{d\mu_0}{dt},\ldots,\dfrac{d^{n-1}}{dt^{n-1}}\mu_0\right)$, $\Delta_0=1$, $\Delta_{-1}=0$ where

$\mathcal{W}(\varphi_1, \varphi_2, \ldots, \varphi_n)$ is the Wronskian given by

$$\mathcal{W}(\varphi_1, \varphi_2, \ldots, \varphi_n) = \begin{vmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_n \\ \varphi_1^{(1)} & \varphi_2^{(1)} & \cdots & \varphi_n^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)} & \varphi_2^{(n-1)} & \cdots & \varphi_n^{(n-1)} \end{vmatrix}, \qquad \varphi_j^{(k)} = \frac{d^k \varphi_j}{dt^k}$$

and this gives rise to the following result.

Theorem 2.6 ([11]) *The recurrence coefficients $\widetilde{\alpha}_n(t)$ and $\widetilde{\beta}_n(t)$ associated with the weight (2.1) are*

$$\widetilde{\alpha}_n(t) = \tfrac{1}{2} q_n(z) + \tfrac{1}{2} t,$$
$$\widetilde{\beta}_n(t) = -\tfrac{1}{8} \frac{dq_n}{dz} - \tfrac{1}{8} q_n^2(z) - \tfrac{1}{4} z q_n(z) + \tfrac{1}{4}\lambda + \tfrac{1}{2} n,$$

with $z = \frac{1}{2}t$, where

$$q_n(z) = -2z + \frac{d}{dz} \ln \frac{\Psi_{n+1,\lambda}(z)}{\Psi_{n,\lambda}(z)}$$

$$\Psi_{n,\lambda}(z) = \mathcal{W}\left(\psi_\lambda, \frac{d\psi_\lambda}{dz}, \ldots, \frac{d^{n-1}\psi_\lambda}{d^{n-1}z}\right), \ \Psi_{0,\lambda}(z) = 1,$$

and

$$\psi_\lambda(z) = \begin{cases} D_{-\lambda-1}\big(-\sqrt{2}\, z\big) \exp\left(\tfrac{1}{2} z^2\right), & \text{if} \quad \lambda \notin \mathbb{N}, \\ \dfrac{d^m}{dz^m} \left\{\big[1 + \operatorname{erf}(z)\big] \exp(z^2)\right\}, & \text{if} \quad \lambda = m \in \mathbb{N}. \end{cases}$$

The zeros of semiclassical Laguerre polynomials satisfy the following properties.

Theorem 2.7 ([12]) *Let $L_n(x; t)$ denote the monic semiclassical Laguerre polynomials orthogonal with respect to*

$$w(x; t) = x^\lambda \exp(-x^2 + tx), \quad x \in \mathbb{R}^+.$$

Then, for $\lambda > -1$ and $t \in \mathbb{R}$, the zeros $x_{1,n} < x_{2,n} < \cdots < x_{n,n}$ of $L_n(x; t)$

(i) *are real, distinct and interlacing with*

$$0 < x_{1,n} < x_{1,n-1} < x_{2,n} < \cdots < x_{n-1,n} < x_{n-1,n-1} < x_{n,n}; \tag{2.4}$$

(ii) *strictly increase with both t and λ;*

(iii) *satisfy*

$$a_n < x_{1,n} < \widetilde{\alpha}_{n-1} < x_{n,n} < b_n,$$

where

$$a_n = \min_{1\le k\le n-1}\left\{\frac{1}{2}(\widetilde{\alpha}_k+\widetilde{\alpha}_{k-1}) - \frac{1}{2}\sqrt{(\widetilde{\alpha}_k+\widetilde{\alpha}_{k-1})^2+4c_n\widetilde{\beta}_k}\right\},$$
$$b_n = \max_{1\le k\le n-1}\left\{\frac{1}{2}(\widetilde{\alpha}_k+\widetilde{\alpha}_{k-1}) + \frac{1}{2}\sqrt{(\widetilde{\alpha}_k+\widetilde{\alpha}_{k-1})^2+4c_n\widetilde{\beta}_k}\right\},$$

with $c_n = 4\cos^2\left(\frac{\pi}{n+1}\right) + \varepsilon$, $\varepsilon > 0$.

Proof

(i) The proof for classical orthogonal polynomials, where $t = 0$, work without change.

(ii) For the semiclassical Laguerre weight

$$w(x;t) = x^{\lambda}\exp(-x^2+tx), \quad x\in\mathbb{R}^+,$$

we have

$$\frac{\partial}{\partial\lambda}\ln w(x;t) = \ln x,$$

an increasing function of x. It follows from Markov's monotonicity theorem that the zeros of $L_n^{(\lambda)}(x;t)$ increase as λ increases.

Similarly, since

$$\frac{\partial}{\partial t}\ln w(x;t) = x,$$

increases with x, it follows that the zeros of $L_n^{(\lambda)}(x;t)$ increase as t increases.

(iii) The inner bound $\widetilde{\alpha}_{n-1}$ for the extreme zeros follows from [15, Cor. 2.2] together with the three-term recurrence

$$\widetilde{\beta}_{n-1}(t)L_{n-2}^{(\lambda)}(x;t) = [x-\widetilde{\alpha}_{n-1}(t)]L_{n-1}^{(\lambda)}(x;t) - L_n^{(\lambda)}(x;t)$$

and (2.4) since $\widetilde{\beta}_{n-1}(t)$ does not depend on x. The outer bounds a_n and b_n for the extreme zeros $x_{1,n}$ and $x_{n,n}$ respectively, follow from the approach based on the Wall–Wetzel Theorem, introduced by Ismail and Li [24] that uses finite chain sequences, applying their results to the three term recurrence relation.

□

2.2 Generalized Freud Polynomials

Polynomials orthogonal with respect to a symmetric moment functional can be generated via quadratic transformation from the classical orthogonal polynomials. For example, Laguerre polynomials generate a class of generalized Hermite polynomials while Jacobi polynomials give rise to a class of generalized Ultraspherical polynomials. Symmetrizing the semiclassical Laguerre weight (2.1) yields a sequence $\{S_n(x;t)\}_{n=0}^{\infty}$ of polynomials orthogonal with respect to the even weight

$$w(x;t) = |x|^{2\lambda+1}\exp(-x^4+tx^2) \quad x\in\mathbb{R} \text{ for } \lambda>-1 \text{ and } t\in\mathbb{R} \tag{2.5}$$

known as the generalized Freud weight.

Exercise Show that the moments of the generalized Freud weight (2.5) all exist for $x, t \in \mathbb{R}$ and $\lambda > -1$.

Generalized Freud orthogonal polynomials with respect to the generalized Freud weight (2.5) satisfy the three-term recurrence relation

$$S_{n+1}(x;t) = xS_n(x;t) - \beta_n(t;\lambda)S_{n-1}(x;t), \quad S_{-1}=0, \quad S_0=1.$$

Expressions for the recurrence coefficients $\beta_n(t;\lambda)$ in terms of Wronskians of parabolic cylinder functions that appear in the description of special function solutions of $\mathrm{P_{IV}}$ were obtained in [13].

The first moment,

$$\begin{aligned}\mu_0(t;\lambda) &= \int_{-\infty}^{\infty} |x|^{2\lambda+1}\exp\left(-x^4+tx^2\right)\underset{\cdot}{x}\\ &= \frac{\Gamma(\lambda+1)}{2^{(\lambda+1)/2}}\exp\left(\tfrac{1}{8}t^2\right)D_{-\lambda-1}\big(-\tfrac{1}{2}\sqrt{2}\,t\big),\end{aligned}$$

can be obtained using the integral representation of a parabolic cylinder function (cf. [34]). The even moments are

$$\begin{aligned}\mu_{2n}(t;\lambda) &= \int_{-\infty}^{\infty} x^{2n}\,|x|^{2\lambda+1}\exp\left(-x^4+tx^2\right)dx\\ &= \frac{d^n}{dt^n}\mu_0(t,\lambda), \quad n=1,2,\ldots\end{aligned}$$

and, since the integrand is odd, the odd ones are

$$\mu_{2n+1}(t;\lambda) = 0, \quad n=1,2,\ldots$$

The explicit expressions for the recurrence coefficients

$$\beta_n = \frac{\Delta_{n+1}\Delta_{n-1}}{\Delta_n^2}$$

in the three term recurrence relation for generalized Freud polynomials are given by (cf. [13])

$$\beta_{2n} = \frac{d}{dt}\ln\frac{\tau_n(t;\lambda+1)}{\tau_n(t;\lambda)},$$
$$\beta_{2n+1} = \frac{d}{dt}\ln\frac{\tau_{n+1}(t;\lambda)}{\tau_n(t;\lambda+1)},$$

for $n \geq 0$, where $\tau_n(t;\lambda)$ is the Hankel determinant given by

$$\tau_n(t;\lambda) = \det\left[\frac{d^{j+k}}{dt^{j+k}}\mu_0(t;\lambda)\right]_{j,k=0}^{n-1}, \quad \tau_0(t;\lambda) = 1.$$

Theorem 2.8 ([13]) *The recurrence coefficients $\beta_n(t;\lambda)$ satisfy the nonlinear difference equation*

$$\beta_n(\beta_n n + 1 + \beta_n + \beta_n n - 1 - \tfrac{1}{2}t) = \frac{2n + (2\lambda+1)[1-(-1)^n]}{8}, \tag{2.6}$$

which is known as discrete Painlevé I.

An overview of the problem of existence and uniqueness of positive solutions of nonlinear difference equations of type (2.6) is given by Alsulami, Nevai, Szabados and van Assche in [2].

Theorem 2.9 ([12]) *For $t \in \mathbb{R}$ and $\beta_0 = 0$, there exists a unique $\beta_1(t;\lambda) > 0$ such that $\{\beta_n(t;\lambda)\}_{n\in\mathbb{N}}$ defined by the nonlinear difference equation (2.6) is a positive sequence and the solution arises when*

$$\beta_1(t;\lambda) = \tfrac{1}{2}t + \tfrac{1}{2}\sqrt{2}\,\frac{D_{-\lambda}\left(-\frac{1}{2}\sqrt{2}\,t\right)}{D_{-\lambda-1}\left(-\frac{1}{2}\sqrt{2}\,t\right)} = \Phi_\lambda(t). \tag{2.7}$$

The solution of the nonlinear discrete equation (2.6) is highly sensitive to the initial conditions $\beta_0(t;\lambda) = 0$ and $\beta_1(t;\lambda) = \Phi_\lambda(t)$ in (2.7) as illustrated in Fig. 1.

The asymptotic expansion of $\beta_n(t;\lambda)$ satisfying the nonlinear discrete equation (2.6) when

- $t = 0$ and $\lambda = -\frac{1}{2}$ was studied by Lew and Quarles [30];
- $t \in \mathbb{R}$ and $\lambda = -\frac{1}{2}$ was given by Clarke and Shizgal [10].

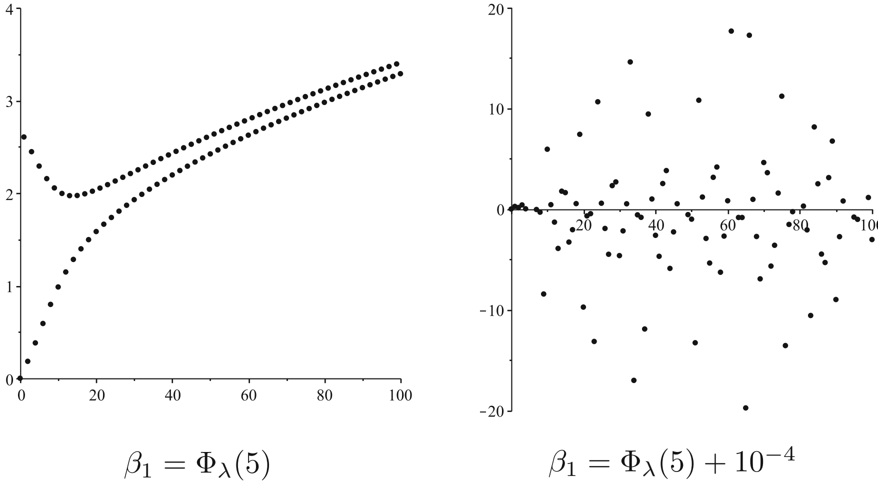


Fig. 1 Plots of the points (n, β_n) where β_n satisfies (2.6) with initial conditions $\beta_0 = 0$ and $\beta_1 = \Phi_\lambda(t) + \epsilon$, with $\epsilon = 0, 10^{-4}$, for $t = 5$ and $\lambda = \frac{1}{2}$

Theorem 2.10 ([12]) *Let $t, \lambda \in \mathbb{R}$, then as $n \to \infty$, the recurrence coefficient β_n associated with monic generalized Freud polynomials*

$$\beta_{2n}(t;\lambda) = \frac{\sqrt{6}\,n^{1/2}}{6}\left\{1 + \frac{\sqrt{6}\,t}{12n^{1/2}} + \frac{t^2}{48n} - \frac{t^4-48}{4608n^2} + \mathcal{O}(n^{-5/2})\right\},$$

$$\beta_{2n+1}(t;\lambda) = \frac{\sqrt{3}\,(2n+1)^{1/2}}{6}\left\{1 + \frac{\sqrt{3}\,t}{6(2n+1)^{1/2}} + \frac{t^2+12(2\lambda+1)}{24(2n+1)} - \frac{t^4+24(2\lambda+1)t^2+96(6\lambda^2+6\lambda+1)}{1152(2n+1)^2} + \mathcal{O}(n^{-5/2})\right\}. \qquad (2.8)$$

The first few terms of the asymptotic expansions of $\beta_{2n}(t;\lambda)$ and $\beta_{2n+1}(t;\lambda)$, considered as functions of a continuous variable n, are compared to the actual values of the points (n, b_n) satisfying (2.8) in Fig. 2.

As $t \to \infty$, the recurrence coefficient $\beta_n(t;\lambda)$ has the asymptotic expansion

$$\beta_{2n}(t;\lambda) = \frac{n}{t} - \frac{2n(2\lambda-n+1)}{t^3} + \mathcal{O}\big(t^{-5}\big),$$

$$\beta_{2n+1}(t;\lambda) = \frac{t}{2} + \frac{\lambda-n}{t} - \frac{2(\lambda^2-4\lambda n+n^2-\lambda-n)}{t^3} + \mathcal{O}\big(t^{-5}\big).$$

The plots of $\beta_n(t;\lambda)$, for $n = 1, 2, ..., 10$, with $\lambda = \frac{1}{2}$ given in Fig. 3 clearly illustrate the completely different behaviour for $\beta_n(t;\lambda)$ as $t \to \infty$ when n is even and n is odd.

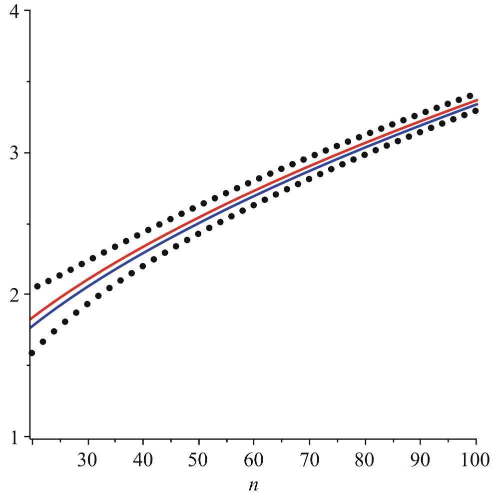


Fig. 2 Plots of points (n, β_n) satisfying (2.6) with initial conditions $\beta_0 = 0$, $\beta_1 = \Phi_\lambda(t)$ and the first few terms of the asymptotic expansions (2.8) of $\beta_{2n}(t; \lambda)$ (blue) and $\beta_{2n+1}(t; \lambda)$ (red) with $\lambda = \frac{1}{2}$ and $t = 5$

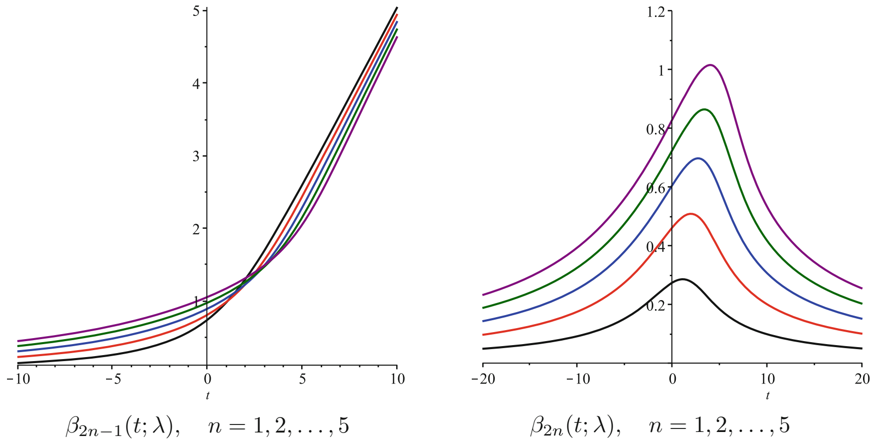


Fig. 3 Plots of the recurrence coefficients $\beta_{2n-1}(t; \frac{1}{2})$ and $\beta_{2n}(t; \frac{1}{2})$, for $n = 1$ (black), $n = 2$ (red), $n = 3$ (blue), $n = 4$ (green) and $n = 5$ (purple)

Further, as $t \to -\infty$

$$\beta_{2n}(t;\lambda) = -\frac{n}{t} + \frac{2n(2\lambda+3n+1)}{t^3} + \mathcal{O}(t^{-5}),$$

$$\beta_{2n+1}(t;\lambda) = -\frac{\lambda+n+1}{t} + \frac{2(\lambda+n+1)(\lambda+3n+2)}{t^3} + \mathcal{O}(t^{-5}).$$

$$\lim_{t\to\infty} \beta_n(t;\lambda) = \tfrac{1}{4}[1-(-1)^n]t.$$

In [41], the asymptotics of $S_n(x;t)$ as $n \to \infty$ are discussed. Asymptotic properties of the extreme zeros of generalized Freud polynomials were studied by Freud [18] and Nevai [33]. Kasuga and Sakai [26] extended and generalized these results. Arceo et al. [3] generalized the electrostatic interpretation of the zero distribution and provided an equation of motion for the distribution of the zeros of a polynomial associated with an Uvarov modification of a quartic Freud type weight ($\lambda = -\frac{1}{2}$).

The generalized Freud weight $w(x;t) = |x|^{2\lambda+1}\exp(-x^4+tx^2)$, is even and the zeros of the corresponding orthogonal polynomials are symmetric about the origin. This implies that the positive and the negative zeros have opposing monotonicity and, as a result of this symmetry, it suffices to study the monotonicity and bounds of the positive zeros.

Theorem 2.11 ([12]) *Let $S_n(x;t)$ be the monic generalized Freud polynomials orthogonal with respect to the weight $w(x;t) = |x|^{2\lambda+1}\exp(-x^4+tx^2)$, and let $x_{n,1}(\lambda,t) < x_{n,2}(\lambda,t) < \cdots < x_{n,[n/2]}(\lambda,t)$ denote the positive zeros of $S_n(x;t)$ where $[m]$ is the largest integer smaller than m. Then, for $\lambda > -1$ and $t \in \mathbb{R}$*

(i) *the zeros of $S_n(x;t)$ are real and distinct and*

$$x_{n,1}(\lambda,t) < x_{n-1,1}(\lambda,t) < x_{n,2}(\lambda,t) < \cdots < x_{n,[n/2]}(\lambda,t);$$

(ii) *the νth zero $x_{n,\nu}(\lambda,t)$, ν fixed, is an increasing function of both λ and t;*

(iii) *the largest zero satisfies the inequality*

$$x_{n,[n/2]}(\lambda,t) < \max_{1\le k\le n-1} \sqrt{c_n\beta_k(t;\lambda)},$$

where $c_n = 4\cos^2\left(\frac{\pi}{n+1}\right) + \varepsilon$, $\varepsilon > 0$.

The interlacing of the zeros of consecutive terms in the sequence $\{S_n(x;t)\}_{n=0}^{\infty}$ of generalized Freud polynomials, described in Theorem 2.11(i), is clearly illustrated in Fig. 4.

Figure 5 shows the monotocity of the zeros of $S_n(x;t)$ for n fixed with increasing t.

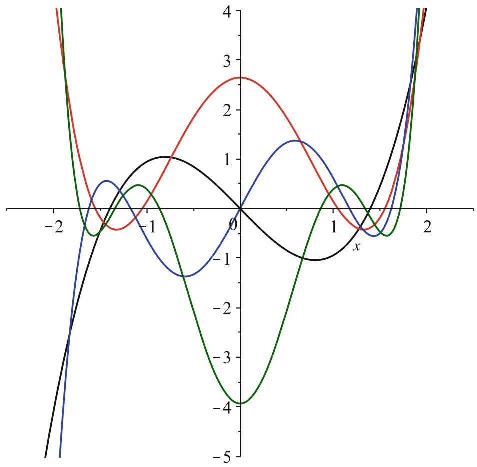


Fig. 4 Plots of the polynomials $S_3(x;t)$ (black), $S_4(x;t)$ (red), $S_5(x;t)$ (blue), $S_6(x;t)$ (green) for $t=3$, with $\lambda=\frac{1}{2}$

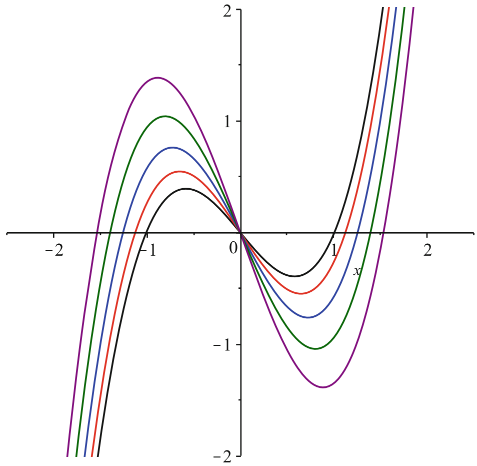


Fig. 5 Plots of the polynomial $S_n(x;t)$, $n=3$ for $t=0$ (black), $t=1$ (red), $t=2$ (blue), $t=3$ (green) and $t=4$ (purple), with $\lambda=\frac{1}{2}$

Consider the differential-difference equation satisfied by monic orthogonal polynomials $S_n(x;t)$ with respect to the generalized Freud weight

$$\pi(x)\frac{d}{dx}S_n(x;t)=\sum_{j=-t}^{s}a_{n,n+j}S_{n+j}(x;t),\quad n=1,2,\ldots \tag{2.9}$$

Shohat [35] gave a procedure using quasi-orthogonality to derive (2.9) for weights $w(x;t)$ such that $w'(x;t)/w(x;t)$ is a rational function. The method of ladder operators was introduced by Chen and Ismail [8] and adapted by Chen and Feigin [7] to the situation where the weight function vanishes at one point. Clarkson, Jordaan and Kelil [13] generalized the work by Chen and Feigin, giving a more explicit expression for the coefficients in (2.9) when the weight function is positive on the real line except for one point.

Theorem 2.12 ([13]) *For the generalized Freud weight*

$$w(x;t) = |x|^{2\lambda+1} \exp\left(-x^4 + tx^2\right), \qquad x \in \mathbb{R},\ \lambda > 0$$

the monic orthogonal polynomials $S_n(x;t)$ *satisfy*

$$x\frac{d}{dx}S_n(x;t) = \sum_{j=-1}^{0} a_{n,n+j} S_{n+j}(x;t)$$

with

$$\begin{aligned} a_{n,n-1} &= 4\beta_n x(x^2 - \tfrac{1}{2}t + \beta_n + \beta_{n+1}), \\ a_{n,n} &= -4x^2\beta_n - \frac{(2\lambda+1)[1-(-1)^n]}{2}. \end{aligned}$$

3 A Characterisation of Askey–Wilson Polynomials

A sequence of orthogonal polynomials is classical if the sequence $\{p_n(x)\}_{n=0}^{\infty}$ as well as $D^m p_{n+m}$, $m \in \mathbb{N}$, where D is the usual derivative $\frac{d}{dx}$ or one of its extensions, including the difference operator, q-difference operator and divided-difference operator, satisfies a three term recurrence of the form ensuring orthogonality by the spectral theorem.

Consider a structural relation of type

$$\pi(x)p_n(x) = \sum_{j=-t}^{s} a_{n,n+j}\, p_{n+j}(x), \quad n = 1, 2, \ldots \tag{3.1}$$

where $\pi(x)$ is a polynomial and S is a linear operator that maps a polynomial of precise degree n to a polynomial of degree $n-1$.

Askey raised the problem of characterizing the orthogonal polynomials satisfying a structure relation of the form (3.1) when $S = \dfrac{d}{dx}$ (cf. [1, p. 69]).

Al-Salam and Chihara [1] characterized Jacobi, Laguerre and Hermite as the only orthogonal polynomials with a structure relation of form (3.1) with $t = s = 1$ where $\pi(x)$ is a polynomial of degree at most two. Replacing S in (3.1) by the difference operator $\Delta f(s) = f(s+1) - f(s)$, García et al. [19] proved that Hahn, Krawtchouk, Meixner and Charlier polynomials are the only orthogonal polynomial sequences satisfying

$$\pi(x)\Delta p_n(x) = \sum_{j=-1}^{1} a_{n,n+j}\, p_{n+j}(x), \quad n = 1, 2, \ldots,$$

with $\pi(x)$ a polynomial of degree at most two. More recently, replacing the operator S in (3.1) by the Hahn operator, also known as the q-difference operator or Jackson derivative,

$$(D_q f)(x) = \frac{f(x) - f(qx)}{(1-q)x},$$

Datta and Griffin [14] characterized the big q-Jacobi polynomial or one of its special or limiting cases (Al-Salam-Carlitz 1, little and big q-Laguerre, little q-Jacobi, and q-Bessel polynomials) as the only orthogonal polynomials that satisfy

$$\pi(x) D_q p_n(x) = \sum_{j=-1}^{1} a_{n,n+j}\, p_{n+j}, \quad n = 1, 2, \ldots$$

where $\pi(x)$ is a polynomial of degree at most two.

Although the polynomials mentioned above are all special or limiting cases of the Askey–Wilson polynomials

$$\frac{a^n P_n(x; a, b, c, d|q)}{(ab, ac, ad; q)_n} = {}_4\phi_3\left(\begin{matrix} q^{-n},\ abcdq^{n-1},\ ae^{-i\theta}, ae^{i\theta} \\ ab,\ ac,\ ad \end{matrix}; q, q\right), \ x = \cos\theta,$$

Askey–Wilson polynomials do not satisfy any of these structural relations.

An extension of Askey's problem is to find a structural relation of type (3.1) that characterises the Askey–Wilson polynomials for $S = \mathcal{D}_q$, the Askey–Wilson divided difference operator, taking $e^{i\theta} = q^s$, with

$$\mathcal{D}_q f(x(s)) = \frac{f(x(s+\frac{1}{2})) - f(x(s-\frac{1}{2}))}{x(s+\frac{1}{2}) - x(s-\frac{1}{2})}, \quad x(s) = \frac{q^{-s} + q^s}{2}.$$

Ismail [23] gave an important hint to the solution of this problem by suggesting that $S = \mathcal{D}_q^2$.

Conjecture 3.1 ([23]) Let $\{p_n(x)\}_{n=0}^{\infty}$ be orthogonal polynomials and π a polynomial of degree at most 4. Then $\{p_n(x)\}$ satisfies

$$\pi(x)\mathcal{D}_q^2 p_n(x) = \sum_{j=-t}^{s} a_{n,j}\, p_{n+j}(x)$$

if and only if $\{p_n(x)\}$ are Askey–Wilson polynomials or special cases of them.

In order to complete and prove the above conjecture, we begin by generalizing a result due to Hahn [20], that a sequence of monic orthogonal polynomials $\{p_n(x)\}_{n=0}^{\infty}$ satisfying

$$\frac{1}{n+1}\frac{dp_{n+1}}{dx}(x) = (x-\widetilde{a}_n)\frac{1}{n}\frac{dp_n}{dx}(x) - \frac{\widetilde{b}_n}{n-1}\frac{dp_{n-1}}{dx}(x), \quad \widetilde{a}_n, \widetilde{b}_n \in \mathbb{R}, \widetilde{b}_n \neq 0,$$

satisfies a second order Sturm-Liouville differential equation of the form

$$\phi(x)\frac{d^2}{dx^2}p_n(x) + \psi(x)\frac{d}{dx}p_n(x) + \lambda_n\, p_n = 0. \tag{3.2}$$

where, ϕ and ψ are polynomials independant of n with $\deg(\phi) \le 2$ and $\deg(\psi) = 1$, while λ_n is a constant dependant on n.

Bochner [5] first considered sequences of polynomials satisfying (3.2) and showed that the orthogonal polynomial solutions of (3.2) are Jacobi, Laguerre and Hermite polynomials. Ismail [22] generalized Bochner's theorem to Askey–Wilson divided difference operators. A more general version is due to Vinet and Zhedanov [40].

Lemma 3.2 ([27]) *Let $\{p_n(x)\}_{n=0}^{\infty}$ a sequence of monic orthogonal polynomials. If there are two sequences (a_n') and (b_n') of numbers such that*

$$\frac{1}{\gamma_{n+1}}\mathcal{D}_q p_{n+1}(x) = (x-a_n')\frac{1}{\gamma_n}\mathcal{D}_q p_n(x) - \frac{b_n'}{\gamma_{n-1}}\mathcal{D}_q p_{n-1}(x) + c_n, \ c_n \in \mathbb{R},$$

then, there are two polynomials $\phi(x)$ and $\psi(x)$ of degree at most two and of degree one respectively and a sequence $\{\lambda_n\}_{n=0}^{\infty}$, such that $p_n(x)$ satisfies the divided-difference equation

$$\phi(x)\mathcal{D}_q^2 p_n(x) + \psi(x)\mathcal{S}_q\mathcal{D}_q p_n(x) + \lambda_n p_n(x) = 0, \ n \ge 5$$

where $\mathcal{S}_q$ is the averaging operator.

$$\mathcal{S}_q\, f(x(s)) = \frac{f(x(s+\frac{1}{2})) + f(x(s-\frac{1}{2}))}{2}. \tag{3.3}$$

Theorem 3.3 ([27]) *Let $\{p_n(x)\}_{n=0}^{\infty}$ be a sequence of monic polynomials orthogonal with respect to a positive weight function $w(x)$. The following properties are equivalent:*

(a) *There is a polynomial $\pi(x)$ of degree at most 4 and constants $a_{n,n+j}$, $-2 \le j \le 2$, $n \ge 2$, $a_{n,n-2} \neq 0$ such that p_n satisfies the structure relation*

$$\pi(x)\mathcal{D}_q^2 p_n(x) = \sum_{j=-2}^{2} a_{n,n+j}\, p_{n+j}(x), \quad n = 2, 3, \ldots;$$

(b) *There is a polynomial $\pi(x)$ of degree at most four such that $\{\mathcal{D}_q^2 p_n\}_{n=2}^{\infty}$ is orthogonal with respect to $\pi(x)\, w(x)$;*
(c) *There are two polynomials $\phi(x)$ and $\psi(x)$ of degree at most two and of degree one respectively and a constant λ_n such that*

$$\phi(x)\mathcal{D}_q^2 p_n(x) + \psi(x)\mathcal{S}_q\mathcal{D}_q p_n(x) + \lambda_n p_n(x) = 0, \quad n = 5, 6, \ldots.$$

Corollary 3.4 ([27]) *A sequence of monic orthogonal polynomials satisfies the relation*

$$\pi(x)\mathcal{D}_q^2 p_n(x) = \sum_{j=-2}^{2} a_{n,n+j}\, p_{n+j}(x), \quad a_{n,n-2} \neq 0, x = \cos\theta,$$

where π is a polynomial of degree at most 4, if and only if $p_n(x)$ is a multiple of the Askey–Wilson polynomial for some parameters $a,\ b,\ c,\ d$, including limiting cases as one or more of the parameters tends to ∞.

Even though Askey–Wilson polynomials (3) are a basic hypergeometric analog of the Wilson polynomials (cf. [29, (9.2.1)]),

$$\frac{W_n(x^2; a,\ b,\ c,\ d)}{(a+b)_n(a+c)_n(a+d)_n} = {}_4F_3\left(\begin{matrix} -n, n+a+b+c+d-1, a-ix, a+ix \\ a+b, a+c, a+d \end{matrix}; 1\right),$$

the coefficients in (3.1) for $S = \delta f(x^2) = f((x+\frac{i}{2})^2) - f((x-\frac{i}{2})^2)$, the Wilson operator, as well as its solutions can not easily be deduced from those of Askey–Wilson polynomials. It therefore is necessary to also consider Ismail's Conjecture 3.1 for the Wilson variable $x(z) = z^2$ $(z = is,\ i^2 = -1)$ as was done in [28].

The explicit structure relation characterizing Askey–Wilson polynomials (cf. [27, Prop 4.2]) is useful to derive bounds for the extreme zeros of the Askey–Wilson polynomials by an application of Theorem 5.1 in the lecture on *Zeros of Orthogonal Polynomials*.

Proposition 3.5 ([27]) *Let $n \in \mathbb{N}$ be fixed and $(x_{n,n})$ be the largest zero of the monic Askey–Wilson polynomial $p_n(x, a, b, c, d|q)$. Then a lower bound for $x_{n,n}$ is*

$$\frac{2(q^{n-1}+1)\left(Aq^{n-1}-a-b-c-d\right)\left(abcdq^{n-1}-1\right)+\sqrt{D_n}}{8\left(abcdq^{2n-2}-1\right)\left(abcdq^{n-1}-1\right)}$$

where

$$A = (abc + abd + acd + bcd)$$

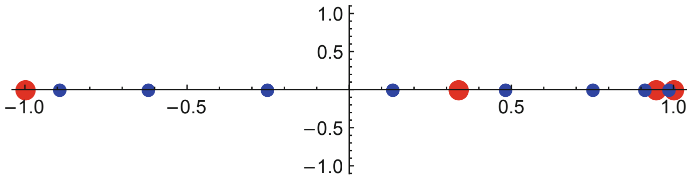


Fig. 6 Plots of zeros of $p_n(x; a, b, c, d|q)$ (blue) for $n = 8$, $a = \frac{6}{7}$, $b = \frac{5}{7}$, $c = \frac{4}{7}$, $d = \frac{3}{7}$, $q = \frac{1}{9}$ and the bounds for the extreme zeros (red)

and

$$\begin{aligned}
D_n = &-(4\left(-q^{3n-3}abcd-1\right)(abcd-ab-ac-ad-bc-bd-cd+1)\\
&+4\,((b^2c^2d^2+b^2c^2+b^2cd+b^2d^2+bc^2d+bcd^2+c^2d^2-bc-bd-cd)a^2\\
&+(bc+bd+cd)\,(bcd-b-c-d)\,a+bcd\,(bcd-b-c-d))q^{2n-2}\\
&+4\,((-bc-bd-cd+1)\,a^2-(bc+bd+cd-1)\,(d+c+b)\,a-b^2cd-bc^2d\\
&-bcd^2+b^2+bc+bd+c^2+cd+d^2+1)q^{n-1})\left(4\,abcdq^{2n-2}-4\right)\left(abcdq^{n-1}-1\right)\\
&+4\left(q^{n-1}+1\right)^2\left(q^{n-1}abc+q^{n-1}abd+q^{n-1}acd+q^{n-1}bcd-a-b-c-d\right)^2\\
&\left(abcdq^{n-1}-1\right)^2
\end{aligned}$$

The zeros of monic Askey–Wilson polynomials $p_n(x; \frac{6}{7}, \frac{5}{7}, \frac{4}{7}, \frac{3}{7}|\frac{1}{9})$ and the bounds obtained from the structure relation for the extreme zeros are illustrated in Fig. 6.

References

1. W.A. Al-Salam, T.S. Chihara, Another characterization of the classical orthogonal polynomials. SIAM J. Math. Anal. **3**, 65–70 (1972)
2. S.M. Alsulami, P. Nevai, J. Szabados, W. Van Assche, A family of nonlinear difference equations: existence, uniqueness, and asymptotic behavior of positive solutions. J. Approx. Theory **193**, 39–55 (2015)
3. A. Arceo, E.J. Huertas, F. Marcellán, On polynomials associated with an Uvarov modification of a quartic potential Freud-like weight. Appl. Math. Comput. **281**, 102–120 (2016)
4. A.F. Beardon, The theorems of Stieltjes and Favard. Comput. Methods Funct. Theory **11**, 247–262 (2011)
5. S. Bochner S, Über Sturm-Liouvillesche Polynomsysteme. Math. Z. **29(1)**, 730–736 (1929)
6. L. Boelen, W. Van Assche, Discrete Painlevé equations for recurrence relations of semiclassical Laguerre polynomials. Proc. Am. Math. Soc. **138**, 1317–1331 (2011)
7. Y. Chen, M.V. Feigin, Painlevé IV and degenerate Gaussian Unitary Ensembles. J. Phys. A Math. Gen. **39**, 12381–12393 (2006)

8. Y. Chen, M.E.H. Ismail, Ladder operators and differential equations for orthogonal polynomials. J. Phys. A **30**, 7817–7829 (1997)
9. T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Gordon and Breach, New York, 1978). Reprinted by Dover Publications (2011)
10. A.S. Clarke, B. Shizgal, On the generation of orthogonal polynomials using asymptotic methods for recurrence coefficients. J. Comput. Phys. **104**, 140–149 (1993)
11. P.A. Clarkson, K. Jordaan, The relationship between semiclassical Laguerre polynomials and the fourth Painlevé equation. Constr. Approx. **39**, 223–254 (2014)
12. P.A. Clarkson, K. Jordaan, Properties of generalized Freud polynomials. J. Approx. Theory **225**, 148–175 (2018)
13. P.A. Clarkson, K. Jordaan, A. Kelil, A generalized Freud weight. Stud. Appl. Math. **136**, 288–320 (2016)
14. S. Datta, J. Griffin, A characterization of some q-orthogonal polynomials. Ramanujan J. **12**(3), 425–437 (2006)
15. K. Driver, K. Jordaan, Bounds for extreme zeros of some classical orthogonal polynomials. J. Approx. Theory **164**, 1200–1204 (2012)
16. J. Favard, Sur les polynomes de Tchebicheff. C. R. Acad. Sci. Paris **200**, 2052–2053 (1935)
17. G. Filipuk, W. Van Assche, L. Zhang, The recurrence coefficients of semiclassical Laguerre polynomials and the fourth Painlevé equation. J. Phys. A **45**, 205201 (2012)
18. G. Freud, On the coefficients in the recursion formulae of orthogonal polynomials. Proc. R. Irish Acad. Sect. A **76**, 1–6 (1976)
19. A.G. García, F. Marcellán, L. Salto, A distributional study of discrete classical orthogonal polynomials. J. Comput. Appl. Math. **57**(1–2), 147–162 (1995)
20. W. Hahn, Über die Jacobischen Polynome und zwei verwandte Polynomklassen. Math. Z. **39(1)**, 634–638 (1935)
21. E. Hendriksen, H. van Rossum, Semi-classical orthogonal polynomials, in *Polynômes Orthogonaux et Applications*, ed. by C. Brezinski, A. Draux, A.P. Magnus, P. Maroni, A. Ronveaux. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 354–361
22. M.E.H. Ismail, A generalization of a theorem of Bochner. J. Comput. Appl. Math. **159**(2), 319–324 (2003)
23. M.E.H. Ismail, in *Classical and Quantum Orthogonal Polynomials in One Variable*. Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
24. M.E.H. Ismail, X. Li, Bounds for extreme zeros of orthogonal polynomials. Proc. Am. Math. Soc. **115**, 131–140 (1992)
25. J. Jost, *Postmodern Aanalysis* (Springer, Berlin, 2006)
26. T. Kasuga, R. Sakai, Orthonormal polynomials with generalized Freud-type weights. J. Approx. Theory **121**, 13–53 (2003)
27. M. Kenfack Nangho, K. Jordaan, A characterization of Askey–Wilson polynomials. Proc. Am. Math. Soc. **147**(6), 2465–2480 (2019)
28. M. Kenfack Nangho, K. Jordaan. Structure relations of classical orthogonal polynomials of the quadratic and$q$-quadratic variable. SIGMA, **14**, 26 (2018)
29. R. Koekoek, P.A. Lesky, R.F. Swarttouw, in *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010)
30. J.S. Lew, D.A. Quarles, Nonnegative solutions of a nonlinear recurrence. J. Approx. Theory **38**, 357–379 (1983)
31. F. Marcellán, R. Álvarez-Nodarse, On the “Favard theorem” and its extensions. J. Comput. Appl. Math. **127**, 231–254 (2001)
32. I.P. Natanson, *Konstruktive Funktionentheorie* (Akademie-Verlag, Berlin, 1955)
33. P. Nevai, Géza Freud, orthogonal polynomials and Christoffel functions. A case study. J. Approx. Theory **48**, 3–167 (1986)
34. F.W.J. Olver, D.W. Lozier, R.F. Boisvert, C.W. Clark (eds), *DLMF Handbook of Mathematical Functions* (Cambridge University Press, Cambridge, 2010)

35. J. Shohat, The Relation of the classical orthogonal polynomials to the polynomials of appell. Am. J. Math. **58**, 453–464 (1936)
36. J. Shohat, Sur les polynômes orthogonaux généralises. C. R. Acad. Sci. Paris **207**, 556–558 (1938)
37. M.H. Stone. *Linear transformations in Hilbert space and their applications to analysis*. AMS Colloquium Publications, vol. 15 (American Mathematical Society, Providence, 1932)
38. T.J. Stieltjes, Recherches sur les fractions continues. Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. **8**, J1–J122 (1894); A5–A47 (1895)
39. W. Van Assche, The impact of Stieltjes work on continued fractions and orthogonal polynomials, in *Collected Papers*, ed. by G. van Dijk (Springer, Berlin, 1993), pp. 5–37
40. L. Vinet, A. Zhedanov, Generalized Bochner theorem: characterization of the Askey–Wilson polynomials. J. Comput Appl. Math. **211**(1), 45–56 (2008)
41. Z.T. Wen, R. Wong and S.X. Xu, Global asymptotics of orthogonal polynomials associated with a generalized Freud weight. Chin. Ann. Math. Ser. B, **39**, 553–596 (2018)
42. A. Wintner, *Spektraltheorie der unendlichen Matrizen* (S. Hirzel, Leipzig, 1929)

Orthogonal Polynomials and Computer Algebra

Wolfram Koepf

Abstract Classical orthogonal polynomials of the Askey–Wilson scheme have extremely many different properties, e.g. satisfying differential equations, recurrence equations, having hypergeometric representations, Rodrigues formulas, generating functions, moment representations etc. Using computer algebra it is possible to switch between one representation and another algorithmically. Such algorithms will be discussed and implementations are presented using Maple.

Keywords Computer algebra · Classical orthogonal polynomials · Askey–Wilson scheme

Mathematics Subject Classification (2000) Primary 33C20, 33F10; Secondary 30B10, 68W30

1 Orthogonal Polynomials

Given: a *scalar product*

$$\langle f, g \rangle := \int_{\alpha}^{\beta} f(x)g(x)\, d\mu(x)$$

with non-negative Borel measure $\mu(x)$ supported in the interval $[\alpha, \beta]$. The following special cases are most important:

- *absolutely continuous* measure $d\mu(x) = \rho(x)\, dx$ with weight function $\rho(x)$,
- *discrete* measure $\mu(x) = \rho(x)$ supported in $\mathbb{Z}$,
- discrete measure $\mu(x) = \rho(x)$ supported in $q^{\mathbb{Z}}$.

W. Koepf (✉)
Institute of Mathematics, University of Kassel, Kassel, Germany
e-mail: koepf@mathematik.uni-kassel.de

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_19

A system of polynomials $(P_n(x))_{n \geqq 0}$

$$P_n(x) = k_n x^n + k'_n x^{n-1} + k''_n x^{n-2} + \cdots , \quad k_n \neq 0 \tag{1.1}$$

is called *orthogonal* (OPS) w.r.t. the *positive-definite* measure $d\mu(x)$, if

$$\langle P_m, P_n \rangle = \begin{cases} 0 & \text{if } m \neq n \\ h_n > 0 & \text{if } m = n \end{cases} .$$

Using the Gram-Schmidt orthogonalization procedure one can compute the orthogonal polynomials $P_n(x)$ iteratively up to a constant standardization factor. One option is to compute the monic system with $k_n = 1$.

We define the scalar product of the Legendre polynomials $P_n(x)$:

```
> ScalarProduct:=proc(f,g,x) int(g*f,x=-1..1) end proc:
```

and declare the Gram-Schmidt procedure

```
> GramSchmidt := proc (n, x) local j, k, g, liste;
>     liste := [seq(x^j, j = 0 .. n)]; g(0) := 1;
>     for j to n do g(j):=op(j+1,liste)-
>     add(ScalarProduct(op(j+1,liste),g(k),x)*g(k)/
>     ScalarProduct(g(k), g(k), x), k = 0 .. j-1) end do;
>      [seq(g(j), j = 0 .. n)]
> end proc:
```

Now we can compute using the Gram-Schmidt procedure.

```
> SEQ1 := GramSchmidt(10, x);
```

$$SEQ1 := [1, x, x^2 - 1/3, x^3 - 3/5\,x, x^4 + \frac{3}{35} - 6/7\,x^2, x^5 + \frac{5\,x}{21} - \frac{10\,x^3}{9},$$
$$x^6 - \frac{5}{231} + \frac{5\,x^2}{11} - \frac{15\,x^4}{11}, x^7 - \frac{35\,x}{429} + \frac{105\,x^3}{143} - \frac{21\,x^5}{13},$$
$$x^8 + \frac{7}{1287} - \frac{28\,x^2}{143} + \frac{14\,x^4}{13} - \frac{28\,x^6}{15}, x^9 + \frac{63\,x}{2431} - \frac{84\,x^3}{221}$$
$$+ \frac{126\,x^5}{85} - \frac{36\,x^7}{17}, x^{10} - \frac{63}{46189} + \frac{315\,x^2}{4199} - \frac{210\,x^4}{323} + \frac{630\,x^6}{323} - \frac{45\,x^8}{19}]$$

This computation has created the first 11 monic Legendre polynomials. Of course, *Maple* knows the Legendre polynomials internally as `LegendreP(k,x)`:

```
> SEQ2:=expand([seq(LegendreP(k,x),k=0..10)]);
```

$$SEQ2 := [1, x, -\frac{1}{2} + \frac{3}{2}\,x^2, \frac{5}{2}\,x^3 - \frac{3}{2}\,x, \frac{3}{8} + \frac{35\,x^4}{8} - \frac{15\,x^2}{4}, \frac{63\,x^5}{8} - \frac{35\,x^3}{4} + \frac{15\,x}{8},$$
$$-\frac{5}{16} + \frac{231\,x^6}{16} - \frac{315\,x^4}{16} + \frac{105\,x^2}{16}, \frac{429\,x^7}{16} - \frac{693\,x^5}{16} + \frac{315\,x^3}{16} - \frac{35\,x}{16}, \frac{35}{128} + \frac{6435\,x^8}{128}$$
$$- \frac{3003\,x^6}{32} + \frac{3465\,x^4}{64} - \frac{315\,x^2}{32}, \frac{12155\,x^9}{128} - \frac{6435\,x^7}{32} + \frac{9009\,x^5}{64} - \frac{1155\,x^3}{32} + \frac{315\,x}{128},$$
$$-\frac{63}{256} + \frac{46189\,x^{10}}{256} - \frac{109395\,x^8}{256} + \frac{45045\,x^6}{128} - \frac{15015\,x^4}{128} + \frac{3465\,x^2}{256}]$$

Obviously the ratios of the corresponding polynomials must be constant:

```
> normal([seq(op(k,SEQ2)/op(k,SEQ1),k=1..11)]);
```

$$[1, 1, 3/2, 5/2, \frac{35}{8}, \frac{63}{8}, \frac{231}{16}, \frac{429}{16}, \frac{6435}{128}, \frac{12155}{128}, \frac{46189}{256}]$$

Every OPS has the following main properties:

- **(Three-term Recurrence)** Every OPS satisfies

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

- **(Zeros)** All zeros of an OPS are simple, lie in the interior of $[\alpha, \beta]$ and have some nice interlacing properties.
- **(Hankel Matrix: Representation by Moments)**

$$P_n(x) = C_n \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n+1} & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix},$$

where $\mu_n := \int_a^b x^n d\mu(x)$ denote the moments of $d\mu(x)$.

We can compute the moment matrix by

```
> Momentmatrix := proc (n)
> local j, k;
> convert([seq([seq(ScalarProduct(x^j, x^k, x), j = 0 .. n)],
> k = 0 .. n)], Matrix)
> end proc:
```

```
> H := Momentmatrix(5);
```

$$\begin{pmatrix} 2 & 0 & 2/3 & 0 & 2/5 & 0 \\ 0 & 2/3 & 0 & 2/5 & 0 & 2/7 \\ 2/3 & 0 & 2/5 & 0 & 2/7 & 0 \\ 0 & 2/5 & 0 & 2/7 & 0 & 2/9 \\ 2/5 & 0 & 2/7 & 0 & 2/9 & 0 \\ 0 & 2/7 & 0 & 2/9 & 0 & 2/11 \end{pmatrix}$$

and the Hankel matrix is given by

```
> Hankelmatrix := proc (n)
> local j, k, m;
> m := [seq([seq(ScalarProduct(t^j,t^k,t),j=0..n)],k=0..n-1)];
> m := [op(m), [seq(x^k, k = 0 .. n)]];
> convert(m, Matrix)
> end proc:
```

Its determinant gives a multiple of the orthogonal polynomial whose degree is the size of the square matrix minus 1. The determinant of the following matrix is therefore a multiple of $P_5(x)$:

```
> Hankelmatrix(5);
```

$$\begin{pmatrix} 2 & 0 & 2/3 & 0 & 2/5 & 0 \\ 0 & 2/3 & 0 & 2/5 & 0 & 2/7 \\ 2/3 & 0 & 2/5 & 0 & 2/7 & 0 \\ 0 & 2/5 & 0 & 2/7 & 0 & 2/9 \\ 2/5 & 0 & 2/7 & 0 & 2/9 & 0 \\ 1 & x & x^2 & x^3 & x^4 & x^5 \end{pmatrix}$$

and the first six multiples of the Legendre polynomials are given by

```
> SEQ3:=[seq(LinearAlgebra[Determinant](Hankelmatrix(n)),
>     n=0..5)];
```

$$SEQ3 := \Big[1, 2\,x, 4/3\,x^2 - 4/9, \frac{32\,x^3}{135} - \frac{32\,x}{225}, \frac{256\,x^4}{23625} - \frac{512\,x^2}{55125} + \frac{256}{275625}, \frac{32768\,x^5}{260465625} - \frac{65536\,x^3}{468838125} + \frac{32768\,x}{1093955625}\Big]$$

Again, the ratios of the corresponding polynomials must be constant:

```
> normal([seq(op(k,SEQ3)/op(k,SEQ1),k=1..6)]);
```

$$[1, 2, 4/3, \frac{32}{135}, \frac{256}{23625}, \frac{32768}{260465625}]$$

2 Classical Orthogonal Polynomials

The *classical* OPS $(P_n(x))_{n \geqq 0}$ can be defined as the polynomial solutions of the *differential equation*:

$$\sigma(x) P_n''(x) + \tau(x) P_n'(x) - \lambda_n P_n(x) = 0 \,. \tag{2.1}$$

Substituting (1.1) into (2.1), we conclude:

- $n = 1$ yields $\tau(x) = dx + e, d \neq 0$,
- $n = 2$ yields $\sigma(x) = ax^2 + bx + c$,
- The coefficient of x^n yields $\lambda_n = n(a(n-1) + d)$.

These classical families can be classified (modulo linear transformations) according to the following scheme (Bochner [2])

- $\sigma(x) = 0$ powers x^n,
- $\sigma(x) = 1$ *Hermite* polynomials,
- $\sigma(x) = x$ *Laguerre* polynomials,
- $\sigma(x) = 1 - x^2$ *Jacobi* polynomials,
- $\sigma(x) = x^2$ *Bessel* polynomials.

For the theory one needs

- a representing basis $f_n(x)$, here the powers $f_n(x) = x^n$;
- an operator, here the derivative operator D, with $D\, f_n(x) = n\, f_{n-1}(x)$.

The corresponding weight function $\rho(x)$ satisfies the *Pearson Differential Equation*

$$\frac{d}{dx}\Big(\sigma(x)\rho(x)\Big) = \tau(x)\rho(x)\ . \tag{2.2}$$

Hence the weight function is given by

$$\rho(x) = \frac{C}{\sigma(x)} e^{\int \frac{\tau(x)}{\sigma(x)} dx}\ .$$

The following properties are equivalent, each defining the classical continuous families:

- Differential equation (2.1) for $(P_n(x))_{n \geqq 0}$.
- Pearson differential equation (2.2) $(\sigma\ \rho)' = \tau\ \rho$ for the weight $\rho(x)$.
- With $(P_n(x))_{n \geqq 0}$ also $(P'_{n+1}(x))_{n \geqq 0}$ is an OPS.
- **Derivative Rule**:

$$\sigma(x)\ P'_n(x) = \alpha_n\ P_{n+1}(x) + \beta_n\ P_n(x) + \gamma_n\ P_{n-1}(x)\ .$$

- **Structure Relation**: $P_n(x)$ satisfies

$$P_n(x) = \widehat{a}_n\ P'_{n+1}(x) + \widehat{b}_n\ P'_n(x) + \widehat{c}_n\ P'_{n-1}(x)\ .$$

- **Rodrigues Formula**: $P_n(x)$ is given as

$$P_n(x) = \frac{E_n}{\rho(x)} \frac{d^n}{dx^n}\Big(\rho(x)\,\sigma(x)^n\Big)\ .$$

3 Classical Discrete Orthogonal Polynomials

The *classical discrete* OPS can be analogously defined as the solutions of the *difference equation* [12]:

$$\sigma(x)\Delta\nabla P_n(x) + \tau(x)\Delta P_n(x) - \lambda_n P_n(x) = 0 \tag{3.1}$$

where $\Delta f(x) = f(x+1) - f(x)$ and $\nabla f(x) = f(x) - f(x-1)$ denote the forward and backward difference operators. As in the continuous case, we get

- $n = 1$ yields $\tau(x) = dx + e, d \neq 0$,
- $n = 2$ yields $\sigma(x) = ax^2 + bx + c$,
- The coefficient of x^n yields $\lambda_n = n(a(n-1) + d)$.

The classical discrete families can be classified (modulo linear transformations) according to the following scheme ([12], see also [14]):

- $\sigma(x) = 0$ falling factorials $x^{\underline{n}} = x(x-1)\cdots(x-n+1)$,
- $\sigma(x) = 1$ shifted **Charlier** polynomials,
- $\sigma(x) = x$ **Charlier**, **Meixner**, **Krawtchouk** polynomials,
- $\deg(\sigma(x), x) = 2$ **Hahn** polynomials.

For the theory one needs

- a representing basis $f_n(x)$, here the falling factorial $f_n(x) = x^{\underline{n}}$;
- an operator, here the operator Δ, with $\Delta\, f_n(x) = n\, f_{n-1}(x)$.

The corresponding discrete weight function $\rho(x)$ satisfies the Pearson difference equation

$$\Delta\Big(\sigma(x)\rho(x)\Big) = \tau(x)\rho(x)\ .$$

Hence it is given by the term ratio

$$\frac{\rho(x+1)}{\rho(x)} = \frac{\sigma(x)+\tau(x)}{\sigma(x+1)}\ . \tag{3.2}$$

We would like to put our results into the general framework of hypergeometric functions.

The power series

$${}_pF_q\left(\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix} \,\middle|\, z\right) = \sum_{k=0}^{\infty} A_k\, z^k\ ,$$

whose summands $\alpha_k = A_k z^k$ have a rational term ratio

$$\frac{\alpha_{k+1}}{\alpha_k} = \frac{A_{k+1}\, z^{k+1}}{A_k\, z^k} = \frac{(k+a_1)\cdots(k+a_p)}{(k+b_1)\cdots(k+b_q)}\,\frac{z}{(k+1)}\ ,$$

is called the *generalized hypergeometric series*. The summand $\alpha_k = A_k z^k$ of a hypergeometric series is called a *hypergeometric term*.

The relation (3.2) therefore tells that the weight function $\rho(x)$ of the classical discrete orthogonal polynomials is a hypergeometric term w.r.t. the variable x.

For the coefficients of the generalized hypergeometric series one gets the following formula

$${}_pF_q\left(\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix} \,\middle|\, z\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k}\,\frac{z^k}{k!}$$

using the *Pochhammer symbol* $(a)_k = a(a+1)\cdots(a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)}$.

From the differential equation (2.1) one can compute a recurrence equation for the corresponding power series coefficients [16]. Using *Maple*, we get

```
> sigma := a*x^2+b*x+c; tau := d*x+e;
```

$$\sigma := ax^2 + xb + c$$

$$\tau := dx + e$$

```
> DE := sigma*(diff(F(x), x$2))+tau*(diff(F(x), x))
>     -n*(a*n-a+d)*F(x);
```

$$DE := \left(ax^2 + xb + c\right) \frac{d^2}{dx^2} F(x) + (dx + e) \frac{d}{dx} F(x) - n(an - a + d) F(x)$$

This differential equation is converted towards the recurrence equation

```
> RE := gfun[diffeqtorec](DE, F(x), A(k));
```

$$RE := \left(ak^2 + (-a + d)\, k - an^2 + an - dn\right) A(k)$$
$$+ \left(bk^2 + (b + e)\, k + e\right) A(k + 1) + \left(ck^2 + 3\,c\,k + 2\,c\right) A(k + 2)$$

The Laguerre polynomials have the data

```
> laguerre := {a = 0, b = 1, c = 0, d = -1, e = alpha+1};
```

$$laguerre := \{a = 0, b = 1, c = 0, d = -1, e = \alpha + 1\}$$

so that we get for their power series coefficients A_k

```
> laguerreRE := subs(laguerre, RE);
```

$$laguerreRE := (-k + n)\, A(k) + \left(k^2\, (2 + \alpha)\, k + \alpha + 1\right) A(k + 1)$$

Therefore their quotient A_{k+1}/A_k is given in factored form by

```
> quotient := factor(solve(laguerreRE, A(k+1))/A(k));
```

$$quotient := \frac{k - n}{(k + 1)\,(k + \alpha + 1)}$$

from which one can read off directly the hypergeometric representation

```
> lag := hypergeom([-n], [alpha+1], x);
```

$$lag := {}_1\mathrm{F}_1(-n;\ \alpha + 1;\ x)$$

By an internal command, *Maple* can convert this towards

```
> convert(lag, StandardFunctions);
```

$$\frac{\Gamma(n + 1)\,\Gamma(\alpha + 1)\,LaguerreL(n, \alpha, x)}{\Gamma(n + \alpha + 1)}$$

back, again. We have therefore seen that using this approach one gets for the Laguerre polynomials

$$L_n^{\alpha}(x) = \binom{n+\alpha}{n} {}_1F_1\left(\left.\begin{array}{c} -n \\ \alpha + 1 \end{array}\right| x\right) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n+\alpha}{n-k} x^k\,.$$

Another example are the Hahn polynomials which are given by

$$Q_n^{(\alpha,\beta)}(x,N) = {}_3F_2\left(\left.\begin{array}{c} -n, -x, n+1+\alpha+\beta \\ \alpha+1, -N \end{array}\right| 1\right).$$

Similarly, all the other classical systems have a hypergeometric representation. These can be found in every book about OPS, e.g. in [7], and on the CAOP web page [9].

4 Classical q-Orthogonal Polynomials and the Askey–Wilson Scheme

The *classical* q-OPS can be analogously defined as the polynomial solutions of the q-*difference equation* (Hahn [6]):

$$\sigma(x)D_qD_{1/q}P_n(x) + \tau(x)D_qP_n(x) - \lambda_{n,q}P_n(x) = 0 \tag{4.1}$$

where $D_q f(x) = \frac{f(x)-f(qx)}{(1-q)x}$ is the Hahn operator. As before, we can conclude

- $n = 1$ yields $\tau(x) = dx + e, d \neq 0$,
- $n = 2$ yields $\sigma(x) = ax^2 + bx + c$,
- The coefficient of x^n yields $\lambda_{n,q} = [n]_q(a[n-1]_q + d[n]_q)$ where $[n]_q = \frac{1-q^n}{1-q}$ is the q-bracket.

The classical q-discrete families of the Hahn class considered can be classified (modulo linear transformations) according to the following list: *Big q-Jacobi polynomials, q-Hahn polynomials, Big q-Laguerre polynomials, Al-Salam-Carlitz I polynomials, discrete q-Hermite I polynomials, Little q-Jacobi polynomials, alternative q-Charlier polynomials, Little q-Laguerre polynomials, q-Meixner polynomials, Stieltjes-Wigert polynomials, q-Laguerre polynomials, q-Charlier polynomials, Al-Salam-Carlitz II polynomials, discrete q-Hermite II polynomials*, see e.g. [7] and [9].

For the theory one needs:

- two representing bases $f_n(x)$, here $f_n(x) = x^n$ and $g_n(x) = (x;q)_n$ where $(x;q)_n = (1-x)(1-xq)\cdots(1-xq^{n-1})$ is the *q-Pochhammer symbol*;
- an operator, here the operator D_q, with $D_q\, f_n(x) = [n]_q\, f_{n-1}(x)$ and a similar relation for $g_n(x)$.

The corresponding q-discrete weight function $\rho(x)$ satisfies the Pearson q-difference equation

$$D_q\Big(\sigma(x)\rho(x)\Big) = \tau(x)\rho(x)\ .$$

Hence it is given by the term ratio:

$$\frac{\rho(qx)}{\rho(x)} = \frac{\sigma(x) + (q-1)x\tau(x)}{\sigma(qx)} . \tag{4.2}$$

The power series

$${}_r\phi_s \left(\begin{array}{c} a_1, \ldots, a_r \\ b_1, \ldots, b_s \end{array} \middle| \, q \, ; \, z \right) = \sum_{k=0}^{\infty} A_k \, z^k ,$$

whose summands $\alpha_k = A_k z^k$ are given by

$$A_k \, z^k = \frac{(a_1; q)_k \cdots (a_r; q)_k}{(b_1; q)_k \cdots (b_s; q)_k} \frac{z^k}{(q; q)_k} \left((-1)^k q^{\binom{k}{2}} \right)^{1+s-r}$$

is called the *basic hypergeometric series*. The summand $\alpha_k = A_k z^k$ of a basic hypergeometric series is called a *q-hypergeometric term*.

The relation (4.2) therefore tells that the weight function $\rho(x)$ of the classical q-orthogonal polynomials is a q-hypergeometric term w.r.t. the variable x.

In CAOP [9] you saw all the families of the *Askey–Wilson Scheme*. This scheme contains

- continuous measures supported in an interval (classical continuous OPS);
- discrete measures supported in $\mathbb{Z}$ (classical discrete OPS);
- discrete measures supported in $q^{\mathbb{Z}}$ (**Hahn tableau**);
- discrete measures supported on a quadratic lattice (**Wilson tableau**);
- discrete measures supported on a q-quadratic lattice (**Askey–Wilson tableau**).

It turns out that the last two classes can be treated in a similar way as the continuous and the discrete cases resulting in a similar theory [3].

5 Computer Algebra Applied to Classical Orthogonal Polynomials

Using linear algebra one can compute the coefficients of the following identities—expressed through the parameters a, b, c, d and e from the defining equations (2.1), (3.1) or (4.1)—(Lesky [13]):

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

(DR) $$\sigma(x) \, P_n'(x) = \alpha_n \, P_{n+1}(x) + \beta_n \, P_n(x) + \gamma_n \, P_{n-1}(x)$$

(SR) $$P_n(x) = \widehat{a}_n \, P_{n+1}'(x) + \widehat{b}_n \, P_n'(x) + \widehat{c}_n \, P_{n-1}'(x)$$

We define $P_n(x)$, given by (1.1) in *Maple*, and substitute the three highest coefficients—which will be sufficient for our purposes—into the differential equation:

```
> p := k[n]*x^n+kprime[n]*x^(n-1)+kprimeprime[n]*x^(n-2);
```

$$p := k_n x^n + kprime_n x^{n-1} + kprimeprime_n x^{n-2}$$

```
> DE := sigma*(diff(p, x$2))+tau*(diff(p, x))-lambda[n]*p:
```

We divide by x^{n-4} so that the result is a polynomial of degree 4 (whose three highest coefficients are those of x^4, x^3 and x^2)

```
> de := collect(simplify(DE/x^(n-4)), x):
```

Equating the highest coefficients yields the relation for λ_n that we already met:

```
> rule1 := lambda[n] = solve(coeff(de, x, 4), lambda[n]);
```

$$rule1 := \lambda_n = n\,(an - a + d)$$

Next, we substitute λ_n into the differential equation and equate the next highest coefficient. This shows that the second highest coefficient k_n' of $P_n(x)$ is a rational multiple of the leading coefficient k_n:

```
> de := expand(subs(rule1, de)):
> rule2 := kprime[n] = solve(coeff(de, x, 3), kprime[n]);
```

$$rule2 := kprime_n = \frac{nk_n\,(bn - b + e)}{2\,an - 2\,a + d}$$

In the last step we deduce that generically k_n'' is also a rational multiple of k_n:

```
> rule3 := kprimeprime[n] =
> solve(coeff(subs(rule2, de), x, 2), kprimeprime[n]);
```

$$rule3 := kprimeprime_n = k_n\,n \cdot \frac{\left(b^2n^3+2\,acn^2-4b^2n^2+2ben^2-4acn+5b^2n-5ben+cdn+e^2n+2ac-2b^2+3be-cd-e^2\right)}{2\,(2\,an-2\,a+d)\,(2\,an-3\,a+d)}$$

In the sequel, we consider without loss of generality the monic case.

```
> k[n] := 1;
```

$$k_n := 1$$

To get information about the coefficients of the recurrence equation, we put it in the following form to be zero.

```
> RE := x*P(n)-a[n]*P(n+1)-b[n]*P(n)-c[n]*P(n-1);
```

$$RE := x\,P\,(n) - a_n\,P\,(n+1) - b_n\,P\,(n) - c_n\,P\,(n-1)$$

After substituting $P_n(x)$, given by (1.1), and the previous results about k_n' (`rule2`) and k_n'' (`rule3`), and by equating again the three highest coefficients, we get:

```
> RE:=subs({P(n)=p,P(n-1)=subs(n=n-1,p),P(n+1)=subs(n=n+1,p)},
  RE):
> RE:=subs({rule2,rule3,subs(n=n-1,rule2),subs(n=n-1,rule3),
> subs(n = n+1, rule2), subs(n = n+1, rule3)}, RE):
> re := simplify(numer(normal(RE))/x^(n-3)):
```

```
> rule4 := a[n] = solve(coeff(re, x, 4), a[n]);
```

$$rule4 := a_n = 1$$

```
> rule5 := b[n] = factor(solve(subs(rule4, coeff(re, x, 3)),
  b[n]));
```

$$rule5 := b_n = -\frac{2\,abn^2 - 2\,abn + 2\,bdn - 2\,ae + de}{(2\,an - 2\,a + d)\,(2\,an + d)}$$

```
> rule6 := c[n] =
> factor(solve(subs(rule5, subs(rule4, coeff(re, x, 2))),
>    c[n]));
```

$$rule6 := c_n = -\frac{n\,(an - 2\,a + d)}{(2\,an - 2\,a + d)^2\,(2\,an - 3\,a + d)\,(2\,an - a + d)} \cdot$$
$$\left(4a^2cn^2 - ab^2n^2 - 8a^2cn + 2ab^2n + 4acdn - b^2dn + 4a^2c - ab^2 - 4acd + ae^2 + b^2d - bde + cd^2\right)$$

Such relations were given generically in the paper [10] for the continuous and the discrete cases, and in later papers extended to the q-case ([4, 11]) and to the quadratic case ([5, 15, 18]).

They can be used to compute power series coefficients, inversion coefficients, connection coefficients and parameter derivatives, as e.g. [10]

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

We have shown that the coefficients of the recurrence equation of the classical systems can be written in terms of the coefficients $a, b, c, d,$ and e of the differential/difference equation.

If one uses these formulas in the backward direction, then one can determine the possible differential/difference equations from a given recurrence. For this purpose one must solve a *non-linear system*.

Assume the following recurrence equation is given:

$$P_{n+2}(x) - (x - n - 1)\, P_{n+1}(x) + \alpha(n+1)^2 P_n(x) = 0 \,.$$

Does this equation have classical OPS solutions?

We find out [11] that the solutions of this equation are shifted Laguerre polynomials for $\alpha = 1/4$. For $\alpha < 1/4$ the recurrence has Meixner and Krawtchouk polynomial solutions.

```
> read "hsum17.mpl";
   `Package "Hypergeometric Summation", Maple V - Maple 17`
  `Copyright 1998-2013, Wolfram Koepf, University of Kassel`
> read "retode.mpl";
             `Package "REtoDE", Maple V - Maple 8`
  `Copyright 2000-2002, Wolfram Koepf, University of Kassel`
> RE := P(n+2)-(x-n-1)*P(n+1)+alpha*(n+1)^2*P(n) = 0;
```

$$RE := P(n+2) - (x - n - 1)\, P(n+1) + \alpha\, (n+1)^2\, P(n) = 0$$

```
> REtoDE(RE, P(n), x);
```

'Warning: parameters have the values'
$\{a = 0, \alpha = 1/4, b = -d/2, c = -d/4, d = d, e = 0\}$

$$\left[1/2\ (2\,x+1)\,\frac{\partial^2}{\partial x^2}P\,(n,x) - 2\,x\,\frac{\partial}{\partial x}P\,(n,x) + 2\,n\,P\,(n,x) = 0,\right.$$
$$\left.\left[I = [-1/2, \infty], \rho\,(x) = 2\,\mathrm{e}^{-2\,x}\right], \frac{k_{n+1}}{k_n} = 1\right]$$

```
> REtodiscreteDE(RE, P(n), x);
```

For the last computation we omit the lengthy output and just state that the difference equation and weight of the Meixner and Krawtchouk polynomials is discovered.

Recently Walter Van Assche asked me the question to find all OPS of the Askey–Wilson scheme that satisfy a certain recurrence equation, see [19]. Dr. Daniel Tcheutia solved this question completely [17] by extending the above algorithm to the quadratic lattice. The answer is: The adapted algorithm finds the solutions to the first question (that were already know to Walter Van Assche). This algorithm also *proves that the second recurrence equation does not have such solutions.*

The *Legendre Polynomials* $P_n(x)$ of the Jacobi class have several representations as series:

$$\begin{aligned} P_n(x) &= \sum_{k=0}^{n} \binom{n}{k}\binom{-n-1}{k}\left(\frac{1-x}{2}\right)^k \\ &= \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}^2 (x-1)^{n-k}(x+1)^k \\ &= \frac{1}{2^n}\sum_{k=0}^{\lfloor n/2\rfloor}(-1)^k\binom{n}{k}\binom{2n-2k}{n}x^{n-2k}\ . \end{aligned}$$

It is already non-trivial to identify that these three series represent the same functions, but Zeilberger's algorithm [8] computes the desired normal forms, namely the corresponding (and identical) recurrence equations jointly with enough initial values.

```
> legendreterm1:=binomial(n,k)*binomial(-n-1,k)
>     *((1-x)*(1/2))^k;
```

$$\mathit{legendreterm1} := \binom{n}{k}\binom{-n-1}{k}(1/2 - x/2)^k$$

```
> legendreterm2 := binomial(n, k)^2*(x-1)^(n-k)*(x+1)^k/2^n;
```

$$\mathit{legendreterm2} := \frac{\left(\binom{n}{k}\right)^2 (x-1)^{-k+n}\,(x+1)^k}{2^n}$$

```
> legendreterm3 :=
> (-1)^k*binomial(n, k)*binomial(2*n-2*k, n)*x^(n-2*k)/2^n;
```

$$\mathit{legendreterm3} := \frac{(-1)^k\,\binom{n}{k}\binom{2\,n-2\,k}{n}x^{n-2\,k}}{2^n}$$

```
> RE1 := {sumrecursion(legendreterm1, k, P(n)),
> P(0) = add(subs(n = 0, legendreterm1), k = 0 .. 0),
> P(1) = add(subs(n = 1, legendreterm1), k = 0 .. 1)};
```

$$RE1 := \{(n+2)\,P(n+2) - x\,(2\,n+3)\,P(n+1) + (n+1)\,P(n) = 0, P(0) = 1, P(1) = x\}$$

```
> RE2 := {sumrecursion(legendreterm2, k, P(n)),
> P(0) = add(subs(n = 0, legendreterm2), k = 0 .. 0),
> P(1) = add(subs(n = 1, legendreterm2), k = 0 .. 1)};
```

$$RE2 := \{(n+2)\,P(n+2) - x\,(2\,n+3)\,P(n+1) + (n+1)\,P(n) = 0, P(0) = 1, P(1) = x\}$$

```
> RE3 := {sumrecursion(legendreterm3, k, P(n)),
> P(0) = expand(add(subs(n = 0, legendreterm3), k = 0 .. 0)),
> P(1) = expand(add(subs(n = 1, legendreterm3), k = 0 .. 1))};
```

$$RE3 := \{(n+2)\,P(n+2) - x\,(2\,n+3)\,P(n+1) + (n+1)\,P(n) = 0, P(0) = 1, P(1) = x\}$$

The above computations have computed the normal forms of each of the three different series representations. Since they agree, we have proved that the series represent the same family of functions.

Next, we compute their hypergeometric representations.

```
> Sumtohyper(legendreterm1, k);
```

$$\mathit{Hypergeom}\,([n+1, -n], [1], 1/2 - x/2)$$

```
> Sumtohyper(legendreterm2, k);
```

$$\frac{(x-1)^n}{2^n} \mathit{Hypergeom}\left([-n, -n], [1], \frac{x+1}{x-1}\right)$$

```
> convert(Sumtohyper(legendreterm3, k), binomial);
```

$$\binom{2n}{n} \frac{x^n\,\mathit{Hypergeom}([-n/2, 1/2 - n/2], [-n+1/2], x^{-2})}{2^n}$$

It can also be easily shown that they satisfy the same differential equation.

```
> DE1 := sumdiffeq(legendreterm1, k, P(x));
```

$$DE1 := \left(\frac{d^2}{dx^2} P(x)\right)(x-1)(x+1) + 2\,x \frac{d}{dx} P(x) - n\,(n+1)\,P(x) = 0$$

```
> DE2 := sumdiffeq(legendreterm2, k, P(x));
```

$$DE2 := \left(\frac{d^2}{dx^2} P(x)\right)(x-1)(x+1) + 2\,x \frac{d}{dx} P(x) - n\,(n+1)\,P(x) = 0$$

```
> DE3 := sumdiffeq(legendreterm3, k, P(x));
```

$$DE3 := \left(\frac{d^2}{dx^2} P(x)\right)(x-1)(x+1) + 2\,x \frac{d}{dx} P(x) - n\,(n+1)\,P(x) = 0$$

In the talk given by Naoures Ayadi [1], she introduced the Meixner type polynomials

$$\widehat{M}_n^{\beta_1,\beta_2}(x,c) = (\beta_1)_k\,(\beta_2)_k\,{}_2F_2\left(\begin{array}{c} -n, -x \\ \beta_1, \beta_2 \end{array} \middle| \, \frac{1}{c}\right).$$

Using Zeilberger's algorithm, it is easy to get a recurrence equation for $M_n^{\beta_1,\beta_2}(x,c)$.

```
> meixnersummand := pochhammer(beta[1], n)*pochhammer(beta[2], n)*
> hyperterm([-n, -x], [beta[1], beta[2]], 1/c, k);
```

$$\mathit{meixnersummand} := \frac{\mathit{pochhammer}(\beta_1,n)\,\mathit{pochhammer}(\beta_2,n)\,\mathit{pochhammer}(-n,k)\,\mathit{pochhammer}(-x,k)\,\left(c^{-1}\right)^k}{\mathit{pochhammer}(\beta_1,k)\,\mathit{pochhammer}(\beta_2,k)\,k!}$$

```
> MeixnerRE := sumrecursion(meixnersummand, k, M(n));
```

$$\begin{aligned}\mathit{MeixnerRE} &:= -cM(n+3)\\ &+\left(3\,cn^2+2\,cn\beta_1+2\,cn\beta_2+c\beta_1\beta_2+11\,cn+4\,c\beta_1+4\,c\beta_2+10\,c-n+x-2\right)M(n+2)\\ &-(n+2)(1+\beta_2+n)(1+\beta_1+n)(3\,cn+c\beta_1+c\beta_2+4\,c-1)\,M(n+1)\\ &+c\,(n+2)(n+1)(1+\beta_2+n)(\beta_2+n)(1+\beta_1+n)(\beta_1+n)\,M(n)=0\end{aligned}$$

The more complicated family

$$M_n^{\beta_1,\beta_2,\beta_3}(x,c)=(\beta_1)_k\,(\beta_2)_k\,(\beta_3)_k\,{}_2F_3\left(\left.\begin{array}{c}-n,-x\\ \beta_1,\beta_2,\beta_3\end{array}\right|\frac{1}{c}\right)$$

is similarly feasible.

6 Epilogue

Software developers love when their software is used. But they need your support. Hence my suggestion: If you use one of the packages mentioned for your scientific work, please cite its use!

References

1. N. Ayadi, Classical discrete d-orthogonal polynomials. Contributed Talk 27 at the AIMS-Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications, Douala, Cameroon, October 5–12, 2018. http://www.aims-volkswagen-workshops.org/workshop-information.html
2. S. Bochner, Über Sturm-Liouvillesche Polynomsysteme. Math. Z. **29**, 730–736 (1929)
3. M. Foupouagnigni, On difference equations for orthogonal polynomials on nonuniform lattices. J. Differ. Equ. Appl. **14**, 127–174 (2008)
4. M. Foupouagnigni, W. Koepf, D.D. Tcheutia, P. Njionou Sadjang, Representations of q-orthogonal polynomials. J. Symbolic Comput. **47**, 1347–1371 (2012)
5. M. Foupouagnigni, W. Koepf, D.D. Tcheutia, Connection and linearization coefficients of the Askey–Wilson polynomials. J. Symbolic Comput. **53**, 96–118 (2013)
6. W. Hahn, Über Orthogonalsysteme, die q-Differenzengleichungen genügen. Math. Nachr. **2**, 4–34 (1949)
7. R. Koekoek, P. Lesky, R.F. Swarttouw, in *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010)

8. W. Koepf, *Hypergeometric Summation. An Algorithmic Approach to Summation and Special Function Identities*, 1st edn. (Vieweg, Braunschweig–Wiesbaden, 1998); 2nd edn, Springer Universitext (Springer, London, 2014)
9. W. Koepf, *Computer Algebra and Orthogonal Polynomials*. www.caop.org
10. W. Koepf, D. Schmersau, Representations of orthogonal polynomials. J. Comput. Appl. Math. **90**, 57–94 (1998)
11. W. Koepf, D. Schmersau, Recurrence equations and their classical orthogonal polynomial solutions. Appl. Math. Comput. **128**, 303–327 (2002)
12. P. Lesky, Über Polynomsysteme, die Sturm-Liouvilleschen Differenzengleichungen genügen. Math. Z. **78**, 439–445 (1962)
13. P. Lesky, Über Polynomlösungen von Differentialgleichungen und Differenzengleichungen zweiter Ordnung. Anzeiger der Österreichischen Akademie der Wissenschaften. Math.-naturwiss. Klasse **121**, 29–33 (1985)
14. A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable* (Springer, Berlin, 1991)
15. P. Njionou Sadjang, W. Koepf, M. Foupouagnigni, On structure formulas for Wilson polynomials. Integral Transform. Spec. Funct. **26**, 1000–1014 (2015)
16. B. Salvy, P. Zimmermann, GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. ACM Trans. Math. Softw. **20**, 163–177 (1994)
17. D. Tcheutia, Recurrence equations and their classical orthogonal polynomial solutions on a quadratic or q-quadratic lattice. J. Differ. Equ. Appl. 2019. https://doi.org/10.1080/10236198.2019.1627346
18. D.D. Tcheutia, P. Njionou Sadjang, W. Koepf, M. Foupouagnigni, Divided-difference equation, inversion, connection, multiplication and linearization formulae of the continuous Hahn and the Meixner-Pollaczek Polynomials. Ramanujan J. **45**, 33–56 (2018)
19. W. Van Assche, Solution of an open problem about two families of orthogonal polynomials. SIGMA **15**, 005 (2019)

Spin Chains, Graphs and State Revival

Hiroshi Miki, Satoshi Tsujimoto, and Luc Vinet

Abstract Connections between the 1-excitation dynamics of spin lattices and quantum walks on graphs will be surveyed. Attention will be paid to perfect state transfer (PST) and fractional revival (FR) as well as to the role played by orthogonal polynomials in the study of these phenomena. Included is a discussion of the ordered Hamming scheme, its relation to multivariate Krawtchouk polynomials of the Tratnik type, the exploration of quantum walks on graphs of this association scheme and their projection to spin lattices with PST and FR.

Keywords Quantum walk · Quantum State transfer · Orthogonal polynomials

Mathematics Subject Classification (2000) Primary 33C45; Secondary 81R30

1 Introduction

The main objective of this lecture is to illustrate the role that orthogonal polynomials play in the analysis of spin network dynamics which is of relevance for quantum

---

The research of ST is supported by JSPS KAKENHI (Grant Numbers 16K13761) and that of LV by a discovery grant of the Natural Sciences and Engineering Research Council (NSERC) of Canada.

H. Miki
Meteorological College, Asahi-Cho, Kashiwa, Japan
e-mail: hmiki@mc-jma.go.jp

S. Tsujimoto
Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Sakyo-Ku, Kyoto, Japan
e-mail: tsujimoto.satoshi.5s@kyoto-u.jp

L. Vinet (✉)
Centre de Recherches Mathématiques, Université de Montréal, Montréal, QC, Canada
e-mail: vinet@crm.umontreal.ca; luc.vinet@umontreal.ca

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_20

information. The transfer of quantum states and the generation of entangled states are two important tasks in this context. Let us first indicate broadly how we shall describe these mathematically in the following.

Take quite generally, a finite set of sites labelled by the integers $n = 0, 1, \ldots, N$. Let $|n)$ be the characteristic vector in $\mathbb{C}^{N+1}$ which has entry 1 at position n and zeros everywhere else. Consider this bra as the vector representing the quantum state of interest at the location n. The evolution operator is the unitary $U(t) = e^{-itH}$ with t the time and H some Hermitian Hamiltonian operator; therefore the state $|n)$ at time $t = 0$ will become $U(t)\,|n)$ at time t. Pick a reference site say 0. We shall wish to have a dynamics such that for some $t = T$,

$$U(T)\,|0) = \alpha\,|0) + \beta\,|N)\,, \quad |\alpha|^2 + |\beta|^2 = 1. \tag{1.1}$$

If this happens, we shall say that we have Fractional Revival (FR) at two sites. Two special cases are of particular interest:

– Perfect State Transfer (PST)
  If $\alpha = 0$ and $\beta = e^{i\phi}$, one has $U(T)\,|0) = e^{i\phi}\,|N)$, that is the state $|0)$ at time $t = 0$ is found at time $t = T$ with probability 1 at the site N; one thus say that it has been perfectly transferred.
– Generation of maximal entanglement
  If the norm of both α and β is equal to $\frac{1}{\sqrt{2}}$, the resulting state at $t = T$ is equivalent to $\frac{1}{\sqrt{2}}\,(|0) + |N))$ which can be pictured as the sum of two vectors: one with a spin up at site 0 and spin down at all the other sites and the other with its only spin up at the site N. This is manifestly a maximally entangled state (that cannot be written as the product of two vectors). By evolving the vector $|0)$ under $U(t)$ we have thus generated maximal entanglement at time T.

These two tasks can be achieved in properly engineered spin chains if one focuses on one-excitation dynamics [2, 3, 16, 25]. PST has in fact been modelled and experimentally realized in photonic waveguide arrays [5, 21].

In the language of Graph Theory, such processes can be viewed as Quantum Walks on weighted paths (one-dimensional graphs) [6, 7, 12, 17]. Quantum walks on other types of graphs have been used in the design of algorithms such as the ones of Grover or Ambianis for example [8, 10]. The interest in PST has prompted the examination of this process on spin networks deployed on higher dimensional graphs with the Hamiltonian taken to be the adjacency matrix (or the Laplacian).

The goals of this lecture are the following:

1. To describe situations where one-excitation dynamics in spin lattices with PST correspond to quantum walks on graphs.
2. To show that spin systems with PST can conversely be identified through this correspondence i.e. from projecting from graphs.
3. To consider for illustration graphs that belong to the Hamming and generalized Hamming schemes.

4. To show that the "go-betweens" are orthogonal polynomials of the Krawtchouk type in one or more variables. As a bonus, this will offer a review of their role in spectral graph theory.

The outline of the lecture is as follows. In the next section, we shall discuss in some details how spin chains can be engineered so as to exhibit FR and/or PST. As shall be seen, this will involve conditions on the spectrum and the reconstruction of the Hamiltonians from these constrained data. Orthogonal polynomials will be seen to play a central role in such inverse spectral problems. The spin chains associated to two families of orthogonal polynomials will be introduced as examples in Sect. 3. This will give the occasion to review some properties of the Krawtchouk polynomials as well as of the novel para-Krawtchouk polynomials. It will be noted that this last family stems naturally from the exploration of fractional revival. In Sect. 4, after an elementary review of the binary Hamming scheme and of its connection to Krawtchouk polynomials, it will be seen that the quantum walk on one graph of the scheme, namely the hypercube, precisely identifies with the 1-excitation dynamics of the spin chain associated to the Krawtchouk polynomials. Section 5 will offer a primer on bivariate Krawtchouk polynomials and their algebraic interpretation as matrix elements of the rotation group $SO(3)$ on the state vectors of the harmonic oscillator in three dimensions. Furthermore, the ordered 2-Hamming scheme will be presented with some of its combinatorics. We shall find out that the adjacency matrices of this scheme are expressed as some bivariate Krawtchouk polynomials (of the Tratnik type) in two elementary matrices. Section 6 will be dedicated to the identification of a (weighted) graph in the ordered 2-Hamming scheme that has both FR and PST and which projects to a spin lattice with these two phenomena. We shall recap our findings in the last section.

2 Fractional Revival (FR) and Perfect State Transfer (PST) in a One Dimensional Spin Chain

We shall here briefly review how the design of spin chains with FR and PST is carried out with the help of orthogonal polynomial theory. Consider the XX chain with the Hamiltonian on $(\mathbb{C}^2)^{N+1}$

$$H = \frac{1}{2}\sum_{l=0}^{N-1} J_{l+1}(\sigma_l^x\sigma_{l+1}^x + \sigma_l^y\sigma_{l+1}^y) + \frac{1}{2}\sum_{l=0}^{N} B_l(\sigma_l^z + 1),$$

where $\sigma_l^x, \sigma_l^y, \sigma_l^z$ are the Pauli matrices acting at site l as follows on the canonical basis $\{|0\rangle, |1\rangle\}$ for $\mathbb{C}^2$:

$$\begin{aligned}
&\sigma^x|1\rangle = |0\rangle, \quad \sigma^y|1\rangle = i|0\rangle, \quad \sigma^z|1\rangle = |1\rangle,\\
&\sigma^x|0\rangle = |1\rangle, \quad \sigma^y|0\rangle = -i|1\rangle, \quad \sigma^z|0\rangle = -|0\rangle.
\end{aligned}$$

It is not difficult to see that

$$\left[H, \sum_{l=0}^{N} \sigma_l^z\right] = 0,$$

in other words that the number of spins that are up is conserved. By convention we shall say that a spin is up if the state is described by the eigenvector of σ^z that has eigenvalue $+1$. We shall use this property to restrict our considerations to the one-excitation subspace, that is to chain states where there is only one spin up. A basis for this subspace will be provided by the following vectors in $\mathbb{C}^{N+1}$:

$$|e_l) = (0, 0, \dots, 0, 1, 0, \dots, 0) \quad l = 0, 1, \dots, N,$$

that have their single 1 entry at the position corresponding to the site where the only spin up is located. In view of the above conservation law, the Hamiltonian preserves the span of these vectors and is seen to act as follows on this basis:

$$H\,|e_l) = J_{l+1}\,|e_{l+1}) + B_l\,|e_l) + J_l\,|e_{l-1})\,, \quad l = 0, 1, \dots, N$$

with $J_0 = J_{N+1} = 0$. In other words, in the occupation basis, the Hamiltonian in the one-excitation subspace takes the form of the following symmetric Jacobi matrix J

$$J = \begin{pmatrix} B_0 & J_1 & & & \\ J_1 & B_1 & J_2 & & \\ & J_2 & \ddots & \ddots & \\ & & \ddots & B_{N-1} & J_N \\ & & & J_N & B_N \end{pmatrix}.$$

Let us now consider the eigenvalue problem

$$H\,|s\rangle = x_s\,|s\rangle\,, \quad s = 0, 1, \dots, N. \tag{2.1}$$

and expand the eigenvectors in terms of the occupation basis vectors:

$$|s\rangle = \sum_{n=0}^{N} \sqrt{w_s}\,\chi_n(x_s)\,|e_n)\,, \quad \chi_0(x_s) = 1.$$

Since $H = J$ on the 1-excitation subspace, it is readily seen that the set $\{\chi_n(x)\}_{n=0}^{N}$ is one of orthogonal polynomials whose orthogonality is given by

$$\sum_{s=0}^{N} \chi_n(x_s)\chi_m(x_s)w_s = \delta_{mn}.$$

From (2.1), one also finds the corresponding three-term recurrence relation

$$x_s \chi_n(x_s) = J_{n+1}\chi_{n+1}(x_s) + B_n \chi_n(x_s) + J_n \chi_{n-1}(x_s).$$

Since the matrix $(\sqrt{w_s}\chi_n(x_s))_{n,s=0}^{N}$ is an orthonormal matrix assuming that the eigenvectors are normalized, we have the inverse expansion

$$|e_n) = \sum_{s=0}^{N} \sqrt{w_s}\chi_n(x_s)\,|s\rangle\,. \tag{2.2}$$

We can now write down how the FR and PST conditions translate on the spectra. From the FR one namely,

$$e^{-iTH}\,|e_0) = \mu\,|e_0) + \nu\,|e_N)\,, \quad |\mu|^2 + |\nu|^2 = 1,$$

one finds with the help of (2.2)

$$e^{-iTx_s} = \mu + \nu\chi_N(x_s) \tag{2.3}$$

since $\chi_0(x) = 1$. Let us begin with the analysis of PST which occurs when $\mu = 0$ and $\nu = e^{i\phi}$. It is straightforward to see that

$$\chi_N(x_s) = e^{-i\phi}e^{-iTx_s}, \quad \phi \in \mathbb{R},$$

which implies that $\chi_N(x_s) = \pm 1$ since $\chi_N(x_s)$ is real. A simple argument using the interlacing properties of zeros of orthogonal polynomials and the positivity of weight function yields

$$\chi_N(x_s) = (-1)^{N+s} \tag{2.4}$$

as a necessary condition for PST. It has also be shown [25] that the condition (2.4) amounts in terms of the coefficients of the recurrence relation as the following mirror-symmetric (or persymmetric) requirement

$$J_{N-n+1} = J_n, \quad B_{N-n} = B_n. \tag{2.5}$$

We then proceed to find spin chains for which both (2.3) and (2.4) are satisfied, which means spin chains where both FR and PST take place. This is readily seen to imply

$$e^{-iTx_s} = e^{i\phi}(\cos\theta + i(-1)^{N+s}\sin\theta). \tag{2.6}$$

It should be remarked that the case $\theta = \frac{\pi}{2}$ corresponds to PST since the amplitude μ is zero.

The problem at this point amounts to find the corresponding parameters J_n and B_n given a spectral set that satisfy the FR requirements (2.6). We can find the solution to this problem by constructing the associated monic orthogonal polynomials

$$P_n(x) = \sqrt{J_1 J_2 \cdots J_n}\,\chi_n(x)$$

whose recurrence coefficients will give J_n and B_n. Briefly, this can be done as follows. From the spectrum data $\{x_s\}_{s=0}^{N}$, we can introduce the characteristic polynomials

$$P_{N+1}(x) = (x - x_0)(x - x_1)\cdots(x - x_N),$$

which is orthogonal to all other $P_n(x) \quad (n = 0, 1, \ldots, N)$. The condition (2.4) provides values for $P_N(x) \propto \chi_N(x)$ at $N+1$ points which fixes $P_N(x)$ by Lagrange interpolation. Once two OPs $P_{N+1}(x)$, $P_N(x)$ are known, all the others are obtained by the recurrence relation. (For more details, the readers may consult [11, 25])

3 Para-Krawtchouk and Krawtchouk Models

Let us see how this algorithm applies in an example. Consider the following spectrum

$$x_s = \beta\left(s + \frac{1}{2}(\delta - 1)(1 - (-1)^s) - \frac{1}{2}(N - 1 + \delta)\right), \quad s = 0, 1, \ldots, N \tag{3.1}$$

which can be viewed as the affine transformation of the superposition of 2 regular lattices of step 2 with spacing δ. It can be checked that this set of spectral points satisfies the FR condition (2.6) with

$$T = \frac{\pi}{\beta}, \quad \theta = (-1)^N \frac{\pi}{2}\delta.$$

In order to have also PST, there must be a $t = T'$ such that

$$e^{-iT'x_s} = e^{i\phi}(-1)^{N+s},$$

which requires $\delta = \frac{q}{p}$ with p, q are coprime integers (and p also odd). In this parametrization,

$$T' = qT.$$

The spectrum (3.1) will therefore correspond to a spin chain with both PST and FR. Using the reconstruction method for the Jacobi matrix that we described briefly at the end of the last section we can obtain the chain specifications through the resulting recurrence coefficients of the associated polynomials. In the present case, when N is odd one finds:

$$J_n = \frac{\beta}{2}\sqrt{\frac{n(N+1-n)((N+1-2n)^2-\delta^2)}{(N-2n)(N-2n+2)}}, \quad B_n = 0.$$

One observes that these couplings are indeed mirror symmetric. Similar expressions are obtained for N even. Remarkably, these explicit recurrence coefficients define orthogonal polynomials that had not really been studied. We have called them para-Krawtchouk polynomials in particular because their orthogonality grid (3.1) has resemblances with the spectrum of the parabosonic oscillator [11, 26]. Quite strikingly they emerge naturally when one looks for fractional revival [1].

From here reorganize a bit the part on Krawtchouk OPs with inclusion of angular momentum connection. If we set $\delta = 1$, the recurrence coefficients become

$$J_n = \frac{\beta\sqrt{n(N+1-n)}}{2}, \quad B_n = 0, \tag{3.2}$$

which are the coefficients of Krawtchouk polynomials given in terms of hypergeometric series:

$$K_n^N(x;p) = {}_2F_1\left(\begin{matrix}-n, -x\\ -N\end{matrix}; \frac{1}{p}\right) = \sum_{k=0}^{N} \frac{(-n)_k(-x)_k}{k!(-N)_k}\left(\frac{1}{p}\right)^k, \quad (0 < p < 1)$$

with $p = \frac{1}{2}$. The corresponding spectrum of the Krawtchouk polynomials are of course

$$x_s = \beta\left(s - \frac{N}{2}\right)$$

and

$$\theta = (-1)^N \frac{\pi}{2}.$$

Therefore, only PST (not FR) can be observed in the spin chain associated with Krawtchouk polynomials.

4 Quantum Walk on the Hypercube

We have seen that the Krawtchouk model (3.2) exhibits PST. It will be instructive to understand how this relates to quantum walks on the hypercube viewed as a graph of the (binary) Hamming scheme and to see how the Krawtchouk polynomials appear in this picture. This will allow us to review basic facts about a standard example of association schemes [4].

4.1 A Brief Review of the Hamming Scheme

Recall that a graph $G = (V; E)$ is defined by the set of vertices V and the set of edges E, which are 2-element subsets of V. Let $|V|$ be a cardinality of V. The adjacency matrix A of G is a $|V| \times |V|$ matrix whose (x, y) element $A_{xy} = \langle x|A|y\rangle$ is given by the number of edges between vertices x and y. Now set $V = \{0, 1\}^N$ which consists of N-tuples of 0 and 1. For these vertices, we can introduce the Hamming distance $d(x, y)$ between $x, y \in V$ which is the number of positions where x and y differ. Using the Hamming distance, we can also introduce the graph G_i $(i = 0, 1, \dots, N)$ whose edges connect all pairs vertices with Hamming distance i. It should be remarked that G_1 is nothing but the N-dimensional hypercube. Let A_i be the adjacency matrix of G_i and $p_{ij}^k (= p_{ji}^k)$ be the intersection numbers which count the number of $z \in V$ such that

$$d(x, z) = i, \quad d(y, z) = j \quad \text{if} \quad d(x, y) = k.$$

The set of matrices $\{A_i\}_{i=0}^N$ is known to satisfy the Bose–Mesner algebra:

$$A_i A_j = \sum_{k=0}^{N} p_{ij}^k A_k,$$

which is a defining condition for an association scheme. The set of graphs $\{G_i\}_{i=0}^N$ belongs to the association scheme known as the binary Hamming scheme $\mathcal{H}(N, 2)$. In this case, we especially have

$$A_1 A_i = (i + 1)A_{i+1} + (N - i + 1)A_{i-1}$$

which implies $A_i = p_i(A_1)$, where $p_i(x)$ is a polynomial of degree i. One can further see that the polynomial $p_i(x)$ is a Krawtchouk polynomial [23]:

$$p_i(\lambda_s) = \binom{N}{i} K_i^N\left(s; \frac{1}{2}\right) \quad s = 0, 1, \dots, N$$

with $\lambda_s = N - 2s$.

4.2 Projection of the Quantum Walk on the Hypercube to the Krawtchouk Model

Let us now explain how quantum walks on the hypercube can be projected to walks on a weighted path that can be identified with the 1-excitation dynamics of the Krawtchouk spin model [6].

Let us consider the N-dimensional hypercube, i.e. the graph G_1 and denote its adjacency matrix by A_1. The unitary operator

$$U(t) = e^{-itA_1} \tag{4.1}$$

defines a (continuous-time) quantum walk on the hypercube. We pick the vertex which corresponds to $(0) \equiv (0, 0, \ldots, 0)$ as reference vertex and organize V as a set of $N + 1$ columns $V_n \quad (n = 0, 1, \ldots, N)$ defined by

$$V_n = \{x \in V \mid d(0, x) = n\}$$

whose cardinality is $|V_n| = k_n = \binom{N}{n}$. We denote the vertices in V_n by $V_{n,m} \quad (m = 1, 2, \ldots, k_n)$. They all have n 1's. It is not difficult to see that each $V_{n,m}$ is connected to the $N - n$ elements of column V_{n+1} obtained by converting a 0 of $V_{n,m}$ to a 1.

To the vertices $x \in V = \{0, 1\}^N$,we shall associate orthonormalized vectors $|x\rangle \in \mathbb{C}^{|V|}$ such that

$$\langle x|y\rangle = \begin{cases} 1 & \text{if} \quad d(x, y) = 0 \\ 0 & \text{otherwise} \end{cases} \qquad (x, y \in V)$$

and introduce the linear span of the following column vectors:

$$|\mathrm{col}\ n\rangle = \frac{1}{\sqrt{k_n}} \sum_{m=1}^{k_n} \left|V_{n,m}\right\rangle.$$

The key observation here is that the evolution (4.1) preserves column space because of distance-regularity, i.e. each vertex in V_n is connected to the same number of vertices in V_{n+1} and vice-versa. In light of this observation, it is possible to project quantum walks on the hypercube to quantum walk along the columns. We can realize this quotient by computing the matrix elements of A_1 between the column vectors. The non-zero elements are obtained as follows:

$$\begin{aligned} \langle \mathrm{col}\ n+1|A_1|\mathrm{col}\ n\rangle &= \frac{1}{\sqrt{k_n k_{n+1}}} \sum_{m'=1}^{k_{n+1}} \sum_{m=1}^{k_n} \left\langle V_{n+1,m'}|A_1|V_{n,m}\right\rangle \\ &= \frac{k_n(N-n)}{\sqrt{k_n k_{n+1}}} = \sqrt{(n+1)(N-n)}. \end{aligned}$$

By symmetry, we also have

$$\langle \text{col}\, n-1|A_1|\text{col}\, n\rangle = \langle \text{col}\, n|A_1|\text{col}\, n-1\rangle = \sqrt{n(N-n+1)}.$$

One thus finds that A_1 has the action

$$A_1\,|\text{col}\, n\rangle = J_{n+1}\,|\text{col}\, n+1\rangle + J_n\,|\text{col}\, n-1\rangle$$

with $J_n = \sqrt{n(N-n+1)}$, which coincides with that of H on $|e_n)$ in the Krawtchouk model (up to a constant factor). It turns out that there is PST on the hypercube but no FR infers the same properties for the Krawtchouk model.

5 Bivariate Krawtchouk Polynomials

In Sect. 3 we have indicated how one-dimensional spin chains with PST and possibly FR could be engineered by identifying the couplings and local magnetic fields along the chain with the recurrence coefficients of suitable orthogonal polynomials. We have presented in some details models associated to the para-Krawtchouk and Krawtchouk polynomials in one variable. This suggests that spin lattices in dimensions higher than one could be constructed with the help of orthogonal polynomials in many variables. We shall focus on two dimensions in the following. With an eye to finding an example of a two-dimensional spin lattice with PST-like properties that extends the simplest system in 1D, we shall review results that concern bivariate Krawtchouk polynomials.

While univariate Krawtchouk polynomials are polynomials orthogonal with respect to binomial distribution function as follows:

$$\sum_{k=0}^{N} \binom{N}{x} p^x (1-p)^{N-x} K_n(x) K_m(x) = h_n \delta_{m,n},$$

bivariate Krawtchouk polynomials are orthogonal with respect to the trinomial distribution function $w(x,y) = \binom{N}{x,y} p^x q^y (1-p-q)^{N-x-y}$:

$$\sum_{0\le x+y\le N} w(x,y) K_{m_1,n_1}(x,y) K_{m_2,n_2}(x,y) = h_{m_1,n_1}\delta_{m_1,m_2}\delta_{n_1,n_2}. \tag{5.1}$$

It should be noted that the orthogonality condition (5.1) does not define bivariate Krawtchouk polynomials uniquely and hence there are several realizations.

Bivariate Krawtchouk polynomials of Tratnik are obtained by the product of univariate Krawtchouk polynomials [24]:

$$T^N_{m,n}(x,y) = \frac{(n-N)_m (x-N)_n}{(-N)_{m+n}} K^{N-n}_m(x;p) K^{N-x}_n\left(k, \frac{q}{1-p}\right). \tag{5.2}$$

Krawtchouk polynomials of Griffiths, originally introduced by Griffiths [9, 13] and rediscovered by Hoare and Rahman [15], are usually defined in terms of the special case of Aomoto-Gel'fand hypergeometric series [14]:

$$G^N_{m,n}(x, y) = \sum_{0\le i+j+k+l\le N} \frac{(-m)_{i+j}(-n)_{k+l}(-x)_{i+k}(-y)_{j+l}}{i!j!k!l!(-N)_{i+j+k+l}} u_1^i v_1^j u_2^k v_2^l \tag{5.3}$$

with

$$pu_i + qv_i = 1, \quad i = 1, 2,$$
$$pu_1u_2 + qv_1v_2 = 1$$

and

$$\tilde{p}u_1 + \tilde{q}u_2 = 1,$$
$$\tilde{p}v_1 + \tilde{q}v_2 = 1,$$
$$\tilde{p}u_1v_1 + \tilde{q}v_1v_2 = 1.$$

It is obvious from the expression (5.3) that the polynomials $\{G^N_{m,n}(x, y)\}$ have a duality relation under the exchange of the variables x, y and the degree m, n. They are hence orthogonal with respect to the variables x, y as follows:

$$\sum_{0\le m+n\le N} \tilde{w}(m, n)G^N_{m,n}(x_1, y_1)G^N_{m,n}(x_2, y_2) = \tilde{h}_{x_1,y_1}\delta_{x_1,x_2}\delta_{y_1,y_2}$$

with $\tilde{w}(m, n) = \binom{N}{m,n}\bar{p}^m\bar{q}^n(1 - p - q)^{N-m-n}$. It should be remarked here that the series (5.3) becomes (5.2) if we set

$$u_1 = \frac{1}{p}, \quad v_1 = 0, \quad u_2 = 1, \quad v_2 = \frac{1-p}{q}. \tag{5.4}$$

In other words, the Krawtchouk polynomials of Griffiths contains those of Tratnik as a special case.

5.1 Algebraic Interpretation: $SO(3)$

A group-theoretic interpretation of the multivariate Krawtchouk polynomials allows for a cogent derivation of many of their properties. Especially, the Krawtchouk polynomials of Griffiths in d variables can be interpreted as matrix elements of $SO(d + 1)$ unitary representations [11]. We shall give a brief review of this in the case $d = 2$.

Let $a_i, a_i^+ (i = 1, 2, 3)$ be operators of 3 independent oscillators with the action

$$a_i\,|n_i\rangle = \sqrt{n_i}\,|n_i - 1\rangle\,, \quad a_i^+\,|n_i\rangle = \sqrt{n_i + 1}\,|n_i + 1\rangle\,, \quad a_i\,|0\rangle = 0$$

and $|n_1, n_2, n_3\rangle$ be oscillator states defined by

$$|n_1, n_2, n_3\rangle = |n_1\rangle \otimes |n_2\rangle \otimes |n_3\rangle\,.$$

We fix $n_1 + n_2 + n_3 = N$ and write

$$|m, n\rangle_N = |m, n, N - m - n\rangle\,, \quad 0 \le m + n \le N. \tag{5.5}$$

Since the three-dimensional harmonic oscillator Hamiltonian $H = \sum_{i=1}^{3} a_i a_i^+$ is invariant under $SU(3)$ and a fortiori under its $SO(3)$ subgroup, the eigensubspace of energy N spanned by the (orthonormal) basis vectors $|m, n\rangle_N$ forms a representation space for these groups. Let $R \in SO(3)$ and define its unitary representation $U = U(R)$ by

$$U(R) a_i U^+(R) = \sum_{k=1}^{3} R_{ki} a_k. \tag{5.6}$$

The matrix elements of this unitary operator in the basis (5.5) can be cast in the form

$${}_N\langle i, k|U(R)|m, n\rangle_N = w_{i,k;N}\, P^N_{m,n}(i, k)$$

with $P^N_{0,0}(i, k) = 1$ and $w_{i,k;N} = {}_N\langle i, k|U(R)|0, 0\rangle_N$. From the unitarity of U:

$${}_N\langle m', n'|U^+ U|m, n\rangle_N = \sum_{0\le i+k\le N} {}_N\langle m', n'|U^+|i, k\rangle_N\ {}_N\langle i, k|U|m, n\rangle_N = \delta_{m,m'}\delta_{n,n'},$$

it is straightforward to see that $\{P_{m,n}\}_{m,n}$ have the orthogonality relation

$$\sum_{0\le i+k\le N} w^2_{i,k;N}\, P^N_{m,n}(i, k) P^N_{m',n'}(i, k) = \delta_{m,m'}\delta_{n,n'}.$$

The weight function $w^2_{i,k;N}$ can be computed directly as follows. From (5.6) and ${}_{N-1}\langle i, k|U(R)a_1|0, 0\rangle_N = 0$, we see that

$$\begin{aligned}
&{}_{N-1}\langle i, k|U a_1|0, 0\rangle_N \\
&= {}_{N-1}\langle i, k|U a_1 U^+ U a_1|0, 0\rangle_N \\
&= R_{11}\sqrt{i + 1}\ {}_N\langle i + 1, k|U|0, 0\rangle_N + R_{21}\sqrt{k + 1}\ {}_N\langle i, k + 1|U|0, 0\rangle_N \\
&+ R_{31}\sqrt{N - i - k}\ {}_N\langle i, k|U|0, 0\rangle_N\,,
\end{aligned}$$

which results in

$$R_{11}\sqrt{i+1}w_{i+1,k;N} + R_{21}\sqrt{k+1}w_{i,k+1;N} + R_{31}\sqrt{N-i-k}w_{i,k;N} = 0.$$

Similarly using ${}_{N-1}\langle i,k|U(R)a_2|0,0\rangle_N = 0$, one finds

$$R_{12}\sqrt{i+1}w_{i+1,k;N} + R_{22}\sqrt{k+1}w_{i,k+1;N} + R_{32}\sqrt{N-i-k}w_{i,k;N} = 0.$$

It is not difficult to verify that

$$w_{i,k;N} = C\frac{R_{13}^i R_{23}^k R_{33}^{N-i-k}}{\sqrt{i!k!(N-i-k)!}}$$

is a solution to the above difference systems. The constant term C is determined to be $C = \sqrt{N!}$ from the relation

$$\begin{aligned}1 &= {}_N\big\langle 0,0|U^+U|0,0\big\rangle_N \\ &= \sum_{0\le i+k\le N} {}_N\langle 0,0|U|i,k\rangle_N \; {}_N\langle i,k|U|0,0\rangle_N = \sum_{0\le i+k\le N} w_{i,k;N}^2.\end{aligned}$$

As a result, we have

$$w_{i,k;N} = R_{13}^i R_{23}^k R_{33}^{N-i-k}\sqrt{\binom{N}{i,k}}$$

and $P_{m,n}^N(i,k)$ are thus orthogonal with respect to trinomial distribution function $\binom{N}{i,k}p^i q^k(1-p-q)^{N-i-k}$ with

$$p = R_{13}^2, \quad q = R_{23}^2. \tag{5.7}$$

We can conclude that $\{P_{m,n}^N(i,k)\}_{0\le m+n\le N}$ are (orthonormal) bivariate Krawtchouk polynomials.

The group theoretical interpretation enables us to derive several properties of $P_{m,n}^N(i,k)$. For instance, the relations

$${}_N\big\langle i,k|a_1^+a_1U|m,n\big\rangle_N = i\;{}_N\langle i,k|U|m,n\rangle_N = \sum_{r,s=1}^{3} R_{r1}R_{s1}\;{}_N\big\langle i,k|Ua_r^+a_s|m,n\big\rangle_N\,,$$

$${}_N\big\langle i,k|a_2^+a_2U|m,n\big\rangle_N = i\;{}_N\langle i,k|U|m,n\rangle_N = \sum_{r,s=1}^{3} R_{r2}R_{s2}\;{}_N\big\langle i,k|Ua_r^+a_s|m,n\big\rangle_N$$

yield the following two 7-term recurrence relations

$$
\begin{aligned}
&i P_{m,n}^{N}(i,k) = [R_{11}^2 m + R_{12}^2 n + R_{13}^2 (N-m-n)] P_{m,n}^{N}(i,k) \\
&+ R_{11}R_{12}[\sqrt{m(n+1)}P_{m-1,n+1}^{N}(i,k) + \sqrt{n(m+1)}P_{m+1,n-1}^{N}(i,k)] \\
&+ R_{11}R_{13}[\sqrt{m(N-m-n+1)}P_{m-1,n}^{N}(i,k) + \sqrt{(m+1)(N-m-n)}P_{m+1,n}^{N}(i,k)] \\
&+ R_{12}R_{13}[\sqrt{n(N-m-n+1)}P_{m,n-1}^{N}(i,k) + \sqrt{(n+1)(N-m-n)}P_{m,n+1}^{N}(i,k)]
\end{aligned}
\tag{5.8}
$$

and

$$
\begin{aligned}
&k P_{m,n}^{N}(i,k) = [R_{21}^2 m + R_{22}^2 n + R_{23}^2 (N-m-n)] P_{m,n}^{N}(i,k) \\
&+ R_{21}R_{22}[\sqrt{m(n+1)}P_{m-1,n+1}^{N}(i,k) + \sqrt{n(m+1)}P_{m+1,n-1}^{N}(i,k)] \\
&+ R_{21}R_{23}[\sqrt{m(N-m-n+1)}P_{m-1,n}^{N}(i,k) + \sqrt{(m+1)(N-m-n)}P_{m+1,n}^{N}(i,k)] \\
&+ R_{22}R_{23}[\sqrt{n(N-m-n+1)}P_{m,n-1}^{N}(i,k) + \sqrt{(n+1)(N-m-n)}P_{m,n+1}^{N}(i,k)].
\end{aligned}
\tag{5.9}
$$

Furthermore, we can find that $P_{m,n}^{N}(i,k)$ have the explicit expression as the Krawtchouk polynomials of Griffiths (5.3):

$$
P_{m,n}^{N}(i,k) = \sqrt{\binom{N}{m,n}} \left(\frac{R_{31}}{R_{33}}\right)^m \left(\frac{R_{32}}{R_{33}}\right)^n G_{m,n}^{N}(i,k)
$$

with

$$
\begin{aligned}
u_1 &= 1 - \frac{R_{11}R_{33}}{R_{13}R_{31}}, \quad v_1 = 1 - \frac{R_{21}R_{33}}{R_{23}R_{31}}, \\
u_2 &= 1 - \frac{R_{12}R_{33}}{R_{13}R_{32}}, \quad v_2 = 1 - \frac{R_{22}R_{33}}{R_{23}R_{32}}.
\end{aligned}
\tag{5.10}
$$

(For the details of the derivation, see [11].) The Tratnik polynomials (of Krawtchouk type) are specialization of the Griffiths corresponding to particular rotations given by the product of two rotation matrices about two orthogonal axes:

$$
\begin{aligned}
R = R_{yz}R_{xz} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix} \\
&= \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ \sin\theta\sin\phi & \cos\theta & \sin\theta\cos\phi \\ \cos\theta\sin\phi & -\sin\theta & \cos\theta\cos\phi \end{pmatrix},
\end{aligned}
$$

from which $R_{12} = 0$ and the parametrization (5.7) and (5.10) coincides with (5.4). The precise identification with the bivariate polynomials of Krawtchouk type defined in (5.2) involves a normalization factor given by:

$$P_{i,j}^{N}(x,y) = \sqrt{\binom{N}{i,\,j}\tilde{p}^{i}\tilde{q}^{j}(1-p-q)^{-i-j}}\,T_{i,j}^{N}(x,y) \tag{5.11}$$

with

$$\tilde{p} = \frac{p(1-p-q)}{1-p}, \quad \tilde{q} = \frac{q}{1-p}.$$

For later usage, we write down the recurrence relations for $T_{i,j}^{N}(x,y)$:

$$\begin{aligned}
xT_{i,j}^{N}(x,y) &= -p(N-i-j)[T_{i+1,j}^{N}(x,y) - T_{i,j}^{N}(x,y)] \\
&\quad - (1-p)i[T_{i-1,j}^{N}(x,y) - T_{i,j}^{N}(x,y)], \\
yT_{i,j}^{N}(x,y) &= \frac{pq}{1-p}(N-i-j)[T_{i+1,j}^{N}(x,y) - T_{i,j}^{N}(x,y)] \\
&\quad - \frac{q}{1-p}(N-i-j)[T_{i,j+1}^{N}(x,y) - T_{i,j}^{N}(x,y)] \\
&\quad + qi[T_{i-1,j}^{N}(x,y) - T_{i,j}^{N}(x,y)] \\
&\quad - \frac{p(1-p-q)}{1-p}j[T_{i+1,j-1}^{N}(x,y) - T_{i,j}^{N}(x,y)] \\
&\quad - \frac{q}{1-p}i[T_{i-1,j+1}^{N}(x,y) - T_{i,j}^{N}(x,y)].
\end{aligned} \tag{5.12}$$

For later reference, combining the relations (5.12) and keeping (5.11) in mind, in the case $p = \frac{1}{2}$ and $q = \frac{1}{4}$ we find the following relation for the "Hermitian" Tratnik polynomials $P_{i,j}^{N}(x,y)$:

$$\begin{aligned}
&[\alpha(N-2x) + \beta(2N-2x-4y)]P_{i,j}^{N}(x,y) \\
&= \alpha j P_{i,j}^{N}(x,y) + \alpha\sqrt{(i+1)(N-i-j)}P_{i+1,j}^{N}(x,y) \\
&+ \beta\sqrt{2(j+1)(N-i-j)}P_{i,j+1}^{N}(x,y) + \alpha\sqrt{i(N+1-i-j)}P_{i-1,j}^{N}(x,y) \\
&+ \beta\sqrt{2j(N+1-i-j)}P_{i,j-1}^{N}(x,y) + \beta\sqrt{2i(j+1)}P_{i-1,j+1}^{N}(x,y) \\
&+ \beta\sqrt{2(i+1)j}P_{i+1,j-1}^{N}(x,y).
\end{aligned} \tag{5.13}$$

The algebraic interpretation also allows us to obtain the generating function formula for $T^N_{i,j}(x, y)$:

$$\sum_{0\le x+y\le N} \binom{N}{x, y} T^N_{i,j}(x, y) s^x t^y \\ = (1+s+t)^{N-i-j}\left(1+\frac{p-1}{p}s+t\right)^i\left(1+\frac{p+q-1}{q}t\right)^j. \tag{5.14}$$

5.2 Relationship to Generalized Hamming Scheme

In Sect. 4.1, we have seen that univariate Krawtchouk polynomials naturally arise in the binary Hamming scheme $\mathcal{H}(N, 2)$. We shall here introduce the genralization of the Hamming scheme which is usually called the ordered 2-Hamming scheme [18] and show that this scheme brings on the bivariate Krawtchouk polynomials of Tratnik.

Let $Q = \{0, 1\}$ and consider the set $Q^{(N,2)}$ of vectors of dimension $2N$ over Q. The vector $x \in Q^{(N,2)}$ will be presented by 2-binary sequences of length N:

$$x = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N), \quad \bar{x}_j = (x_{j1}, x_{j2}) \in Q^2.$$

For $x \in Q^{(N,2)}$, we can introduce the shape $e(x)$ by

$$\begin{aligned} &e(x) = (e_1, e_2),\\ &e_1 = \#\{j \in \{1, 2, \dots, N\} \mid \bar{x}_j = (1, 0)\},\\ &e_2 = \#\{j \in \{1, 2, \dots, N\} \mid \bar{x}_j = (0, 1), (1, 1)\}. \end{aligned}$$

For example, $x = ((0, 0), (1, 0), (1, 1), (0, 1), (0, 1)) \in Q^{(5,2)}$ and $e(x) = (1, 3)$. We denote the set of the all shapes by

$$E = \{(e_1, e_2) \in (\mathbb{Z}_{\ge 0})^2 \mid 0 \le e_1 + e_2 \le N\}.$$

We can now use shapes to establish relations between vertices. We shall say that two vertices $x, y \in Q^{N,2}$ are related under shape e:

$$x \sim_e y \quad \text{if} \quad e((x - y \mod 2)) = e.$$

For example, let $x = ((0, 0), (1, 0), (0, 0))$ and $y = ((1, 1), (0, 1), (1, 0))$. Then $(x - y \mod 2) = ((1, 1), (1, 1), (1, 0))$ and $e((x - y \mod 2)) = (1, 2)$, we thus have $x \sim_{(1,2)} y$. We can then introduce the graph G_e associated with the shape e as the one where all two vertices (v_x, v_y) in $\{v_x \mid x \in Q^{(N,2)}\}$ are linked if $v_x \sim_e v_y$.

The adjacency matrix A_e of the graph G_e is given by

$$\langle x|A_e|y\rangle = \begin{cases} 1 & (x \sim_e y) \\ 0 & (\text{otherwise}) \end{cases} .$$

The set of the adjacency matrices $\mathcal{A} = \{A_e \mid e \in E\}$ satisfy the Bose–Mesner algebra

$$A_{(i,j)}A_{(k,l)} = \sum_{0\le m+n\le N} \alpha^{(m,n)}_{(i,j),(k,l)} A_{(m,n)}$$

and thus defines an association scheme called the ordered 2-Hamming scheme. The intersection numbers are here defined by

$$\alpha^{(m,n)}_{(i,j),(k,l)} = \#\{z \in Q^{(N,2)} \mid x \sim_{(i,j)} z, y \sim_{(k,l)} z, x \sim_{(m,n)} y\}.$$

and it is found in particular that

$$\begin{aligned} A_{(1,0)}A_{(i,j)} &= (N+1-i-j)A_{(i-1,j)} + jA_{(i,j)} + (i+1)A_{(i+1,j)}, \\ A_{(0,1)}A_{(i,j)} &= 2(N+1-i-j)A_{(i,j-1)} + 2(i+1)A_{(i+1,j-1)} \\ &\quad + (j+1)A_{(i-1,j+1)} + (j+1)A_{(i,j+1)}. \end{aligned} \tag{5.15}$$

The derivation of these relations is not difficult. For instance, we can compute $\alpha^{(i-1,j)}_{(1,0),(i,j)}$ as follows. Let us count the number of $z \in Q^{(N,2)}$ such that

$$e((x - z \mod 2)) = (1,0), \quad e((y - z \mod 2)) = (i, j)$$

if $e((x - y \mod 2)) = ((i-1, j))$. Take

$$x = ((1,0), \dots, (1,0), (0,1), \dots, (0,1), (0,0), \dots, (0,0))$$

with $i-1$ $(1,0)$s, j $(0,1)$s and $N+1-i-j$ $(0,0)$s and $y = ((0,0), \dots, (0,0))$. In order to get a z, a $(0,0)$ should be converted into a $(1,0)$ and hence there are $N+1-i-j$ ways to do that. We then conclude that

$$\alpha^{(i-1,j)}_{(1,0),(i,j)} = N+1-i-j.$$

The other intersection numbers can be computed in the same manner. Quite interestingly, the relations (5.15) coincide with the recurrence relations (5.12) for the bivariate Krawtchouk polynomials of Tratnik with $p = \frac{1}{2}, q = \frac{1}{4}$. Therefore, the bivariate Krawtchouk polynomials of Tratnik arise in the ordered 2-Hamming scheme as univariate Krawtchouk polynomials do in the binary Hamming scheme $\mathcal{H}(N, 2)$.

6 FR and PST in Two-Dimensional Spin Lattices

We have seen in Sect. 4.2 that the quantum walk on the hypercube the Hamming scheme leads to a 1-dimensional spin model with PST. We shall here consider the graph $G_{\alpha,\beta}$ in the ordered 2-Hamming scheme corresponding to the weighted adjacency matrix:

$$A_{\alpha,\beta} = \alpha A_{(1,0)} + \beta A_{(0,1)}$$

and consider quantum walks this matrix generates. Thanks to the distance regularity of the graph, we can project the quantum walk on the graph $G_{\alpha,\beta}$ to walks on a two-dimensional regular lattice of triangular shape in the same fashion as in Sect. 4.2.

Let $(0) \equiv ((0,0), (0,0), \ldots, (0,0))$ be a reference vertex and organize V as the set of $\binom{N+1}{2}$ column $V_{i,j}$ defined by

$$V_{ij} = \{x \in V \mid e(x) = (i,j)\} \quad 0 \le i + j \le N.$$

The cardinality of $V_{i,j}$ is given by

$$k_{i,j} = |V_{i,j}| = \binom{N}{i,\, j} 2^j.$$

Let us label the vertices in column $V_{i,j}$ by $V_{(i,j),k}, \quad k = 1, 2, \ldots, k_{i,j}$. Under the relation corresponding to shape $(1,0)$, each $V_{(i,j),k}$ in $V_{i,j}$ is connected to $N-i-j$ vertices in $V_{i+1,j}$ since exchanging a $(0,0)$ for a $(1,0)$ in $V_{(i,j),k}$ gives a vertex in $V_{i+1,j}$ and there are $N-i-j$ ways. It is not difficult to see similarly that $V_{(i,j),k}$ is connected to j vertices in $V_{i,j}$. For relation corresponding to $(0,1)$, we can also see that each $V_{(i,j),k}$in $V_{i,j}$ is connected to $2(N-i-j)$ vertices in $V_{i,j+1}$ and j vertices in $V_{i+1,j-1}$.

Consider now the column space taken to be the linear span of the column vectors given by

$$|\text{col } i, j\rangle = \frac{1}{\sqrt{k_{i,j}}} \sum_{k=1}^{k_{i,j}} \left|V_{(i,j),k}\right\rangle.$$

Distance regularity assures that $A_{(1,0)}$ and $A_{(0,1)}$ preserve column space and allows to project from the quantum walks on $G_{\alpha,\beta}$ to the simplex labelling the columns. To that end, we compute the matrix elements of $A_{(1,0)}$ and $A_{(0,1)}$ between the column vectors in the same manner as for the hypercube. One finds:

$$\left\langle \text{col } i+1, j \middle| A_{(1,0)} \middle| \text{col } i, j \right\rangle = \sqrt{(i+1)(N-i-j)},$$
$$\left\langle \text{col } i, j \middle| A_{(1,0)} \middle| \text{col } i, j \right\rangle = j,$$

$$\langle \text{col } i, j+1 | A_{(1,0)} | \text{col } i, j \rangle = \sqrt{2(j+1)(N-i-j)},$$
$$\langle \text{col } i+1, j-1 | A_{(1,0)} | \text{col } i, j \rangle = \sqrt{2(i+1)j}.$$

One thus observes that quantum walks on $G_{\alpha,\beta}$ are in correspondence with 1-excitation dynamics of the spin lattice of triangular shape with Hamiltonian [20]

$$\begin{aligned} H = \sum_{0\le i+j\le N} & \alpha\sqrt{(i+1)(N-i-j)}\frac{\sigma^x_{i,j}\sigma^x_{i+1,j}+\sigma^y_{i,j}\sigma^y_{i+1,j}}{2} \\ & + \beta\sqrt{2(j+1)(N-i-j)}\frac{\sigma^x_{i,j}\sigma^x_{i,j+1}+\sigma^y_{i,j}\sigma^y_{i,j+1}}{2} \\ & + \beta\sqrt{2(i+1)j}\frac{\sigma^x_{i,j}\sigma^x_{i+1,j-1}+\sigma^y_{i,j}\sigma^y_{i-1,j+1}}{2} + \alpha j\frac{1+\sigma^z_{i,j}}{2}. \end{aligned} \tag{6.1}$$

Indeed, on the subspace spanned by the 1-excitation basis vectors $\left|e_{i,j}\right) = E_{i,j}$ with $E_{i,j}$ the $(N+1)\times(N+1)$ matrix with 1 in the (i, j) entry and zeros everywhere else, we see that

$$\begin{aligned} H\left|e_{i,j}\right) = & \alpha\sqrt{(i+1)(N-i-j)}\left|e_{i+1,j}\right) + \beta\sqrt{2(j+1)(N-i-j)}\left|e_{i,j+1}\right) \\ & + \alpha\sqrt{i(N+1-i-j)}\left|e_{i-1,j}\right) + \beta\sqrt{2j(N+1-i-j)}\left|e_{i,j-1}\right) \\ & + \beta\sqrt{2(i+1)j}\left|e_{i+1,j-1}\right) + \beta\sqrt{2i(j+1)}\left|e_{i-1,j+1}\right) + \alpha j\left|e_{i,j}\right), \end{aligned} \tag{6.2}$$

which corresponds to $[\alpha A_{(1,0)} + \beta A_{(0,1)}]\,|\text{col } i, j\rangle$. as the coefficients of the relation (6.2) coincide with those in (5.13). On the subspace spanned by the 1-excitation basis, H can hence be diagonalized by the Hermitian Tratnik polynomials (5.11) with $p = \frac{1}{2}$ and $q = \frac{1}{4}$ and the corresponding eigenvalues and eigenvectors are given by

$$\lambda_{x,y} = \alpha(N-2x) + \beta(2N-2x-4y),$$
$$|x, y\rangle = \sum_{0\le i+j\le N} w_{x,y;N} P^N_{i,j}(x, y)\left|e_{i,j}\right), \quad 0 \le x+y \le N$$

respectively.

Let us now consider the evolution under the dynamics of this spin lattice of a single qubit located at the apex $(0, 0)$. The amplitude for finding that qubit at the site (k, l) at time t is given by:

$$f_{(k,l)}(t) = \left(e_{k,l}|e^{-itH}|e_{0,0}\right).$$

This transition amplitude can be computed with the help of the generating function formula (5.14) as follows;

$$\begin{aligned}
f_{(k,l)}(t) &= \left(e_{k,l}|e^{-itH}|e_{0,0}\right) \\
&= \sum_{0\le x+y\le N} \left(e_{k,l}|x,y\right) e^{-it\lambda_{x,y}} \left(x,y|e_{0,0}\right) \\
&= \sum_{0\le x+y\le N} \binom{N}{x,y} \left(\frac{1}{2}\right)^x \left(\frac{1}{4}\right)^y \left(\frac{1}{4}\right)^{N-x-y} P^N_{0,0}(x,y) P^N_{k,l}(x,y) e^{-it\lambda_{x,y}} \\
&= e^{-iN(\alpha+2\beta)t} \frac{\sqrt{2^l}}{4^N} \sqrt{\binom{N}{k,l}} (1+2z_1+z_2)^{N-k-l}(1-2z_1+z_2)^k(1-z_2)^l
\end{aligned}$$

with $z_1 = e^{2i(\alpha+\beta)t}$ and $z_2 = e^{4i\beta t}$. In order to achieve transfer only to site (i, j) with $i + j = N$, we must have for some $t = T$

$$1 + 2z_1 + z_2 = 0.$$

Since $|z_1| = |z_2| = 1$, the relation implies that

$$z_2 = 1, \quad z_1 = -1. \tag{6.3}$$

From last expression for the amplitude $f_{k,l}(t)$, we see that $z_2 = 1$ requires that $j = 0$ at $t = T$ and hence

$$|f_{(i,j)}(T)| = \begin{cases} 1 & (i,j) = (N,0) \\ 0 & \text{(otherwise)} \end{cases} .$$

This is nothing but PST between (0, 0) and $(N, 0)$. The condition (6.3) for PST can be realized in different ways. One simple instance is that

$$\alpha = 1, \quad \beta = 2, \quad T = \frac{\pi}{2}.$$

It should be remarked here that in the case $\alpha = 1$ and $\beta = 2$, we see that at $t = \frac{\pi}{4}$

$$z_2 = 1, \quad z_1 \neq -1,$$

which implies

$$\left|f_{(i,j)}\left(\frac{\pi}{4}\right)\right| = 0, \quad j \neq 0.$$

In other words, at $t = \frac{\pi}{4}$ FR occurs on the sites $(i, 0)$, $i = 0, 1, \ldots, N$. This can be depicted on Fig. 1 where $N = 7$.

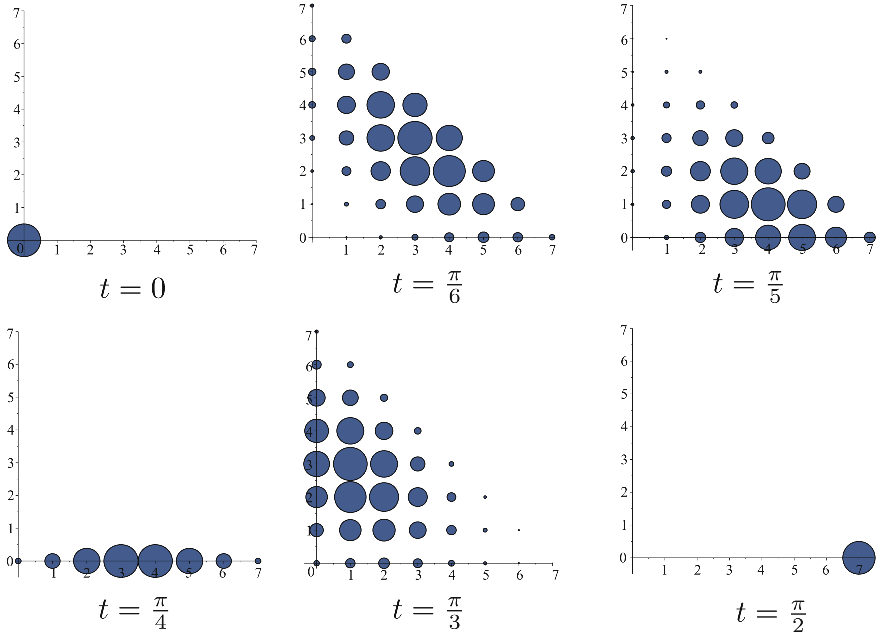


Fig. 1 The transition amplitude $|f_{i,j}(t)|$ for $A_{(1,0)} + 2A_{(0,1)}$ when $N = 7$. The areas of the circles are proportional to $|f_{(i,j)}(t)|$ at the given lattice point (i, j). PST occurs at $\frac{\pi}{2}$ and FR on the set of sites $i = 0, 1, \cdots, N$ and $j = 0$ occurs at $t = \frac{\pi}{4}$

7 Concluding Remarks

Summing up, we have seen how PST and FR can occur on weighted paths associated to spin chains with non-uniform couplings prescribed by orthogonal polynomials. We have observed in particular that the Krawtchouk spin chain model is related with quantum walks on the hypercube. This connection is underscored by the occurrence of Krawtchouk polynomials as eigenfunctions and as matrix eigenvalues of the Hamming scheme. We have reviewed the algebraic interpretation of a two-variable generalization of the Krawtchouk polynomials and discussed the ordered 2-Hamming scheme which features the bivariate Krawthouk polynomials of Tratnik. Such connection enables us to introduce a new two-dimensional spin chain model where PST and FR can be found. It should be stressed here that the spin chain model associated with the more general bivariate Krawtchouk polynomials, i.e. of Griffiths, was considered [19, 22] although PST was not observed in these models. These observations give illustrations of the role that the theory of orthogonal polynomials can play in the analysis of quantum information tasks.

References

1. L. Banchi, E. Compagno, S. Bose, Phys. Rev. A **91**, 052323 (2015)
2. S. Bose, Contemp. Phys. **48**, 13–30 (2007)
3. É.O. Bosse, L. Vinet, SIGMA **13**, 074 (2017)
4. A.E. Brouwer, A.M. Cohen, A. Neumaier, *Distance-Regular Graphs* (Springer, 1989)
5. R.J. Chapman, M. Santandrea, Z. Huang, G. Corrielli, A. Crespi, M. Yung, R. Osellame, A. Peruzz, Nat. Commun. **7**, 11339 (2016)
6. M. Christandl, N. Datta, T.C. Dorlas, A. Ekert, A. Kay, A.J. Landahl, Phys. Rev. A **71**, 032312 (2005)
7. A. Childs, E. Farhi, S. Gutmann, Quant. Inf. Process. **1**, 35–43 (2002)
8. A. Childs, J. Goldstone, Phys. Rev. A **70**, 022314 (2004)
9. P. Diaconis, R. Griffiths, J. Stat. Plan. Inf. **154**, 39–53 (2014)
10. E. Farhi, S. Gutmann, Phys. Rev. A **58**, 915 (1998)
11. V. Genest, L. Vinet, A. Zhedanov, Ann. Phys. **371**, 348–367 (2016)
12. C. Godsil, Disc. Math. **312**, 123–147 (2012)
13. R. Griffiths, Austral. J. Statist. **13**, 27–35 (1971)
14. F.A. Grünbaum, M. Rahman, SIGMA **6**, 090 (2010)
15. M.R. Hoare, M. Rahman, SIGMA **4**, 089 (2008)
16. A. Kay, Perfect, Int. J. Quant. Inf. **8**, 641–676 (2010)
17. V. Kendon, C. Tamon, J. Comput. Theor. Nanosci. **8**, 422–433 (2011)
18. W.J. Martin, D.R. Stinson, Can. J. Math. **51**, 326–346 (1999)
19. H. Miki, S. Tsujimoto, L. Vinet, A. Zhedanov, Phys. Rev. A **85**, 062306 (2012)
20. H. Miki, S. Tsujimoto, L. Vinet, Scipost Phys. **7**, 001 (2019)
21. A. Perez-Leija, R. Keil, A. Kay, H. Moya-Cessa, S. Nolte, L. Kwek, B.M. Rodríguez-Lara, A. Szameit, D.N. Christodoulides, Phys. Rev. A **87**, 012309 (2013)
22. S. Post, Acta Appl. Math. **135**, 209–224 (2014)
23. D. Stanton, in *Special Functions 2000: Current Perspective and Future Directions*, ed. by J. Bustoz, M.E.H. Ismail, S.K. Suslov (Kluwer, Dorchester, 2001), pp. 389–410
24. M.V. Tratnik, J. Math. Phys. **32**, 2337–2342 (1991)
25. L. Vinet, A. Zhedanov, Phys. Rev. A **85**, 012323 (2012)
26. L. Vinet, A. Zhedanov, J. Phys. A Math. Theor. **45**, 265304 (2012)

An Introduction to Special Functions with Some Applications to Quantum Mechanics

Sergei K. Suslov, José M. Vega-Guzmán, and Kamal Barley

Dedicated to the memory of Professor Arnold F. Nikiforov and Professor Vasilii B. Uvarov

Abstract A short review on special functions and solution of the Coulomb problems in quantum mechanics is given. Multiparameter wave functions of linear harmonic oscillator, which cannot be obtained by the standard separation of variables, are discussed. Expectation values in relativistic Coulomb problems are studied by computer algebra methods.

Keywords Nonrelativistic and relativistic Coulomb problems · Schrödinger equation · Dirac equation · Laguerre polynomials · Spherical harmonics · Bessel functions · Hahn polynomials · Chebyshev polynomials of a discrete variable · Generalized hypergeometric series

Mathematics Subject Classification (2000) Primary 33A65, 81C05; Secondary 81C40

In this survey, we aim to provide a concise introduction to the theory of special functions of hypergeometric type and discuss some of their immaculate applications to problems of nonrelativistic and relativistic quantum mechanics. We work in

S. K. Suslov (✉)
Arizona State University, Tempe, AZ, USA
e-mail: sks@asu.edu; sergei@asu.edu

J. M. Vega-Guzmán
Lamar University, Beaumont, TX, USA
e-mail: jvegaguzman@lamar.edu

K. Barley
Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA
e-mail: kamal.barley@stonybrook.edu

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_21

the Nikiforov-Uvarov paradigm [138] when a general transformation of a certain differential equation, together with an integral representation, have been used for developing the theory of special functions of hypergeometric type. On the contrary, we utilize generalized power series expansions which is more traditional in solutions of applied problems. Basic facts about special functions are useful in the study of theoretical and mathematical physics, applied mathematics and modeling. This is why we hope our self-contained presentation of those topics in the state of the art theory of special functions will be very useful for general readers.

In the second part, in addition to traditional topics, a detailed discussion of a "missing" solution of simple harmonic oscillator and relativistic Coulomb integrals are given by computer algebra methods. This will be useful for teaching elementary quantum mechanics and quantum field theory. We are grateful to all our students.

1 An Introduction to Special Functions

The main objective in this section is to present in a compact form main facts about the classical special functions of hypergeometric type, i.e., classical orthogonal polynomials (Jacobi, Laguerre and Hermite) and functions of the second kind, hypergeometric functions, confluent hypergeometric functions and Bessel functions, on the base on the second order differential equation they satisfy. That forms a platform for the study of difference hypergeometric functions, q-orthogonal polynomials, q-beta integrals and biorthogonal rational functions (see [7–9, 12, 15, 16, 79, 103, 139, 147, 174, 195] and references therein).

1.1 Classical Hypergeometric Functions

Classical orthogonal polynomials, hypergeometric functions and Bessel functions are particular solutions of the differential equation

$$\sigma(x)y'' + \tau(x)y' + \lambda y = 0, \tag{1.1}$$

where $\sigma(x)$ and $\tau(x)$ are polynomials of respective degrees at most two and one, and λ is a constant. Equation (1.1) can also be rewritten in the self-adjoint form

$$\left(\sigma\rho y'\right)' + \lambda\rho y = 0, \qquad (\sigma\rho)' = \tau\rho. \tag{1.2}$$

We shall refer to (1.1) as an *equation of hypergeometric type*, and its solutions as *functions of hypergeometric type*. Generally speaking, these functions can be studied in a domain of the complex plane. In this case, we shall usually use the complex variable z instead of x.

1.1.1 Method of Undetermined Coefficients

It is convenient to construct particular solutions of Eq. (1.1) by using the method of undetermined coefficients (see, for example, the classical work of Boole [32, 33]).

Theorem 1.1 *Let a be a root of the equation $\sigma(x) = 0$. Then Eq.* (1.1) *has particular solution of the form*

$$y(x) = \sum_{n=0}^{\infty} c_n (x-a)^n, \tag{1.3}$$

where

$$\frac{c_{n+1}}{c_n} = -\frac{\lambda + n\left(\tau' + (n-1)\,\sigma''/2\right)}{(n+1)\,(\tau(a) + n\sigma'(a))}, \tag{1.4}$$

if:

(i) $\lim\limits_{m\to\infty} \dfrac{d^k}{dx^k}\, y_m(x) = \dfrac{d^k}{dx^k}\, y(x)$ *with* $k = 0, 1, 2$*;*
(ii) $\lim\limits_{m\to\infty} (\lambda - \lambda_m)\, c_m (x-a)^m = 0.$

(*Here* $y_m(x) = \sum\limits_{n=0}^{m} c_n (x-a)^n$ *and* $\lambda_m = -m\tau' - \dfrac{1}{2} m(m-1)\sigma''$*.*)

In the case $\sigma(x) = constant \neq 0$ series (1.3) *satisfies* (1.1) *when a is a root of the equation $\tau(x) = 0$,*

$$\frac{c_{n+2}}{c_n} = -\frac{\lambda + n\tau'}{(n+1)(n+2)\sigma} \tag{1.5}$$

and convergence conditions (i)–(ii) *are valid.*

The proof of Theorem 1.1 follows from the identity

$$\begin{aligned}\rho^{-1}\frac{d}{dx}\left[\sigma\rho\frac{d}{dx}(x-\xi)^n\right] &= n(n-1)\sigma(\xi)(x-\xi)^{n-2} \\ &\quad + n\tau_{n-1}(\xi)(x-\xi)^{n-1} - \lambda_n (x-\xi)^n,\end{aligned} \tag{1.6}$$

where $\tau_m(\xi) = \tau(\xi) + m\sigma'(\xi)$ and $\lambda_n = -n\tau' - \dfrac{1}{2}n(n-1)\sigma''$, which can be easily verified (see Exercise 1).

In fact, for a partial sum of the series (1.3) we can write

$$\left[\sigma(x)\frac{d^2}{dx^2}+\tau(x)\frac{d}{dx}+\lambda\right]y_m(x)$$
$$=\sigma(a)\sum_{n=0}^{m}c_n n(n-1)(x-a)^{n-2}$$
$$+\sum_{n=0}^{m}c_n n\tau_{n-1}(a)(x-a)^{n-1}$$
$$+\sum_{n=0}^{m}c_n\left(\lambda-\lambda_n\right)(x-a)^n.$$

By the hypothesis $\sigma(a)=0$, the first term in the right side is equal to zero. Equating coefficients in the next two terms with the aid of

$$\frac{c_{n+1}}{c_n}=\frac{\lambda_n-\lambda}{(n+1)\tau_n(a)},$$

which is equivalent to (1.4), we get

$$\left[\sigma(x)\frac{d^2}{dx^2}+\tau(x)\frac{d}{dx}+\lambda\right]y_m(x)=c_m\left(\lambda-\lambda_m\right)(x-a)^m. \tag{1.7}$$

Taking the limit $m\to\infty$ we prove the first part of the theorem under convergence conditions (i)–(ii).

When $\sigma=$ constant we can obtain in the same manner that

$$\left[\sigma(x)\frac{d^2}{dx^2}+\tau(x)\frac{d}{dx}+\lambda\right]y_m(x) \tag{1.8}$$
$$=\sigma\sum_{n=0}^{m}c_n n(n-1)(x-a)^{n-2}$$
$$+\sum_{n=0}^{m}c_n\left(\lambda-\lambda_n\right)(x-a)^n$$
$$=c_m\left(\lambda-\lambda_m\right)(x-a)^m,$$

which proves the second part of the theorem in the limit $m\to\infty$. □

Corollary *Equation* (1.1) *has polynomial solutions* $y_m(x)$ *corresponding to the eigenvalues* $\lambda=\lambda_m=-m\tau'-\frac{1}{2}m(m-1)\sigma''$, $m=0,1,2,\ldots$.

(It follows from (1.7) and (1.8).)

Examples With the aid of linear transformations of independent variable equation (1.1) for $\tau' \neq 0$ can be reduced to one of the following *canonical forms*

$$x(1-x)y'' + [\gamma - (\alpha + \beta + 1)x]\,y' - \alpha\beta y = 0, \tag{1.9}$$

$$xy'' + (\gamma - x)y' - \alpha y = 0, \tag{1.10}$$

$$y'' - 2xy' + 2\nu y = 0. \tag{1.11}$$

According to (1.3)–(1.5) the appropriate particular solutions are:
the *hypergeometric function*,

$$y(x) = {}_2F_1(\alpha, \beta; \gamma; x) = \sum_{n=0}^{\infty} \frac{(\alpha)_n (\beta)_n}{(\gamma)_n n!}\, x^n,$$

the *confluent hypergeometric function*,

$$y(x) = {}_1F_1(\alpha; \gamma; x) = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\gamma)_n n!}\, x^n,$$

and the *Hermite function*,

$$\begin{aligned} y(x) = H_\nu(x) &= \frac{2^\nu \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1-\nu}{2}\right)}\, {}_1F_1\left(-\frac{\nu}{2};\ \frac{1}{2};\ x^2\right) \\ &\quad + \frac{2^\nu \Gamma\left(-\frac{1}{2}\right)}{\Gamma\left(-\frac{\nu}{2}\right)}\, x\, {}_1F_1\left(\frac{1-\nu}{2};\ \frac{3}{2};\ x^2\right) \\ &= \frac{1}{2\Gamma(-\nu)} \sum_{n=0}^{\infty} \Gamma\left(\frac{n-\nu}{2}\right) \frac{(-2x)^n}{n!}, \end{aligned}$$

respectively. Here $(a)_n = a(a+1)\ldots(a+n-1) = \Gamma(a+n)\big/\Gamma(a)$, $n \geq 1$, $(a)_0 = 1$ and $\Gamma(a)$ is the gamma function of Euler.

Generally speaking, these solutions arise under some restrictions on the variable and parameters. They can be extended to wider domains by analytic continuation.

Definition All the hypergeometric series above are the special cases of the (*generalized*) *hypergeometric series* with r numerator parameters $\alpha_1, \ldots, \alpha_r$ and s

denominator parameters $\beta_1, \ldots, \beta_s$ defined by [16]

$$
{}_rF_s\left(\alpha_1, \alpha_2, \ldots, \alpha_r; \beta_1, \ldots, \beta_s;\; x\right) = {}_rF_s\left(\begin{matrix}\alpha_1, \alpha_2, \ldots, \alpha_r \\ \\ \beta_1, \ldots, \beta_s\end{matrix}\; ; x\right) \tag{1.12}
$$

$$
= \sum_{n=0}^{\infty} \frac{(\alpha_1)_n\,(\alpha_2)_n \ldots (\alpha_r)_n}{n!\,(\beta_1)_n \ldots (\beta_s)_n}\, x^n,
$$

where $(a)_n = a(a+1)\ldots(a+n-1)$ and $(a)_0 = 1$, as for the Hermite function.

By ratio test, the ${}_rF_s$ series converges absolutely for all complex values of x if $r \leq s$, and for $|x| < 1$ if $r = s+1$. By an extension of the ratio test (Bromwich [38], p. 241), it converges absolutely for $|x| = 1$ if $r = s+1$ and $x \neq 0$ or $r = s+1$ and Re $[\beta_1 + \ldots + \beta_s - (\alpha_1 + \ldots + \alpha_s)] > 0$. If $r > s+1$ and $x \neq 0$ or $r = s+1$ and $|x| > 1$, than this series diverges, unless it terminates.

More Solutions The solution (1.3)–(1.4) can be rewritten in the following explicit form

$$
y(x) = c_0 \sum_{n=0}^{\infty} \prod_{k=0}^{n-1} \frac{(\lambda - \lambda_k)\,(a-x)}{\tau_k(a)(k+1)}, \tag{1.13}
$$

where c_0 is a constant.

Using the expansion

$$
y(x) = \sum_n c_n (x-\xi)^{\alpha+n}, \qquad \frac{c_{n+1}}{c_n} = \frac{\lambda_{\alpha+n} - \lambda}{(\alpha+n+1)\tau_{\alpha+n}(a)}
$$

it is also not difficult to find solutions of a more general form

$$
y(x) = c_0 (x-a)^{\alpha} \sum_{n=0}^{\infty} \prod_{k=0}^{n-1} \frac{(\lambda - \lambda_{\alpha+k})\,(a-x)}{\tau_{\alpha+k}(a)(\alpha+k+1)}, \tag{1.14}
$$

provided that $\sigma(a) = 0$ and $\alpha\,\tau_{\alpha-1}(a) = 0$ (in particular, putting $\alpha = 0$ we recover (1.13)).

We can also satisfy (1.1) by using the series of the form

$$
y(x) = \sum_n \frac{c_n}{(x-\xi)^{\alpha+n}}, \qquad \frac{c_{n+1}}{c_n} = \frac{(\alpha+n)\tau_{-\alpha-n-1}(a)}{\lambda - \lambda_{-\alpha-n-1}},
$$

if $\sigma(a) = 0$ and $\lambda = \lambda_{-\alpha}$. Hence

$$
y(x) = \frac{c_0}{(x-a)^{\alpha}} \sum_{n=0}^{\infty} \prod_{k=0}^{n-1} \frac{(\alpha+k)\tau_{-\alpha-k-1}(a)}{(\lambda - \lambda_{-\alpha-k-1})\,(x-a)}. \tag{1.15}
$$

When σ = constant $\neq 0$ we can write the solution as

$$y(x) = \sum_n \frac{c_n}{(x-a)^{\alpha+n}}, \qquad \frac{c_{n+2}}{c_n} = -\frac{(\alpha+n)(\alpha+n+1)\sigma}{\lambda - \lambda_{-\alpha-n-2}}, \tag{1.16}$$

if $\tau(a) = 0$ and $\lambda = \lambda_{-\alpha}$ (for even integer values of n) or $\lambda = \lambda_{-\alpha-1}$ (for odd integer values of n).

1.1.2 Some Solutions of Hypergeometric Equations

Let us apply the method of undetermined coefficients to the main equations of hypergeometric type.

1. Consider particular solutions of the *hypergeometric equation*,

$$x(1-x)y'' + [c - (a+b+1)x]y' - aby = 0, \tag{1.17}$$

when $\sigma(x) = x(1-x)$, $\tau(x) = c - (a+b+1)x$ and $\lambda = -ab$. Here $\tau_\mu(\xi) = c - (a+b+1)\xi + \mu(1-2\xi)$, $\lambda_\mu = \mu(a+b+\mu)$ and $\lambda_\mu - \lambda = (a+\mu)(b+\mu)$. Equation $\sigma(\xi) = \xi(1-\xi) = 0$ has two solutions.

Solution (1.13) for $\sigma(0) = 0$ is the hypergeometric function,

$$y_1(x) = {}_2F_1(a,\ b;\ c;\ x), \quad |x| < 1; \tag{1.18}$$

for $\sigma(1) = 0$ we get

$$y_2(x) = {}_2F_1(a,\ b;\ a+b-c+1;\ 1-x), \quad |1-x| < 1. \tag{1.19}$$

The corresponding solutions (1.14) are

$$y_3(x) = x^{1-c}\,{}_2F_1(1+a-c,\ 1+b-c;\ 2-c;\ x) \tag{1.20}$$

and

$$y_4(x) = (1-x)^{c-a-b}\,{}_2F_1(c-b,\ c-a;\ c-a-b+1;\ 1-x). \tag{1.21}$$

Solutions (1.15) take the form

$$y_5(x) = x^{-a}\,{}_2F_1\left(a,\ a-c+1;\ a-b+1;\ x^{-1}\right), \quad |x| > 1, \tag{1.22}$$

$$y_6(x) = y_5(x)|_{a\leftrightarrow b}; \tag{1.23}$$

and

$$y_7(x) = (x-1)^{-a}\, {}_2F_1\left(a,\ c-b;\ a-b+1;\ \frac{1}{1-x}\right), \quad |1-x| > 1, \tag{1.24}$$

$$y_8(x) = y_7(x)\big|_{a\leftrightarrow b}\,. \tag{1.25}$$

Any three of these solutions are linearly dependent. For example,

$${}_2F_1(a,\ b;\ c;x) = (1-x)^{c-a-b}\, {}_2F_1(c-a,\ c-b;\ c;\ x) \tag{1.26}$$

and

$$\begin{aligned} &{}_2F_1(a,\ b;\ a+b-c+1;\ 1-x) \\ &\quad = \frac{\Gamma(a+b-c+1)\Gamma(1-c)}{\Gamma(a-c+1)\Gamma(b-c+1)}\, {}_2F_1(a,\ b;\ c;\ x) \\ &\quad + \frac{\Gamma(a+b-c+1)\Gamma(c-1)}{\Gamma(a)\Gamma(b)}\, x^{1-c}\, {}_2F_1(1+a-c,\ 1+b-c;\ 2-c;\ x). \end{aligned} \tag{1.27}$$

2. For the *confluent hypergeometric equation*,

$$xy'' + (c-x)y' - ay = 0, \tag{1.28}$$

we get $\sigma(x) = x$, $\tau(x) = c - x$, $\tau_\mu(x) = c + \mu - x$ and $\lambda = -a$. Particular solutions are

$$y_1(x) = {}_1F_1(a;\ c;\ x), \tag{1.29}$$

$$y_2(x) = x^{1-c}\, {}_1F_1(1+a-c;\ 2-c;\ x) \tag{1.30}$$

and

$$y_3(x) = x^{-a} \sum_{n=0}^{\infty} \frac{(a)_n (1+a-c)_n}{n!\,(-x)^n} \tag{1.31}$$

(this formal series does not converge unless it terminates).

3. In the case of the *Hermite equation*,

$$y'' - 2xy' + 2\nu y = 0, \tag{1.32}$$

when $\sigma(x) = 1$, $\tau(x) = -2x$ and $\lambda = 2\nu$, particular solutions have the forms

$$y_1(x) = H_\nu(x) = \frac{2^\nu \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1-\nu}{2}\right)} {}_1F_1\left(-\frac{\nu}{2};\ \frac{1}{2};\ x^2\right) + \frac{2^\nu \Gamma\left(-\frac{1}{2}\right)}{\Gamma\left(-\frac{\nu}{2}\right)} x\, {}_1F_1\left(\frac{1-\nu}{2};\ \frac{3}{2};\ x^2\right) \tag{1.33}$$

and

$$y_2(x) = x^\nu \sum_{n=0}^{\infty} \frac{\left(-\frac{\nu}{2}\right)_n \left(\frac{1-\nu}{2}\right)_n}{n!\,\left(-x^2\right)^n} \tag{1.34}$$

(this ${}_2F_0$ series diverges, unless it terminates).

4. Finally, let us consider the equation

$$xy'' + cy' - \lambda y = 0 \tag{1.35}$$

with $\sigma(x) = x$, $\tau(x) = c$ and $\tau_\mu = c + \mu$. Solutions (1.13) and (1.14) take the form

$$y_1(x) = {}_0F_1(-,\ c;\ \lambda x) = \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{(c)_n\, n!} \tag{1.36}$$

and

$$y_2(x) = x^{1-c}\, {}_0F_1(-;\ 2-c;\ \lambda x), \tag{1.37}$$

respectively. These functions are closely related to the *Bessel functions*

$$J_\nu(x) = \frac{(x/2)^\nu}{\Gamma(\nu+1)}\, {}_0F_1\left(-;\ \nu+1;\ -x^2\big/4\right). \tag{1.38}$$

1.2 Integral Representations

The special functions of hypergeometric type are also easily studied by means of integral representation, which holds for solutions of the differential equation (1.1).

1.2.1 Transformation to the Simplest Form

Many model problems in atomic, molecular, and nuclear physics lead to differential equations of the form [138]

$$u'' + \frac{\widetilde{\tau}\,(z)}{\sigma\,(z)}\,u' + \frac{\widetilde{\sigma}\,(z)}{\sigma^2\,(z)}\,u = 0, \tag{1.39}$$

where $\sigma\,(z)$ and $\widetilde{\sigma}\,(z)$ are polynomials of degrees at most two, $\widetilde{\tau}\,(z)$ is a polynomial of degree at most one. It is convenient to assume that z is a complex variable and the coefficients of the polynomials $\sigma\,(z)$, $\widetilde{\sigma}\,(z)$ and $\widetilde{\tau}\,(z)$ are arbitrary complex numbers. (If the independent variable takes the real values we shall write x instead of z.)

Let us try to transform the differential equation (1.39) to the simplest form by the change of unknown function $u = \varphi\,(z)\;\; y$ with the help of some special choice of function $\varphi\,(z)$.

Substituting $u = \varphi\,(z)\;\; y$ in (1.39) one gets

$$y'' + \left(\frac{\widetilde{\tau}}{\sigma} + 2\frac{\varphi'}{\varphi}\right)\,y' + \left(\frac{\widetilde{\sigma}}{\sigma^2} + \frac{\widetilde{\tau}}{\sigma}\frac{\varphi'}{\varphi} + \frac{\varphi''}{\varphi}\right)\,y = 0. \tag{1.40}$$

Equation (1.40) should not be more complicated than our original equation (1.39). Thus, it is natural to assume that the coefficient in front of y' has the form $\tau\,(z)\,/\sigma\,(z)$, where $\tau\,(z)$ is a polynomial of at most first degree. This implies the following first order differential equation

$$\frac{\varphi'}{\varphi} = \frac{\pi\,(z)}{\sigma\,(z)} \tag{1.41}$$

for the function $\varphi\,(z)$, where

$$\pi\,(z) = \frac{1}{2}\,(\tau\,(z) - \widetilde{\tau}\,(z)) \tag{1.42}$$

is a polynomial of the most first degree. As a result, Eq. (1.40) takes the form

$$y'' + \frac{\tau\,(z)}{\sigma\,(z)}\,u' + \frac{\overline{\sigma}\,(z)}{\sigma^2\,(z)}\,u = 0, \tag{1.43}$$

where

$$\overline{\sigma}\,(z) = \widetilde{\sigma}\,(z) + \pi^2\,(z) + \pi\,(z)\left[\widetilde{\tau}\,(z) - \sigma'\,(z)\right] + \pi'\,(z)\,\sigma\,(z)\,. \tag{1.44}$$

Functions $\tau\,(z)$ and $\overline{\sigma}\,(z)$ are polynomials of degrees at most one and two in z, respectively. Therefore, Eq. (1.43) is an equation of the same type as our original equation (1.39).

By using a special choice of the polynomial $\pi(z)$ we can reduce (1.43) to the simplest form assuming that

$$\overline{\sigma}(z) = \lambda\, \sigma(z), \tag{1.45}$$

where λ is some constant. Then Eq. (1.43) takes the form

$$\sigma(z)\; y'' + \tau(z)\; y' + \lambda\, y = 0. \tag{1.46}$$

We call Eq. (1.46) as a *differential equation of hypergeometric type* and its solutions as *functions of hypergeometric type*. In this contest, it is natural to call Eq. (1.39) as a *generalized differential equation of hypergeometric type*.

The condition (1.45) can be rewritten as

$$\pi^2 + \left(\widetilde{\tau} - \sigma'\right)\, \pi + \widetilde{\sigma} - k\sigma = 0, \tag{1.47}$$

where

$$k = \lambda - \pi'(z) \tag{1.48}$$

is a constant. Assuming that this constant is known we can find $\pi(z)$ as a solution of the quadratic equation

$$\pi(z) = \frac{\sigma' - \widetilde{\tau}}{2} \pm \sqrt{\left(\frac{\sigma' - \widetilde{\tau}}{2}\right)^2 - \widetilde{\sigma} + k\sigma}. \tag{1.49}$$

But $\pi(z)$ is a polynomial, therefore the second degree polynomial

$$p(z) = \left(\frac{\sigma'(z) - \widetilde{\tau}(z)}{2}\right)^2 - \widetilde{\sigma}(z) + k\sigma(z) \tag{1.50}$$

under the radical should be a square of a linear function and the discriminant of $p(z)$ should be zero. This condition gives an equation for the constant k, which is, generally, a quadratic equation. Given k as a solution of this equation, we find $\pi(z)$ by (1.49), then $\tau(z)$ and λ by (1.42) and (1.48). Finally, we find function $\varphi(z)$ as a solution of (1.41). It is clear that the reduction of Eq. (1.39) to the simplest form (1.46) can be accomplished by a few different ways in accordance with different choices of the constant k and different signs in (1.49) for $\pi(z)$.

The above transformation allows us to restrict ourself to study of the properties of solutions of Eq. (1.46).

1.2.2 Main Theorem

The following integral representation holds.

Theorem 1.2 *Let $\rho(z)$ satisfy the equation*

$$[\sigma(z)\rho(z)]' = \tau(z)\rho(z) \tag{1.51}$$

and let ν be a root of the equation

$$\lambda + \nu\tau' + \frac{1}{2}\,\nu(\nu-1)\sigma'' = 0. \tag{1.52}$$

Then (1.1) has particular solution of the form

$$y = y_\nu(z) = \frac{C_\nu}{\rho(z)} \int\limits_C \frac{\sigma^\nu(s)\rho(s)}{(s-z)^{\nu+1}}\,ds, \tag{1.53}$$

where C_ν is a constant and C is a contour in the complex s-plane, if:

(i) *the derivative of the integral*

$$\varphi_{\nu\mu}(z) = \int\limits_C \frac{\rho_\nu(s)}{(s-z)^{\mu+1}}\,ds \quad \textit{with} \quad \rho_\nu(s) = \sigma^\nu(s)\rho(s)$$

can be evaluated for $\mu = \nu - 1$ and $\mu = \nu$ by using the formula

$$\varphi'_{\nu\mu}(z) = (\mu+1)\varphi_{\nu,\,\mu+1}(z);$$

(ii) *the contour C is chosen so that the equality*

$$\left.\frac{\sigma(s)\rho_\nu(s)}{(s-z)^{\nu+1}}\right|_{s_1}^{s_2} = 0$$

holds, where s_1 and s_2 are the endpoints of the contour C.

Proof The function $\rho_\nu(s) = \sigma^\nu(s)\rho(s)$ satisfies the equation

$$[\sigma(s)\rho_\nu(s)]' = \tau_\nu(s)\rho_\nu(s),$$

where $\tau_\nu(s) = \tau(s) + \nu\sigma'(s)$. We multiply both sides of this equality by $(s-z)^{-\nu-1}$ and integrate over the contour C. Upon integrating by parts we obtain

$$\left.\frac{\sigma(s)\rho_\nu(s)}{(s-z)^{\nu+1}}\right|_{s_1}^{s_2} + (\nu+1)\int\limits_C \frac{\sigma(s)\rho_\nu(s)}{(s-z)^{\nu+2}}\,ds = \int\limits_C \frac{\tau_\nu(s)\rho_\nu(s)}{(s-z)^{\nu+1}}\,ds. \tag{1.54}$$

By hypothesis, the first term is equal to zero. We expand the polynomials $\sigma(s)$ and $\tau_\nu(s)$ in powers of $s - z$:

$$\sigma(s) = \sigma(z) + \sigma'(z)(s-z) + \frac{1}{2}\,\sigma''(s-z)^2,$$
$$\tau_\nu(s) = \tau_\nu(z) + \tau'_\nu(s-z).$$

Taking into account the integral formulas for the functions $\varphi_{\nu,\,\nu-1}$, $\varphi_{\nu\nu}$ and $\varphi_{\nu,\,\nu+1}$, we arrive at the relation

$$(\nu+1)\left[\sigma(z)\varphi_{\nu,\,\nu+1}+\sigma'(z)\varphi_{\nu\nu}+\frac{1}{2}\sigma''\varphi_{\nu,\,\nu-1}\right]=\tau_\nu(z)\varphi_{\nu\nu}+\tau_\nu'\varphi_{\nu,\,\nu-1}.$$

Upon substituting $\tau_\nu=\tau+\nu\sigma'$ and using the formula $\varphi_{\nu\nu}'=(\nu+1)\varphi_{\nu,\,\nu+1}$ we get

$$\sigma\varphi_{\nu\nu}'+\left(\sigma'-\tau\right)\varphi_{\nu\nu}=\left(\tau'+\frac{1}{2}(\nu-1)\sigma''\right)\varphi_{\nu,\,\nu-1}. \tag{1.55}$$

At the same time, by differentiating the relation $\sigma\rho y'=C_\nu\sigma\varphi_{\nu\nu}$ we find that

$$\frac{1}{C_\nu}\sigma\rho y'=\sigma\varphi_{\nu\nu}'+\left(\sigma'-\tau\right)\varphi_{\nu\nu}. \tag{1.56}$$

Comparing (1.55) and (1.56) we obtain

$$\sigma\rho y'=\kappa_\nu C_\nu\varphi_{\nu,\,\nu-1}, \tag{1.57}$$

where $\kappa_\nu=\tau'+(\nu-1)\,\sigma''\big/2$. Upon differentiating (1.57) we arrive at (1.1) in the self-adjoint form

$$\left(\sigma\rho y'\right)'+\lambda\rho y=0,$$

where $\lambda=-\nu\kappa_\nu=-\nu\tau'-\nu(\nu-1)\,\sigma''\big/2$. This proves the theorem. □

In the proof of Theorem 1.2 we have, en route, deduced the formula (1.57), which is a simple *integral representation* for the first derivative of the function of hypergeometric type:

$$y_\nu'(z)=\frac{C_\nu^{(1)}}{\sigma(z)\rho(z)}\int\limits_C\frac{\rho_\nu(s)}{(s-z)^\nu}\,ds, \tag{1.58}$$

where $C_\nu^{(1)}=\kappa_\nu C_\nu=\left(\tau'+\frac{1}{2}(\nu-1)\sigma''\right)C_\nu$. Hence

$$y_\nu^{(k)}(z)=\frac{C_\nu^{(k)}}{\rho_k(z)}\varphi_{\nu,\,\nu-k}(z)=\frac{C_\nu^{(k)}}{\sigma^k(z)\rho(z)}\int\limits_C\frac{\rho_\nu(s)}{(s-z)^{\nu-k+1}}\,ds, \tag{1.59}$$

where $C_\nu^{(k)}=\prod\limits_{p=0}^{k-1}\left(\tau'+\frac{\nu+p-1}{2}\sigma''\right)C_\nu$.

1.2.3 Integrals for Hypergeometric and Bessel Functions

Using Theorem 1.2 we can obtain integral representations for all the most commonly used special functions of hypergeometric type, in particular, for the hypergeometric functions:

$$ {}_2F_1(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 t^{\alpha-1}(1-t)^{\gamma-\alpha-1}(1-zt)^{-\beta}\, dt, \quad (1.60) $$

$$ {}_1F_1(\alpha; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 t^{\alpha-1}(1-t)^{\gamma-\alpha-1}e^{zt}\, dt, \quad (1.61) $$

$$ H_\nu(z) = \frac{1}{\Gamma(-\nu)} \int_0^\infty e^{-t^2-2zt}t^{-\nu-1}\, dt. \quad (1.62) $$

Here $\operatorname{Re}\gamma > \operatorname{Re}\alpha > 0$ and $\operatorname{Re}(-\nu) > 0$.

Let us mention also some solutions of the *Bessel equation*,

$$ z^2u'' + zu' + \left(z^2 - \nu^2\right)u = 0. \quad (1.63) $$

With the aid of the change of the function $u = \varphi(z)y$ when $\varphi(z) = z^\nu e^{iz}$ this equation can be reduced to the hypergeometric form

$$ zy'' + (2iz + 2\nu + 1)y' + i(2\nu + 1)y = 0 \quad (1.64) $$

and on the base on Theorem 1.2 one can obtain the *Poisson integral representations* for the Bessel functions of the first kind, $J_\nu(z)$, and the Hankel functions of the first and second kind, $H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$:

$$ J_\nu(z) = \frac{(z/2)^\nu}{\sqrt{\pi}\,\Gamma(\nu+1/2)} \int_{-1}^1 \left(1-t^2\right)^{\nu-1/2}\cos zt\, dt, \quad (1.65) $$

$$ H_\nu(z) = \sqrt{\frac{2}{\pi z}}\, \frac{e^{\pm i\left(z-\frac{\pi}{2}\nu-\frac{\pi}{4}\right)}}{\Gamma(\nu+1/2)} \int_0^\infty e^{-t}t^{\nu-1/2}\left(1 \pm \frac{it}{2z}\right)^{\nu-1/2} dt, \quad (1.66) $$

where $\operatorname{Re}\nu > -1/2$. It is then possible to deduce from these integral representations all the remaining properties of these functions. (For details, see Nikiforov and Uvarov [138], or Watson [193], Whittaker and Watson [195].)

1.3 Classical Orthogonal Polynomials

The Jacobi, Laguerre and Hermite polynomials are the simplest solutions of Eq. (1.1).

1.3.1 Main Property

By differentiating (1.1) we can easily verify that the function $\upsilon_1(x) = y'(x)$ satisfies the equation of the same type

$$\sigma(x)\upsilon_1'' + \tau_1(x)\upsilon_1' + \mu_1\upsilon_1 = 0, \tag{1.67}$$

where $\tau_1(x) = \tau(x) + \sigma'(x)$ is a polynomial of degree at most one and $\mu_1 = \lambda + \tau'(x)$ is a constant.

The converse is also true: any solution of (1.67) is the derivative of a solution of (1.1) if $\lambda = \mu_1 - \tau' \neq 0$. Let $\upsilon_1(x)$ be a solution of (1.67) and define the function

$$y(x) = -\frac{1}{\lambda}\left(\sigma(x)\upsilon_1' + \tau(x)\upsilon_1\right).$$

We have

$$\lambda y' = -\left(\sigma\upsilon_1'' + \tau_1\upsilon_1' + \tau'\upsilon_1\right) = \lambda\upsilon_1$$

or $\upsilon_1 = y'(x)$ and, therefore, $y(x)$ satisfies (1.1).

1.3.2 Rodrigues Formula

By differentiating (1.1) n times we obtain an equation of hypergeometric type for the function $\upsilon_n(x) = y^{(n)}(x)$,

$$\sigma(x)\upsilon_n'' + \tau_n(x)\upsilon_n' + \mu_n\upsilon_n = 0, \tag{1.68}$$

where

$$\tau_n(x) = \tau(x) + n\sigma'(x), \tag{1.69}$$

$$\mu_n = \lambda + n\tau' + \frac{1}{2}n(n-1)\sigma''. \tag{1.70}$$

This property lets us construct the simplest solutions of (1.1) corresponding to some values of λ. Indeed, when $\mu_n = 0$ Eq. (1.68) has the solution $\upsilon_n =$ constant. Since $\upsilon_n(x) = y^{(n)}(x)$, Eq. (1.1) has a particular solution $y = y_n(x)$ which is a polynomial of degree n if

$$\lambda = \lambda_n = -n\tau' - \frac{1}{2}n(n-1)\sigma'' \quad (n = 0, 1, 2, \ldots). \tag{1.71}$$

To find these polynomials explicitly let us rewrite Eqs. (1.1) and (1.68) in the self-adjoint forms

$$\left(\sigma\rho y'\right)' + \lambda\rho y = 0, \tag{1.72}$$

$$\left(\sigma\rho_n \upsilon_n'\right)' + \mu_n \rho_n \upsilon_n = 0. \tag{1.73}$$

Functions $\rho(x)$ and $\rho_n(x)$ satisfy the first-order differential equations

$$(\sigma\rho)' = \tau\rho, \tag{1.74}$$

$$(\sigma\rho_n)' = \tau_n\rho_n. \tag{1.75}$$

So,

$$\frac{(\sigma\rho_n)'}{\rho_n} = \tau + n\sigma' = \frac{(\sigma\rho)'}{\rho} + n\sigma',$$

whence

$$\frac{\rho_n'}{\rho_n} = \frac{\rho'}{\rho} + n\,\frac{\sigma'}{\sigma}$$

and, consequently,

$$\rho_n(x) = \sigma^n(x)\rho(x). \tag{1.76}$$

Since $\sigma\rho_n = \rho_{n+1}$ and $\upsilon_n' = \upsilon_{n+1}$ we can rewrite (1.73) in the form

$$\rho_n\upsilon_n = -\frac{1}{\mu_n}\left(\rho_{n+1}\upsilon_{n+1}\right)'.$$

Hence we obtain successively

$$\begin{aligned}\rho y = \rho_0 \upsilon_0 &= -\frac{1}{\mu_0}\,(\rho_1\upsilon_1)' \\ &= \left(-\frac{1}{\mu_0}\right)\left(-\frac{1}{\mu_1}\right)(\rho_2\upsilon_2)'' \\ &\;\;\vdots \\ &= \frac{1}{A_n}\,(\rho_n\upsilon_n)^{(n)},\end{aligned}$$

where

$$A_0 = 1, \quad A_n = (-1)^n \prod_{t=0}^{n-1} \mu_k. \tag{1.77}$$

If $y = y_n(x)$ is a polynomial of degree n, then $\upsilon_n = y_n^{(n)}(x) = \text{constant}$ and we arrive at the *Rodrigues formula* for polynomial solutions of (1.1),

$$y_n(x) = \frac{B_n}{\rho(x)}\left[\sigma^n(x)\rho(x)\right]^{(n)}, \qquad n \geq 1, \tag{1.78}$$

where $B_n = A_n^{-1} y_n^{(n)}$ is a constant. These solutions correspond to the eigenvalues (1.71).

Remark We have also found the explicit series representations for polynomials (1.78) in Theorem 1.1.

1.3.3 Orthogonality

The polynomial solutions of (1.1) obey an orthogonality property. Let us write equations for polynomials $y_n(x)$ and $y_m(x)$ in the self-adjoint form,

$$\left(\sigma(x)\rho(x)y_n'(x)\right)' + \lambda_n \rho(x) y_n(x) = 0,$$

$$\left(\sigma(x)\rho(x)y_m'(x)\right)' + \lambda_m \rho(x) y_m(x) = 0,$$

multiply the first equation by $y_m(x)$ and the second by $y_n(x)$, subtract the second equality from the first one and then integrate the result over x on the interval (a, b). Since

$$\begin{aligned}y_m(x)\left(\sigma(x)\rho(x)y_n'(x)\right)' - y_n\left(\sigma(x)\rho(x)y_m'(x)\right)' \\ = \frac{d}{dx}\left(\sigma(x)\rho(x)W\left[y_m(x), y_n(x)\right]\right),\end{aligned}$$

where $W(u, \upsilon) = u\upsilon' - \upsilon u'$ is the Wronskian, we get

$$(\lambda_m - \lambda_n) \int_a^b y_m(x) y_n(x) \rho(x)\, dx \tag{1.79}$$
$$= \left. (\sigma(x)\rho(x) W\,[y_m(x), y_n(x)]) \right|_a^b .$$

If the conditions

$$\left. \sigma(x)\rho(x)x^k \right|_{x=a,b} = 0 \quad (k = 0, 1, 2, \ldots) \tag{1.80}$$

are satisfied for some points a and b, then the right side of (1.79) vanishes because the Wronskian is a polynomial in x. Therefore, we arrive at the *orthogonality property*

$$\int_a^b y_m(x) y_n(x) \rho(x)\, dx = 0 \tag{1.81}$$

provided that $\lambda_n \neq \lambda_m$. We can replace this condition by $m \neq n$ due to the relation

$$\lambda_n - \lambda_m = (m - n) \left(\tau' + \frac{n + m - 1}{2}\, \sigma'' \right),$$

if $\tau' + (m + n - 1)\, \sigma'' \big/ 2 \neq 0$.

We shall refer to polynomial solutions of (1.1) obeying the orthogonality property (1.81) with respect to a positive weight function as *classical orthogonal polynomials*.

1.3.4 Classification

Equation (1.74) for the weight function $\rho(x)$ is usually called the *Pearson equation*. By using the linear transformations of independent variable x we can reduce solutions of (1.74) to the following canonical forms

$$\rho(x) = \begin{cases} (1-x)^\alpha (1+x)^\beta & \text{for} \quad \sigma(x) = 1 - x^2, \\ x^\alpha e^{-x} & \text{for} \quad \sigma(x) = x, \\ e^{-x^2} & \text{for} \quad \sigma(x) = 1. \end{cases}$$

The corresponding orthogonal polynomials are the *Jacobi polynomials* $P_n^{(\alpha,\beta)}(x)$, the *Laguerre polynomials* $L_n^{\alpha}(x)$ and the *Hermite polynomials* $H_n(x)$.

The basic information about classical orthogonal polynomials is presented in the table below, which contains also the leading terms $y_n(x) = a_n x^n + b_n x^{n-1} + \dots$ for these polynomials, squared norms,

$$d_n^2 = \int_a^b y_n^2(x)\rho(x)\,dx, \qquad (1.82)$$

and the coefficients of the three-term recurrence relation

$$x\,y_n(x) = \alpha_n\,y_{n+1}(x) + \beta_n\,y_n(x) + \gamma_n\,y_{n-1}(x), \qquad n \ge 1, \qquad (1.83)$$

where

$$\alpha_n = \frac{a_n}{a_{n+1}}, \quad \beta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}, \quad \gamma_n = \alpha_{n-1}\frac{d_n^2}{d_{n-1}^2}, \qquad n \ge 1. \qquad (1.84)$$

$y_n(x)$	$P_n^{(\alpha,\beta)}(x)\ (\alpha > -1, \beta > -1)$	$L_n^{\alpha}(x)\ (\alpha > -1)$	$H_n(x)$
(a, b)	$(-1, 1)$	$(0, \infty)$	$(-\infty, \infty)$
$\rho(x)$	$(1-x)^{\alpha}(1+x)^{\beta}$	$x^{\alpha}e^{-x}$	e^{-x^2}
$\sigma(x)$	$1 - x^2$	x	1
$\tau(x)$	$\beta - \alpha - (\alpha + \beta + 2)\,x$	$1 + \alpha - x$	$-2x$
λ_n	$n(\alpha + \beta + n + 1)$	n	$2n$
B_n	$\frac{(-1)^n}{2^n n!}$	$\frac{1}{n!}$	$(-1)^n$
a_n	$\frac{\Gamma(\alpha+\beta+2n+1)}{2^n n!\Gamma(\alpha+\beta+n+1)}$	$\frac{(-1)^n}{n!}$	2^n
b_n	$\frac{(\alpha-\beta)\Gamma(\alpha+\beta+2n)}{2^n(n-1)!\Gamma(\alpha+\beta+n+1)}$	$(-1)^{n-1}\frac{\alpha+n}{(n-1)!}$	0
d^2	$\frac{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{n!(\alpha+\beta+2n+1)\Gamma(\alpha+\beta+n+1)}$	$\frac{\Gamma(\alpha+n+1)}{n!}$	$2^n n!\sqrt{\pi}$
α_n	$\frac{2(n+1)(\alpha+\beta+n+1)}{(\alpha+\beta+2n+1)(\alpha+\beta+2n+2)}$	$-(n+1)$	$\frac{1}{2}$
β_n	$\frac{\beta^2-\alpha^2}{(\alpha+\beta+2n)(\alpha+\beta+2n+2)}$	$\alpha + 2n + 1$	0
γ_n	$\frac{2(\alpha+n)(\beta+n)}{(\alpha+\beta+2n)(\alpha+\beta+2n+1)}$	$-(\alpha + n)$	n

1.3.5 Functions of the Second Kind

Consider Eq. (1.1) in the complex z-plane for the eigenvalues (1.71). By using Theorem 1.2 we can choose a particular solution of the form

$$y = y_n(z) = \frac{B_n\, n!}{2\pi i \rho(z)} \int\limits_C \frac{\rho_n(s)\, ds}{(s-z)^{n+1}}, \tag{1.85}$$

where B_n is a constant, $\rho_n(s) = \sigma^n(s)\rho(s)$ and C is a closed contour in the complex s-plane that encloses the point $s = z$. Here the conditions of the theorem hold.

The solution (1.85) defines *classical orthogonal polynomials*. In fact, in view of

$$\frac{d^n}{dz^n}\left(\frac{1}{s-z}\right) = \frac{n!}{(s-z)^{n+1}},$$

we arrive at the Rodrigues formula

$$\begin{aligned} y_n(z) &= \frac{B_n}{2\pi i \rho(z)} \int\limits_C \rho_n(s)\, \frac{d^n}{dz^n}\left(\frac{1}{s-z}\right) ds \\ &= \frac{B_n}{2\pi i \rho(z)} \frac{d^n}{dz^n} \int\limits_C \frac{\rho_n(s)}{s-z}\, ds = \frac{B_n}{\rho(z)} \frac{d^n}{dz^n}\, [\rho_n(z)] \end{aligned}$$

with the aid of the Cauchy integral formula.

As a second linearly independent solution of (1.1) for $\lambda = \lambda_n$ we take a function of the form

$$y = Q_n(z) = \frac{B_n\, n!}{\rho(z)} \int\limits_a^b \frac{\rho_n(s)\, ds}{(s-z)^{n+1}}, \qquad z \neq s, \tag{1.86}$$

where B_n is the constant in the Rodrigues formula (1.78) and $\rho_n(s) = \sigma^n(s)\rho(s)$. The points a and b are chosen so that (1.80) holds, which is the case for the classical orthogonal polynomials. In this case, the conditions of Theorem 1.2 hold.

The functions $Q_n(z)$ defined by (1.86) are called *functions of the second kind*.

We can find a relation between the functions $Q_n(z)$ and the polynomials $y_n(z)$. From definition

$$Q_n(z) = -\frac{B_n(n-1)!}{\rho(z)} \int_a^b \rho_n(s) \frac{d}{ds}\left[\frac{1}{(s-z)^n}\right] ds$$

$$= \frac{B_n(n-1)!}{\rho(z)} \left(- \frac{\rho_n(s)}{(s-z)^n}\Big|_a^b + \int_a^b \frac{[\rho_n(s)]'}{(s-z)^n} ds \right).$$

The first term vanishes by virtue of (1.80). Similarly,

$$Q_n(z) = \frac{B_n}{\rho(z)} \int_a^b \frac{[\rho_n(s)]^{(n)}}{s-z} ds,$$

and we find that

$$Q_n(z) = \frac{1}{\rho(z)} \int_a^b \frac{y_n(s)}{s-z} \rho(s)\, ds \tag{1.87}$$

due to the Rodrigues formula.

By setting $y_n(s) = [y_n(s) - y_n(z)] + y_n(z)$ Eq. (1.87) can be written in the convenient form

$$Q_n(z) = \frac{1}{B_0} y_n(z) Q_0(z) + \frac{1}{\rho(z)} q_{n-1}(z), \tag{1.88}$$

where

$$Q_0(z) = \frac{B_0}{\rho(z)} \int_a^b \frac{\rho(s)}{s-z} ds \tag{1.89}$$

and

$$q_{n-1}(s) = \int_a^b \frac{y_n(s) - y_n(z)}{s-z} \rho(s)\, ds \tag{1.90}$$

is a polynomial of degree $n-1$ in s, which is called a *polynomial of the second kind*. It follows from (1.88) that all the singularities of the second solution $Q_n(z)$ in the complex z-plane are determined by the behavior of the functions $Q_0(z)$ and $1\big/\rho(z)$.

It is possible to derive from (1.86)–(1.88) all the main properties of the functions of the second kind. Using the identity

$$\frac{1}{s-z}=-\frac{1}{z}\sum_{k=0}^{p}\left(\frac{s}{z}\right)^k+\frac{s^{p+1}}{(s-z)z^{p+1}}$$

and the orthogonality property

$$\int_a^b s^k y_n(s)\rho(s)\,ds=0,\quad k<n$$

we obtain from (1.87) an expansion of the form

$$\rho(z)Q_n(z)=-\sum_{k=n}^{p}\frac{1}{z^{k+1}}\int_a^b s^k y_n(s)\rho(s)\,ds+\frac{r_p(z)}{z^{p+1}},\tag{1.91}$$

where

$$r_p(z)=\int_a^b s^{p+1}\rho(s)\,ds.$$

Equation (1.91) determines the *asymptotic behavior* of the functions $Q_n(z)$ as $|z|\to\infty$. In particular, for $p=n$, (1.91) yields

$$Q_n(z)=\frac{\left(d_n^2\big/a_n\right)}{\rho(z)z^{n+1}}\left[1+O\left(\frac{1}{z}\right)\right],\quad z\to\infty.\tag{1.92}$$

Hence, the functions of the second kind $Q_n(z)$ and the classical orthogonal polynomials $y_n(z)$ have different assumptions behaviors at $|z|\to\infty$, so that they are two linearly-independent solutions of (1.1) for $\lambda=\lambda_n$.

According to (1.87), we have

$$zQ_n(z)=\frac{1}{\rho(z)}\int_a^b\frac{s\,y_n(s)}{s-z}\rho(s)\,ds-\frac{1}{\rho(z)}\int_a^b y_n(s)\rho(s)\,ds.$$

Therefore, the functions $Q_n(z)$ satisfy a *three-term recurrence relation*

$$z\,Q_n(z) = \alpha_n\,Q_{n+1}(z) + \beta_n\,Q_n(z) + \gamma_n\,Q_{n-1}(z) \quad \text{if} \quad n \geq 1$$

and

$$z\,Q_0(z) = \alpha_0\,Q_1(z) + \beta_0\,Q_0(z) - \frac{1}{\rho(z)}\,\frac{d_0^2}{B_0} \quad \text{if} \quad n = 0,$$

where α_n, β_n and γ_n are the coefficients of the recurrence relation (1.83) for the classical orthogonal polynomials.

It follows from (1.86) and (1.58) that the derivative of $Q_n(z)$ can be represented in the form

$$Q_n'(z) = \frac{\kappa_n B_n\,n!}{\sigma(z)\rho(z)} \int_a^b \frac{\rho_n(s)}{(s-z)^n}\,ds,$$

where $\kappa_n = \tau' + (n-1)\sigma''/2$. For $n = 0$ this leads to the simple differential equation for $Q_0(z)$,

$$\sigma(z)\rho(z)Q_0'(z) = C_0, \tag{1.93}$$

where $C_0 = \kappa_0 B_0 d_0^2$. According to (1.93) we obtain

$$Q_0(z) = Q_0\,(z_0) + C_0 \int_{z_0}^{z} \frac{ds}{\sigma(s)\rho(s)}, \tag{1.94}$$

where it is convenient to choose for z_0 a value of z for which $Q_0\,(z_0) = 0$.

The functions $Q_n(z)$ have not been defined when $z \in [a, b]$. At this interval it is convenient to set

$$\begin{aligned}\rho(x)Q_n(x) &= \frac{1}{2}\left[\rho(x-i0)Q_n(x-i0) + \rho(x+i0)Q_n(x+i0)\right] \qquad (1.95)\\ &= \lim_{\epsilon\to 0}\frac{1}{2}\left[\rho(x-i\epsilon)Q_n(x-i\epsilon) + \rho(x+i\epsilon)Q_n(x+i\epsilon)\right].\end{aligned}$$

We note also the relation

$$\rho(x-i0)Q_n(x-i0) - \rho(x+i0)Q_n(x+i0) = 2\pi i\rho(x)y_n(x), \tag{1.96}$$

which comes form (1.87) due to

$$\frac{1}{x-i0} - \frac{1}{x+i0} = 2\pi i\,\delta(x)$$

where $\delta(x)$ is Dirac's delta function.

The explicit forms of the functions of the second kind can be easily found by comparing the definition (1.86) with the integral representations for the hypergeometric functions in (1.60), (1.62) and Exercise 9(ii). They are

$$Q_n^{(\alpha,\beta)}(z) = (-1)^n 2^{\alpha+\beta+n+1} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{\Gamma(\alpha+\beta+2n+2)}$$
$$\times (1-z)(1+z)^{-\beta} {}_2F_1\left(\begin{matrix} n+1,\ \alpha+n+1 \\ \\ \alpha+\beta+2n+2 \end{matrix};\ 2(1-z)^{-1}\right), \quad (1.97)$$
$$z \notin [-1, 1];$$
$$Q_n^{\alpha}(z) = e^{-i\pi\alpha}\,\Gamma(\alpha+n+1)$$
$$\times e^z\, G(\alpha+n+1;\ \alpha+1;\ -z), \quad 0 < \arg z < 2\pi; \quad (1.98)$$

and

$$Q_n(z) = 2^{n+1} n! \sqrt{\pi}\, e^{z^2 \mp i\frac{\pi}{2}(n-1)}\, H_{-n-1}(\mp iz) \quad (\mathrm{Im}\, z > 0 \quad \text{or} \quad \mathrm{Im}\, z < 0) \quad (1.99)$$

corresponding to the cases of the Jacobi, Laguerre and Hermite polynomials, respectively.

1.3.6 Complex Orthogonality

Classical orthogonal polynomials obey an interesting orthogonality property with respect to a complex measure. To prove it, let us start with the identity

$$[\sigma(z)\rho(z)W(y_m, y_n)]' = (\lambda_m - \lambda_n)\, y_m(z)y_n(z)\rho(z), \quad (1.100)$$

derived in Sect. 1.3.3. Multiply both sides of (1.100) by the function $Q_0(z)$ defined by (1.89) and integrate the result over a contour C in the complex z-plane. Upon integrating by parts we obtain with the aid of (1.93) that

$$(\lambda_m - \lambda_n) \int_C y_m(z)y_n(z)\rho(z)Q_0(z)\, dz \quad (1.101)$$
$$= \left(\sigma(z)\rho(z)Q_0(z)\, W[y_m(z), y_n(z)]\right)\Big|_{z_1}^{z_2}$$
$$-C_0 \int_C W[y_m(z), y_n(z)]\, dz.$$

Here, z_1 and z_2 are the end points of the contour C.

Suppose first that (a, b) is a finite interval on the real axis. Then for a closed contour C, which encloses interval (a, b) in the complex z-plane, the right side of (1.101) vanishes due to Cauchy's theorem. As a result, we arrive at the *complex orthogonality property*

$$\int_C y_m(z) y_n(z) \rho(z) Q_0(z)\, dz = 0, \quad m \neq n, \tag{1.102}$$

provided that $\tau' + (m+n-1)\sigma''/2 \neq 0$. This case corresponds to the Jacobi polynomials $y_n = P_n^{(\alpha,\beta)}(z)$ if $\alpha + \beta + 1 \neq -1, -2, \ldots$ The complex weight function in (1.102) is a weight function in a wider range of parameters than the real weight function for the Jacobi polynomials. (See Exercise 24.)

In the case $b = +\infty$, which corresponds to the Laguerre polynomials $y_n = L_n^{\alpha}(z)$, consider (1.101) for the contour $C_\epsilon(R)$. Taking the limit $\epsilon \to 0$ $(\epsilon > 0)$ due to (1.96) we obtain

$$\begin{aligned} PV \int_{C(R)} & y_m(z) y_n(z) \rho(z) Q_0(z)\, dz \\ &= \frac{2\pi i}{\lambda_m - \lambda_n} \sigma(R) \rho(R)\, W\left[y_m(R), y_n(R)\right], \end{aligned} \tag{1.103}$$

where PV denotes Cauchy's principal value integral and $C(R) = \lim_{\epsilon \to 0} C_\epsilon(R)$ is the closed contour. In the limit $R \to \infty$ we arrive at a complex orthogonality property, for the Laguerre polynomials $L_n^{\alpha}(z)$ if $\alpha > -1$ (Exercise 25).

Finally, consider the case of the Hermite polynomials $y_n = H_n(z)$ when $a = -\infty$ and $b = +\infty$. In the limit $\epsilon \to 0$ for the two contours $C_\epsilon^{\pm}(R_1, R_2)$, we get

$$\begin{aligned} PV \int_{C(R_1,R_2)} & y_m(z) y_n(z) \rho(z) Q_0(z)\, dz \\ &= \frac{2\pi i}{\lambda_m - \lambda_n} \left(\sigma(R_1)\, \rho(R_1)\; W\left[y_m(R_1), y_n(R_1)\right] \right. \\ &\qquad \left. + \sigma(R_2)\, \rho(R_2)\; W\left[y_m(R_2), y_n(R_2)\right]\right), \end{aligned} \tag{1.104}$$

where $C(R_1, R_2) = \lim_{\epsilon \to 0} C_\epsilon^{+}(R_1, R_2) \bigcup C_\epsilon^{-}(R_1, R_2)$ is the closed contour. Taking the limits $R_{1,2} \to \mp\infty$ we obtain a complex orthogonality property of the Hermite polynomials (Exercise 26).

Exercises

1. (i) By using the Pearson equation $(\sigma\rho)' = \tau\rho$ show that the function $\rho_\nu(s) = \sigma^\nu(s)\rho(s)$ satisfy the equation

$$[\sigma(s)\rho_\nu(s)]' = \tau_\nu(s)\rho_\nu(s),$$

where $\tau_\nu(s) = \tau(s) + \nu\sigma'(s)$;

(ii) using this equation and the expansions

$$\sigma(s) = \sigma(z) + \sigma'(z)(s-z) + \frac{1}{2}\sigma''(s-z)^2,$$
$$\tau_\nu(s) = \tau_\nu(z) + \tau_\nu'(s-z)$$

verify the identity

$$\begin{aligned}\frac{d}{ds}\left[\sigma(s)\rho_\nu(s)(s-z)^\mu\right] &= \mu\sigma(z)(s-z)^{\mu-1}\rho_\nu(s)\\ &\quad + \left[\tau_\nu(z) + \mu\sigma'(z)\right](s-z)^\mu\rho_\nu(s)\\ &\quad + \left(\tau_\nu' + \frac{1}{2}\mu\sigma''\right)(s-z)^{\mu+1}\rho_\nu(s),\end{aligned}$$

where μ and ν are arbitrary complex numbers;

(iii) verify identity (1.6) for arbitrary complex values of n.

2. Derive (1.13)–(1.16).
3. (i) Show that the series (1.14) with arbitrary α satisfies a non-homogeneous differential equation (1.1) with the right side

$$G_\alpha(x) = c_0\alpha\,\tau_{\alpha-1}(a)(x-a)^{\alpha-1}$$

if $\sigma(a) = 0$.

(ii) Show that the series of the form

$$\begin{aligned}u(x) &= x^\alpha \sum_{n=0}^{\infty} \frac{(\alpha+a)_n(\alpha+b)_n}{(\alpha+c)_n(\alpha+1)_n}\, x^n\\ &= x^\alpha\, {}_3F_2\left(\begin{matrix}1,\ \alpha+a,\ \alpha+b\\ \\ \alpha+1,\ \alpha+c\end{matrix}\ ;\ x\right), \qquad |x|<1,\end{aligned}$$

satisfies the non-homogeneous equation

$$x(1-x)u'' + [c-(a+b+1)x]u' - abu = \alpha(c+\alpha-1)x^{\alpha-1}$$

(putting $\alpha = 0$ or $\alpha = 1 - c$, we recover (1.18) and (1.20), respectively). Find the general solution of this non-homogeneous equation.

(iii) Show that the series

$$\upsilon(x) = x^{\alpha} \sum_{n=0}^{\infty} \frac{(\alpha + a)_n}{(\alpha + c)_n (\alpha + 1)_n} x^n = x^{\alpha}\, {}_2F_2 \left(\begin{array}{c} 1,\ \alpha + a \\ \\ \alpha + 1,\ \alpha + c \end{array} ;\ x \right), \quad |x| < 1,$$

satisfies the non-homogeneous equation

$$x\upsilon'' + (c - x)\upsilon' - a\upsilon = \alpha(c + \alpha - 1)x^{\alpha - 1}.$$

Find the general solution of this equation.

(iv) Find the general solution of the homogeneous equation of the form

$$x^2 w'' + (a + b)xw' + a(b - 1)w = 0.$$

4. (i) Show that the series (1.15) with arbitrary α satisfies the non-homogeneous differential equation (1.1) with the right side

$$G_{\alpha}(x) = c_0 \frac{\lambda - \lambda_{-\alpha}}{(x - a)^{\alpha}}$$

if $\sigma(a) = 0$.

(ii) Show that the series of the form

$$u(x) = x^{-\alpha} \sum_{n=0}^{\infty} \frac{(\alpha)_n (\alpha - c + 1)_n}{(\alpha - a + 1)_n (\alpha - b + 1)_n} x^{-n} = x^{-\alpha}\, {}_3F_2 \left(\begin{array}{c} 1,\ \alpha,\ \alpha - c + 1 \\ \\ \alpha - a + 1,\ \alpha - b + 1 \end{array} ;\ x^{-1} \right), \quad |x| > 1,$$

satisfies the non-homogeneous equation

$$x(1 - x)u'' + [c - (a + b + 1)x]u' - abu = -(\alpha - a)(\alpha - b)x^{-\alpha}.$$

Find the general solution of this equation.

5. Prove that

$$J_\nu(x) = \frac{(x/2)^\nu}{\Gamma(\nu+1)}\, e^{-ix}\, {}_1F_1(\nu+1/2;\ 2\nu+1;\ 2ix)$$
$$= \frac{(x/2)^\nu}{\Gamma(\nu+1)}\, {}_0F_1\left(-;\ \nu+1;\ -x^2\Big/4\right).$$

6. By using the special case of (1.6), prove that elementary functions

$$u_0(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$
$$u_+(x) = \cos x = \sum_{n=0}^{\infty} (-1)^n\, \frac{x^{2n}}{(2n)!},$$
$$u_-(x) = \sin x = \sum_{n=0}^{\infty} (-1)^n\, \frac{x^{2n+1}}{(2n+1)!},$$

satisfy the equations $u'' = u$ and $u'' = -u,$ respectively.

7. Classify all solutions of the hypergeometric-type equation (1.1) depending on degrees of $\sigma(x)$ and $\tau(x)$.
8. Prove that functions (1.53) satisfy (under proper boundary conditions) the differentiation formula

$$\sigma(z)y_\nu'(z) = \frac{\kappa_\nu}{\tau_\nu'}\left[(\nu+1)\,\frac{C_\nu}{C_{\nu+1}}\, y_{\nu+1}(z) - \tau_\nu(z)y_\nu(z)\right]$$

and the recurrence relation

$$z\, y_\nu(z) = \alpha_\nu\, y_{\nu+1}(z) + \beta_\nu\, y_\nu(z) + \gamma_\nu\, y_{\nu-1}(z),$$

where

$$\alpha_\nu = -\frac{(\nu+1)\kappa_\nu C_\nu}{\tau_\nu'\tau_{\nu-1/2}' C_{\nu+1}},$$
$$\beta_\nu = \nu\,\frac{\tau_{\nu-1}(0)}{\tau_{\nu-1}'} - (\nu+1)\,\frac{\tau_\nu(0)}{\tau_\nu'}$$
$$= \frac{\left(\sigma''-\tau'\right)\tau(0) - \nu\left(2\tau' - (\nu-1)\sigma''\right)\sigma'(0)}{\tau_{\nu-1}'\tau_\nu'},$$

and

$$\gamma_\nu = -\frac{\tau'_{\nu-1}\sigma\left(-\tau_{\nu-1}(0)\big/\tau'_{\nu-1}\right)C_\nu}{\tau'_{\nu-1/2}C_{\nu-1}}.$$

Here $\kappa_\nu = \tau' + (\nu-1)\sigma''\big/2$ [138, 174, 175].

9. (i) Derive integral representations (1.60)–(1.62).

   (ii) Prove that the function

$$G(\alpha;\gamma;z) = \frac{z^{-\alpha}}{\Gamma(\alpha)}\int_0^\infty e^{-t}t^{\alpha-1}\left(1+\frac{t}{z}\right)^{\gamma-\alpha-1} dt, \quad |\arg z| < \pi, \quad \mathrm{Re}\,(\alpha) > 0,$$

   satisfies the confluent hypergeometric equation.

   (iii) Prove that

$$G(\alpha;\gamma;z) = \frac{\Gamma(1-\gamma)}{\Gamma(\alpha-\gamma+1)}\,{}_1F_1(\alpha;\gamma;z) + \frac{\Gamma(\gamma-1)}{\Gamma(\alpha)}\,z^{1-\gamma}\,{}_1F_1(1+\alpha-\gamma;2-\gamma;z).$$

10. (i) Transform the Bessel equation (1.63) to the form (1.64) and derive integral representations (1.65)–(1.66).

    (ii) Prove that

$$\frac{d}{dz}\left(z^{\pm\nu}J_\nu(z)\right) = \pm z^{\pm\nu}J_{\nu\mp1}(z)$$

$$J_{\nu-1}(z) + J_{\nu+1}(z) = \frac{2\nu}{z}\,J_\nu(z),$$

$$J_{\nu-1}(z) - J_{\nu+1}(z) = 2J'_\nu(z).$$

11. Prove that

$$y_n^{(m)}(x) = \frac{A_{mn}B_n}{\sigma^m(x)\rho(x)}\,\frac{d^{n-m}}{dx^{n-m}}\left[\sigma^n(x)\rho(x)\right],$$

where

$$A_{mn} = (-1)^m\prod_{k=0}^{m-1}(\lambda_n-\lambda_k) = \frac{n\,!}{(n-m)\,!}\prod_{k=0}^{m-1}\left(\tau' + \frac{n+k-1}{2}\,\sigma''\right), \; A_{0n} = 1.$$

12. Prove that the differentiation formula

$$\sigma(x)y_n'(x) = \frac{\lambda_n}{n\tau_n'}\left[\tau_n(x)y_n(x) - \frac{B_n}{B_{n+1}}\,y_{n+1}(x)\right]$$

is valid for classical orthogonal polynomials and the functions of the second kind.
13. Prove (1.83)–(1.84) for general orthogonal polynomials.
14. Find the following expressions

$$a_n = \frac{A_{nn}B_n}{n\,!} = B_n \prod_{k=0}^{n-1}\left(\tau' + \frac{n+k-1}{2}\,\sigma''\right), \quad a_0 = B_0;$$
$$\frac{b_n}{a_n} = n\,\frac{\tau_{n-1}(0)}{\tau_{n-1}'}$$

for the leading terms $y_n(x) = a_n x^n + b_n x^{n-1} + \ldots$ of the classical orthogonal polynomials.
15. Show that the following relation

$$d_n^2 = (-1)^n A_{nn} B_n^2 \int_a^b \sigma^n(x)\rho(x)\,dx$$

is valid for the squared norms (1.82) of the classical orthogonal polynomials.
16. Prove that

$$\frac{d}{dx}\,P_n^{(\alpha,\,\beta)}(x) = \frac{1}{2}\,(\alpha+\beta+n+1)\;P_{n-1}^{(\alpha+1,\,\beta+1)}(x),$$
$$\frac{d}{dx}\,L_n^{\alpha}(x) = -L_{n-1}^{\alpha+1}(x),$$
$$\frac{d}{dx}\,H_n(x) = 2n\;H_{n-1}(x).$$

17. Prove the symmetry relations

$$P_n^{(\alpha,\,\beta)}(-x) = (-1)^n\;P_n^{(\beta,\,\alpha)}(x),$$
$$H_n(-x) = (-1)^n\;H_n(x).$$

18. Prove that

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

$$\begin{aligned} L_n^{\alpha}(x) &= \frac{(\alpha+1)_n}{n\,!}\ {}_1F_1(-n, \alpha+1; x) \\ &= \frac{(-1)^n}{n\,!}\, x^n\, {}_2F_0\left(-n, -n-\alpha; -; -\frac{1}{x}\right); \end{aligned}$$

and

$$\begin{aligned} H_{2n}(x) &= (-1)^n\, 2^{2n} \left(\frac{1}{2}\right)_n {}_1F_1\left(-n;\ \frac{1}{2};\ x^2\right), \qquad n \geq 0, \\ H_{2n+1}(x) &= (-1)^n\, 2^{2n+1} \left(\frac{3}{2}\right)_n x\, {}_1F_1\left(-n;\ \frac{3}{2};\ x^2\right), \qquad n \geq 0, \\ H_n(x) &= (2x)^n\, {}_2F_0\left(-n/2,\ (1-n)/2; -; -x^{-2}\right), \qquad n \geq 0. \end{aligned}$$

19. Define moments for classical orthogonal polynomials by

$$C_m = \int_a^b (s-\xi)^m \rho(s)\, ds.$$

By using the identity from Exercise 1(i), prove that

$$\frac{C_{m+1}}{C_m} = -\frac{\tau(a) + m\sigma'(a)}{\tau' + m\sigma''/2},$$

if $\sigma(\xi) = 0$ and $\left(\sigma(s)\rho(s)(s-\xi)^n\right)\big|_a^b = 0$, $n = 0, 1, 2, \ldots$ [14].

20. Find relations between moments for the Jacobi, Laguerre and Hermite polynomials.
21. Let $\{\varphi_k\}_{k=0}^{\infty}$ and $\{\psi_k\}_{k=0}^{\infty}$ be sequences of polynomials $\varphi_k(x)$ and $\psi_k(x)$ of exact degree k. Prove that the orthogonal polynomials $p_n(x)$ for a given distribution $d\mu$ can be expressed as Gram determinants

$$p_n(x) = \begin{vmatrix} C_{0,0} & C_{0,1} & \ldots & C_{0,n} \\ C_{1,0} & \ldots & & C_{1,n} \\ \vdots & & & \\ C_{n-1,0} & \ldots & & C_{n-1,n} \\ \varphi_0(x) & \varphi_1(x) & \ldots & \varphi_n(x) \end{vmatrix}$$

with $C_{i,k} = \displaystyle\int \psi_i \varphi_k d\mu$.

Derive the explicit series representation for the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ from the Gram determinant [198, 199].

22. Prove (1.97)–(1.99).
23. Let $a, b < \infty$. Prove that

$$\frac{1}{2\pi i} \int\limits_C y_m(z) y_n(z) \rho(z) Q_0(z)\, dz$$
$$= -B_0 \int\limits_a^b y_m(s) y_n(s) \rho(s)\, ds,$$

where $Q_0(z)$ defined by (1.89) and C is a closed counter-clockwise contour which encloses the interval $[a, b]$ in the complex z-plane [80, 95].
24. Prove the complex orthogonality relation for the Jacobi polynomials,

$$\frac{1}{2\pi i}\, PV \int\limits_C P_n^{(\alpha,\,\beta)}(z)\, P_m^{(\alpha,\,\beta)}(z)\, \chi^{(\alpha,\,\beta)}(z)\, dz$$
$$= 2^{\alpha+\beta+1} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)\, n!\, \Gamma(\alpha+\beta+n+1)}\, \delta_{mn},$$

provided that $\alpha+\beta+1 \neq -1, -2, \ldots$. Here

$$\chi^{(\alpha,\,\beta)}(z) = -(1-z)^\alpha (1+z)^\beta\, Q_0^{(\alpha,\,\beta)}(z)$$
$$= \begin{cases} -2^{\alpha+\beta+1} \dfrac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)(1-z)}\, {}_2F_1\left(\begin{matrix} 1, \alpha+1 \\ \\ \alpha+\beta+2 \end{matrix} ;\ \dfrac{2}{(1-z)}\right), & |z-1| > 2, \\ 2^{\alpha+\beta+1} \dfrac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)(1-z)}\, {}_2F_1\left(\begin{matrix} 1, \beta+1 \\ \\ \alpha+\beta+2 \end{matrix} ;\ \dfrac{2}{(1+z)}\right), & |z+1| > 2, \end{cases}$$

and C is a counter-clockwise closed contour which encloses the interval $[-1, 1]$ in the complex z-plane[95, 153].
25. (i) Let $\alpha \neq -k$ and $k = 1, 2, \ldots$. Prove the complex orthogonality relation for the Laguerre polynomials,

$$\lim_{R\to\infty} \frac{1}{2\pi i} \int\limits_{C(R)} L_n^\alpha(z)\, L_m^\alpha(z)\, \chi^\alpha(z)\, dz = \frac{(\alpha+1)_n}{n\,!}\, \delta_{mn},$$

where

$$\chi^{\alpha}(z) = -\frac{z^{\alpha} e^{-z}}{\Gamma(\alpha+1)}\, Q_0^{\alpha}(z) = G(1; 1-\alpha; -z)$$

and $C(R) = \lim_{\epsilon\to 0} C_{\epsilon}(R)$ is a closed contour.

(ii) Let $\alpha = -k$ and $k = 1, 2, \ldots$. Prove that a finite number of the Laguerre polynomials $\left\{L_n^{-k}(x)\right\}_{n=0}^{k-1}$ satisfies the orthogonality relation

$$\int_{-\infty}^{\infty} L_n^{-k}(x)\, L_m^{-k}(x)\, d\mu(x) = \frac{(1-k)_n}{n\,!}\,\delta_{mn},$$

where the measure may be expressed as a real distribution with support at $x = 0$:

$$d\mu(x) = \left(1 + \frac{d}{dx}\right)^{k-1} \delta(x)\, dx = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell}\, \delta^{(\ell)}(x)\, dx.$$

(Ismail et al. [95].)

26. Prove the complex orthogonality relation for the Hermite polynomials,

$$\lim_{R_{1,2}\to\mp\infty} \frac{1}{\pi i}\, PV \int_{C(R_1,R_2)} H_n(z)\, H_m(z)\, \chi(z)\, dz = 2^n\, n!\, \delta_{mn},$$

where

$$\chi(z) = -\frac{e^{-z^2}}{2\sqrt{\pi}}\, Q_0(z) = \mp i H_{-1}(\mp i z) \quad (\mathrm{Im}(z) > 0 \quad \text{or} \quad \mathrm{Im}(z) < 0)$$

and $C\,(R_1, R_2) = \lim_{\epsilon\to 0} C_{\epsilon}^{+}\,(R_1, R_2) \bigcup C_{\epsilon}^{-}\,(R_1, R_2)$ is the counter-clockwise contour (details are left to the reader).

2 Some Problems of Nonrelativistic and Relativistic Quantum Mechanics

In this section, we give a short summary of Nikiforov and Uvarov's approach to special functions of mathematical physics and their applications in quantum mechanics [138].

2.1 Generalized Equation of Hypergeometric Type

The second order differential equation of the form

$$u'' + \frac{\widetilde{\tau}(z)}{\sigma(z)}\, u' + \frac{\widetilde{\sigma}(z)}{\sigma^2(z)}\, u = 0, \tag{2.1}$$

where $\sigma(z)$ and $\widetilde{\sigma}(z)$ are polynomials of degree at most 2 and $\widetilde{\tau}(z)$ is a polynomial of degree at most 1 of a complex variable z, is called the *generalized equation of hypergeometric type*. By the substitution $u = \varphi(z)\, y$ Eq. (2.1) can be reduced to the *equation of hypergeometric type*

$$\sigma(z)\, y'' + \tau(z)\, y' + \lambda y = 0, \tag{2.2}$$

where $\tau(z)$ is a polynomial of degree at most 1, and λ is a constant. The factor $\varphi(z)$ here satisfies

$$\frac{\varphi'}{\varphi} = \frac{\pi(z)}{\sigma(z)}, \tag{2.3}$$

where $\pi(z)$ is a polynomial of degree at most 1 given by a quadratic formula

$$\pi(z) = \frac{\sigma' - \widetilde{\tau}}{2} \pm \sqrt{\left(\frac{\sigma' - \widetilde{\tau}}{2}\right)^2 - \widetilde{\sigma} + k\sigma} \tag{2.4}$$

and constant k is determined by the condition that the discriminant of the quadratic polynomial under the square root sign is zero. Then $\tau(z)$ and λ are determined by

$$\tau(z) = \widetilde{\tau}(z) + 2\pi(z), \qquad \lambda = k + \pi'(z). \tag{2.5}$$

Two exceptions are [138]:

1. If $\sigma(z)$ has a double root, $\sigma(z) = (z-a)^2$, the original equation can be carried out into a generalized equation of hypergeometric type with $\sigma(s) = s$, by a substitution $s = (z-a)^{-1}$.
2. If $\sigma(z) = 1$ and $(\widetilde{\tau}(z)/2)^2 - \widetilde{\sigma}(z)$ is a polynomial of degree 1, the substitution $\pi(z) = -\widetilde{\tau}(z)/2$ reduces the original equation to the form

$$y'' + (az + b)\, y = 0. \tag{2.6}$$

Solutions of (2.1)–(2.2) are known as special functions of hypergeometric type; they include classical orthogonal polynomials, hypergeometric and confluent hypergeometric functions, Hermite functions, Bessel functions and spherical harmonics. These functions are often called special functions of mathematical physics.

2.2 Classical Orthogonal Polynomials and Eigenvalue Problems

The following theorem is a useful tool for finding of the square integrable solutions of basic problems in quantum mechanics [138].

Theorem 2.1 *Let $y = y(x)$ be a solution of the equation of hypergeometric type (2.2) and let $\rho(x)$, a solution of the Pearson equation*

$$(\sigma(x)\rho(x))' = \tau(x)\rho(x) \tag{2.7}$$

be bounded on the interval (a, b) and satisfy the boundary conditions

$$\sigma(x)\rho(x)x^k\Big|_{x=a,b} = 0, \qquad k = 0, 1, 2, \ \ldots \tag{2.8}$$

Then nontrivial solutions of (2.2) such that $y(x)\sqrt{\rho(x)}$ is bounded and of integrable square on (a, b) exist only for the eigenvalues given by

$$\lambda = \lambda_n = -n\tau' - \frac{1}{2}n(n-1)\sigma'' \qquad (n = 0, 1, 2, \ \ldots). \tag{2.9}$$

They are the corresponding classical orthogonal polynomials on (a, b) and can be found by the Rodrigues-type formula

$$y_n(z) = \frac{B_n}{\rho(z)}\left(\sigma^n(z)\rho(z)\right)^{(n)}. \tag{2.10}$$

The proof is given in [138].

2.2.1 Example: Linear Harmonic Oscillator

Let us consider one-dimensional stationary Schrödinger equation for the harmonic oscillator,

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2x^2\psi = E\psi \tag{2.11}$$

with the orthonormal real-valued wave function

$$\int_{-\infty}^{\infty}\psi^2(x)\ dx = 1. \tag{2.12}$$

Introducing dimensionless variables

$$\psi(x) = u(\xi), \qquad x = \xi\sqrt{\frac{\hbar}{m\omega}}, \quad E = \hbar\omega\varepsilon \tag{2.13}$$

one gets

$$u'' + \left(2\varepsilon - \xi^2\right) u = 0. \tag{2.14}$$

Here, $\sigma(\xi) = 1$, $\widetilde{\tau}(\xi) = 0$, and $\widetilde{\sigma}(\xi) = 2\varepsilon - \xi^2$. Therefore,

$$\pi(\xi) = \pm\sqrt{k - 2\varepsilon + \xi^2} = \pm\xi, \qquad k = 2\varepsilon. \tag{2.15}$$

We pick $\pi = -\xi$, which gives a negative derivative for

$$\tau(\xi) = \widetilde{\tau}(\xi) + 2\pi(\xi) = -2\xi. \tag{2.16}$$

Then

$$\frac{\varphi'}{\varphi} = \frac{\pi(\xi)}{\sigma(\xi)} = -\xi, \qquad \varphi(\xi) = e^{-\xi^2/2} \tag{2.17}$$

and $\lambda = 2\varepsilon - 1$, $\rho(\xi) = e^{-\xi^2}$. The energy levels are $\varepsilon = \varepsilon_n = n + 1/2$, $(n = 0, 1, 2, \dots)$ from (2.9). The eigenfunctions,

$$y_n(\xi) = B_n e^{\xi^2} \frac{d^n}{d\xi^n}\left(e^{-\xi^2}\right), \tag{2.18}$$

are, up a normalization, the Hermite polynomials.

As a result, the orthonormal wave functions are given by Schrödinger [158]

$$\psi(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) H_n\left(x\sqrt{\frac{m\omega}{\hbar}}\right), \tag{2.19}$$

corresponding to the discrete energy levels

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \qquad (n = 0, 1, 2, \dots), \tag{2.20}$$

in Gaussian units. (More general, “missing”, solutions of the time-dependent Schrödinger equation will be discussed in section 4 [121].)

2.3 Method of Separation of Variables and Its Extension

In this part, we give an extension of the method of separation of variables, that is used in theoretical and mathematical physics for solving partial differential equations, from a single equation to a system of partial differential equations which we call Dirac-type system [183].

2.3.1 Method of Separation of Variables

We follow [138] and give an extension for suitable Dirac's systems. The method of separation of the variables helps to find particular solutions of equation

$$\mathcal{L}u = 0 \tag{2.21}$$

if the operator $\mathcal{L}$ can be represented in the form

$$\mathcal{L} = \mathcal{M}_1\mathcal{N}_1 + \mathcal{M}_2\mathcal{N}_2. \tag{2.22}$$

Here the operators $\mathcal{M}_1$ and $\mathcal{M}_2$ act only on one subset of the variables, and the operators $\mathcal{N}_1$ and $\mathcal{N}_2$ act on the others; a product of operators $\mathcal{M}_i\mathcal{N}_k$ means the result of applying them successively $(\mathcal{M}_i\mathcal{N}_k)\, u = \mathcal{M}_i\, (\mathcal{N}_k u)$ with $i, k = 1, 2$; it is assumed that the operators $\mathcal{M}_i$ and $\mathcal{N}_i$ are linear operators.

We look for solutions of Eq. (2.21) in the form

$$u = f\ g, \tag{2.23}$$

where the first unknown function f depends only on the first set of variables and the second function g depends on the others. Since

$$\begin{aligned}\mathcal{M}_i\mathcal{N}_k u &= (\mathcal{M}_i\mathcal{N}_k)\,(f\ g) = \mathcal{M}_i\,(\mathcal{N}_k\,(f\ g))\\ &= \mathcal{M}_i\,(f\ \ (\mathcal{N}_k g)) = (\mathcal{M}_i f)\ \ (\mathcal{N}_k g)\end{aligned}$$

the equation $\mathcal{L}u = 0$ can be rewritten in the form

$$\frac{\mathcal{M}_1 f}{\mathcal{M}_2 f} = -\frac{\mathcal{N}_2 g}{\mathcal{N}_1 g},$$

where the left hand side is independent of the second group of the variables and the right hand side is independent of the first ones. Thus, we must have

$$\frac{\mathcal{M}_1 f}{\mathcal{M}_2 f} = -\frac{\mathcal{N}_2 g}{\mathcal{N}_1 g} = \lambda,$$

where λ is a constant, and one obtains equations

$$\mathcal{M}_1 f = \lambda \mathcal{M}_2 f, \qquad \mathcal{N}_2 g = -\lambda \mathcal{N}_1 g \tag{2.24}$$

each containing functions of only some of the variables. Since $\mathcal{L}$ is linear, a linear combination of solutions,

$$u = \sum_k c_k \ f_k g_k \tag{2.25}$$

with some constants c_k, corresponding to all admissible values of $\lambda = \lambda_k$, will be a solution of the original equation (2.21). Under certain condition of the completeness of the constructed set of particular solutions, every solution of (2.21) can be represented in the form (2.25). The method of separation of variables is very useful in theoretical and mathematical physics and partial differential equations—including solutions of the nonrelativistic Schrödinger equation—but it should be modified in the case of the Dirac equation.

Example The stationary Schrödinger equation in the central field with the potential energy $U(r)$ is given by

$$\Delta\psi + \frac{2m}{\hbar^2}\left(E - U(r)\right)\psi = 0. \tag{2.26}$$

The Laplace operator in the spherical coordinates $r, \ \theta, \ \varphi$ has the form [138, 139]

$$\Delta = \Delta_r + \frac{1}{r^2}\Delta_\omega \tag{2.27}$$

with

$$\Delta_r = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right), \quad \Delta_\omega = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{2.28}$$

Thus

$$\mathcal{M}_1 = \Delta_r + \frac{2m}{\hbar^2}\left(E - U(r)\right), \quad \mathcal{M}_2 = \frac{1}{r^2} \tag{2.29}$$

$$\mathcal{N}_1 = \mathrm{id} = I, \qquad \mathcal{N}_2 = \Delta_\omega \tag{2.30}$$

and separation of the variables $\psi = R(r)\, Y(\theta, \varphi)$ gives

$$\Delta_\omega Y(\theta, \varphi) + \lambda Y(\theta, \varphi) = 0, \tag{2.31}$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right) + \left(\frac{2m}{\hbar^2}\left(E - U(r)\right) - \frac{\lambda}{r^2}\right)R(r) = 0. \tag{2.32}$$

Bounded single-valued solutions of Eq. (2.31) on the sphere S^2 exist only when $\lambda = l\,(l+1)$ with $l = 0, 1, 2, \ \dots$. They are the spherical harmonics $Y = Y_{lm}\,(\theta, \varphi)$. (See [138, 139, 189, 190, 196, 197] for more details.)

2.3.2 Dirac-Type Systems

Let us consider the system of two equations [183]

$$\mathcal{P}u = \alpha v, \tag{2.33}$$

$$\mathcal{P}v = \beta u, \tag{2.34}$$

where $u = u\,(\mathbf{x})$ and $v = v\,(\mathbf{x})$ are some unknown (complex) vector valued functions on $\mathbb{R}^n$ (or $\mathbb{C}^n$). Here operator $\mathcal{P}$ has the following structure

$$\mathcal{P} = \mathcal{N}\,(\mathbf{n})\,(\mathcal{D}_1\,(r)\,\mathcal{L}_1\,(\mathbf{n}) + \mathcal{D}_2\,(r)\,\mathcal{L}_2\,(\mathbf{n}))\,, \quad \mathcal{N}^2\,(\mathbf{n}) = \mathrm{id} = I, \tag{2.35}$$

where $\mathcal{D}_i = \mathcal{D}_i\,(r)\,,\ \mathcal{L}_k = \mathcal{L}_k\,(\mathbf{n})$ and $\mathcal{N} = \mathcal{N}\,(\mathbf{n})$ are linear operators acting with respect to two different subsets of variables, say “radial” r and “angular” $\mathbf{n}$ variables, respectively (in the case of the hyperspherical coordinates in $\mathbb{R}^n$ [139] one gets $\mathbf{x} = r\mathbf{n}$ and $\mathbf{n}^2 = 1$, which justifies our terminology). The following algebraic properties hold

$$[\mathcal{D}_i, \mathcal{L}_k] = [\mathcal{D}_i, \mathcal{N}] = 0, \tag{2.36}$$

$$[\mathcal{N}, \mathcal{L}_1] = [\mathcal{L}_1, \mathcal{L}_2] = 0, \tag{2.37}$$

$$\mathcal{N}\mathcal{L}_2 + \mathcal{L}_2\mathcal{N} = \gamma\mathcal{N}, \tag{2.38}$$

where $[\mathcal{A},\ \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$ is the commutator and γ is some constant.

We look for solutions of (2.33) and (2.34) in the form

$$u = \mathcal{Y}\,(\mathbf{n})\,R\,(r)\,, \tag{2.39}$$

$$v = (\mathcal{N}\mathcal{Y}\,(\mathbf{n}))\,S\,(r)\,, \tag{2.40}$$

where $\mathcal{Y}$ is the common eigenfunction of commuting operators $\mathcal{L}_1$ and $\mathcal{L}_2$:

$$\mathcal{L}_1\mathcal{Y} = \kappa_1\mathcal{Y}, \qquad \mathcal{L}_2\mathcal{Y} = \kappa_2\mathcal{Y}. \tag{2.41}$$

If $w = w\,(\mathbf{x}) = F\,(\mathbf{n})\,G\,(r)\,,$ we define the action of the “radial” and “angular” operators in (2.35) as follows

$$\mathcal{L}_i w = (\mathcal{L}_i F)\,G, \quad \mathcal{N}w = (\mathcal{N}F)\,G, \quad \mathcal{D}_k w = F\,(\mathcal{D}_k G)\,. \tag{2.42}$$

The Ansatz (2.39) and (2.40) results in two equation for our "radial" functions R and S :

$$\kappa_1 \mathcal{D}_1 R + \kappa_2 \mathcal{D}_2 R = \alpha S, \tag{2.43}$$

$$\kappa_1 \mathcal{D}_1 S + (\gamma - \kappa_2)\, \mathcal{D}_2 S = \beta R. \tag{2.44}$$

Indeed, in view of (2.33) and (2.34) and (2.39)–(2.41) one gets

$$\begin{aligned} \mathcal{P}u &= \mathcal{N}\,(\mathcal{D}_1\mathcal{L}_1 + \mathcal{D}_2\mathcal{L}_2)\,\mathcal{Y}R \\ &= \mathcal{N}\,((\mathcal{L}_1\mathcal{Y})\,(\mathcal{D}_1 R) + (\mathcal{L}_2\mathcal{Y})\,(\mathcal{D}_2 R)) \\ &= (\mathcal{N}\mathcal{Y})\,(\kappa_1\mathcal{D}_1 R + \kappa_2\mathcal{D}_2 R) \\ &= \alpha\,(\mathcal{N}\mathcal{Y})\,S = \alpha v, \end{aligned}$$

which gives (2.43). In a similar fashion, with the aid of (2.38)

$$\begin{aligned} \mathcal{P}v &= \mathcal{N}\,(\mathcal{D}_1\mathcal{L}_1 + \mathcal{D}_2\mathcal{L}_2)\,(\mathcal{N}\mathcal{Y})\,S \\ &= \mathcal{N}\,((\mathcal{L}_1\mathcal{N}\mathcal{Y})\,(\mathcal{D}_1 S) + (\mathcal{L}_2\mathcal{N}\mathcal{Y})\,(\mathcal{D}_2 S)) \\ &= (\mathcal{N}\mathcal{L}_1\mathcal{N}\mathcal{Y})\,(\mathcal{D}_1 S) + (\mathcal{N}\mathcal{L}_2\mathcal{N}\mathcal{Y})\,(\mathcal{D}_2 S) \\ &= \left(\mathcal{N}^2\mathcal{L}_1\mathcal{Y}\right)(\mathcal{D}_1 S) + \left((\gamma - \mathcal{L}_2)\,\mathcal{N}^2\mathcal{Y}\right)(\mathcal{D}_2 S) \\ &= \mathcal{Y}\,(\kappa_1\mathcal{D}_1 S + (\gamma - \kappa_2)\,\mathcal{D}_2 S) = \beta\mathcal{Y}R = \beta u, \end{aligned}$$

which results in the second equation (2.44) and our proof is complete.

Example The original Dirac system (3.44) and (3.45) below has

$$\psi = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix}$$

and

$$\mathcal{P} = c\boldsymbol{\sigma}\boldsymbol{p} = \hbar c\,(\boldsymbol{\sigma}\mathbf{n})\left(\frac{1}{i}\frac{\partial}{\partial r} + \frac{i}{r}\boldsymbol{\sigma}\boldsymbol{l}\right). \tag{2.45}$$

Here

$$\mathcal{N} = \boldsymbol{\sigma}\mathbf{n}, \quad \mathcal{D}_1 = \frac{\hbar c}{i}\frac{\partial}{\partial r}, \quad \mathcal{L}_1 = \mathrm{id} = I, \quad \mathcal{D}_2 = \frac{i\hbar c}{r}, \quad \mathcal{L}_2 = \boldsymbol{\sigma}\boldsymbol{l} \tag{2.46}$$

and

$$\alpha\,(r) = E + mc^2 - U\,(r)\,, \quad \beta\,(r) = E - mc^2 - U\,(r)\,, \quad \gamma = -2 \tag{2.47}$$

by (3.50). Moreover, $\kappa_1 = 1$, $\kappa_2 = -(1+\kappa)$ and we use $R = F(r)$, $S = -iG(r)$. The Ansazts (3.51) and (3.52) give the familiar radial equations (3.53) and (3.54).

2.4 Nonrelativistic Coulomb Problem

As an example, we give a detailed solution of the nonrelativistic Schrödinger equation for Coulomb potential [138].

2.4.1 Radial Equation

In view of identity

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = \frac{1}{r}\frac{d^2}{dr^2}(rR),$$

the substitution $F(r) = rR(r)$ results in the standard radial equation

$$F'' + \left[\frac{2m_e}{\hbar^2}(E - U(r)) - \frac{l(l+1)}{r^2}\right]F = 0, \quad U(r) = -\frac{Ze^2}{r} \tag{2.48}$$

for the nonrelativistic Coulomb problem in spherical coordinates. In dimensional units,

$$F(r) = u(x), \quad x = \frac{r}{a_0}, \quad \varepsilon = \frac{E}{E_0} \quad \left(a_0 = \frac{\hbar^2}{m_e e^2}, \quad E_0 = \frac{e^2}{a_0}\right) \tag{2.49}$$

the radial equation is a generalized equation of hypergeometric type,

$$u'' + \left[2\left(\varepsilon + \frac{Z}{x}\right) - \frac{l(l+1)}{x^2}\right]u = 0, \tag{2.50}$$

where

$$\sigma(x) = x, \qquad \widetilde{\tau}(x) = 0, \qquad \widetilde{\sigma}(x) = 2\varepsilon x^2 + 2Zx - l(l+1). \tag{2.51}$$

Thus one can utilize Nikiforov and Uvarov's approach in order to determine the corresponding wave functions and discrete energy levels.

2.4.2 Quantization

We transform (2.50) to the equation of hypergeometric type,

$$\sigma(x)\, y'' + \tau(x)\, y' + \lambda y = 0, \tag{2.52}$$

by means of the substitution $u(x) = \varphi(x)\, y(x)$. Here,

$$\frac{\varphi'}{\varphi} = \frac{\pi(x)}{\sigma(x)}, \quad \pi(x) = \frac{\sigma' - \widetilde{\tau}}{2} \pm \sqrt{\left(\frac{\sigma' - \widetilde{\tau}}{2}\right)^2 - \widetilde{\sigma} + k\sigma}, \tag{2.53}$$

and the linear function $\pi(x)$ takes the form

$$\pi(x) = \frac{1}{2} \pm \sqrt{\frac{1}{4} - 2\varepsilon x^2 - 2x + l(l+1) + kx}, \tag{2.54}$$

or

$$\pi(x) = \frac{1}{2} \pm \begin{cases} \sqrt{-2\varepsilon}\, x + l + 1/2, & k = 2Z + (2l+1)\sqrt{-2\varepsilon} \\ \sqrt{-2\varepsilon}\, x - l - 1/2, & k = 2Z - (2l+1)\sqrt{-2\varepsilon} \end{cases} \tag{2.55}$$

where we should choose the case when the linear function $\tau = \widetilde{\tau} + 2\pi$ will have a negative derivative and a zero on $(0, +\infty)$:

$$\tau(x) = 2\left(l + 1 - x\sqrt{-2\varepsilon}\right).$$

This choice corresponds to

$$\pi(x) = l + 1 - x\sqrt{-2\varepsilon}, \quad \varphi(x) = x^{l+1} \exp\left(-x\sqrt{-2\varepsilon}\right)$$

and

$$\lambda = k + \pi' = 2\left[Z - (l+1)\sqrt{-2\varepsilon}\,\right].$$

The energy values are determined by the equation

$$\lambda + n_r \tau' + \frac{1}{2} n_r (n_r - 1)\, \sigma'' = 0 \qquad (n_r = 0, 1, 2, \ldots)$$

resulting in

$$\varepsilon = \frac{E}{E_0} = -\frac{Z^2}{2(n_r + l + 1)^2}, \qquad E_0 = \frac{e^2}{a_0}. \tag{2.56}$$

Here, $n = n_r + l + 1$ is known as the principal quantum number.

In order to use the Rodrigues formula, one finds

$$\frac{\rho'}{\rho} = \frac{\tau - \sigma'}{\sigma} = \frac{2l+1}{x} - \frac{2Z}{n},$$

or

$$\rho(x) = x^{2l+1} \exp\left(-\frac{2Z}{n}x\right), \qquad x = \frac{r}{a_0}.$$

Therefore,

$$y_{n_r}(x) = \frac{B_{n_r}}{x^{2l+1}e^{-\eta}} \frac{d^{n_r}}{dx^{n_r}} \left(x^{n_r+2l+1}e^{-\eta}\right) = L_{n_r}^{2l+1}(\eta), \tag{2.57}$$

where

$$\eta = \frac{2Z}{n}x = \frac{2Z}{n}\left(\frac{r}{a_0}\right),$$

and, up to a constant,

$$F(r) = rR(r) = C_{nl}\ \eta^{l+1}e^{-\eta/2}L_{n_r}^{2l+1}(\eta). \tag{2.58}$$

In view of the normalization condition

$$1 = \int_0^\infty F^2\, dr = C_{nl}^2 \left(\frac{na_0}{2Z}\right) \int_0^\infty \eta^{2l+2}e^{-\eta}\left[L_{n_r}^{2l+1}(\eta)\right]^2\, d\eta,$$

the three-term recurrence relation

$$\eta L_n^\alpha = -(n+1)L_{n+1}^\alpha + (\alpha + 2n + 1)L_n^\alpha - (\alpha + n)L_{n-1}^\alpha,$$

and the orthogonality property of the Laguerre polynomials, one gets

$$C_{nl}^2 = \frac{Z}{a_0 n^2} \frac{(n-l-1)!}{(n+l)!}. \tag{2.59}$$

(We can also utilize a special case of the integral (A.5); see (A.6).)

2.4.3 Summary: Wave Functions and Energy Levels

The nonrelativistic Coulomb wave functions obtained by the method of separation of the variables in spherical coordinates, see above, are

$$\psi = \psi_{nlm}(\mathbf{r}) = R_{nl}(r)\ Y_{lm}(\theta, \varphi), \tag{2.60}$$

where $Y_{lm}(\theta, \varphi)$ are the spherical harmonics, the radial functions $R_{nl}(r)$ are given in terms the Laguerre polynomials [28, 113, 138, 156]

$$R(r) = R_{nl}(r) = \frac{2}{n^2}\left(\frac{Z}{a_0}\right)^{3/2}\sqrt{\frac{(n-l-1)!}{(n+l)!}}\, e^{-\eta/2}\eta^l\, L_{n-l-1}^{2l+1}(\eta) \tag{2.61}$$

with

$$\eta = \frac{2Z}{n}\left(\frac{r}{a_0}\right), \qquad a_0 = \frac{\hbar^2}{m_e e^2} \tag{2.62}$$

and the normalization is

$$\int_0^\infty R_{nl}^2(r)\, r^2\, dr = 1. \tag{2.63}$$

Here $n = 1, 2, 3, \ \ldots$ is the principal quantum number of the hydrogenlike atom in the nonrelativistic Schrödinger theory; $l = 0, 1, \ \ldots\ , n-1$ and $m = -l, -l+1, \ \ldots\ , l-1, l$ are the quantum numbers of the angular momentum and its projection on the z-axis, respectively. The corresponding discrete energy levels in the cgs units are given by Bohr's formula

$$E = E_n = -\frac{m_e Z^2 e^4}{2\hbar^2 n^2}, \tag{2.64}$$

where $n = 1, 2, 3, \ \ldots$ is the principal quantum number; they do not depend on the quantum number of the orbital angular momenta l due to a "hidden" $SO(4)$-symmetry of the Hamiltonian of the nonrelativistic hydrogen atom; see, for example, [61, 113, 173] and references therein and the original paper by Fock [75] and Bargmann [19].

2.5 Matrix Elements

The following integral evaluations are significant in applications.

2.5.1 General Results

Here we evaluate the mean values

$$\langle r^p \rangle = \frac{\displaystyle\int_{\mathbb{R}^3} |\psi_{nlm}(\mathbf{r})|^2\ r^p\ dv}{\displaystyle\int_{\mathbb{R}^3} |\psi_{nlm}(\mathbf{r})|^2\ dv} = \frac{\displaystyle\int_0^\infty R_{nl}^2(r)\, r^{p+2}\, dr}{\displaystyle\int_0^\infty R_{nl}^2(r)\, r^2\, dr}, \qquad dv = r^2 dr d\omega \tag{2.65}$$

in terms of Chebyshev polynomials of a discrete variable $t_k(x, N) = h_k^{(0,\,0)}(x, N)$ [185, 186] and [187]. We have used the orthogonality relation for the spherical harmonics [138, 189],

$$\int_{S^2} Y_{lm}^*(\theta, \varphi)\, Y_{l'm'}(\theta, \varphi)\; d\omega = \delta_{ll'}\delta_{mm'} \tag{2.66}$$

with $d\omega = \sin\theta\; d\theta d\varphi$ and $0 \le \theta \le \pi,\, 0 \le \varphi \le 2\pi$. The end results are

$$\left\langle r^{k-1}\right\rangle = \frac{1}{2n}\left(\frac{na_0}{2Z}\right)^{k-1} t_k\left(n-l-1, -2l-1\right), \tag{2.67}$$

when $k = 0, 1, 2, \ldots$ and

$$\left\langle \frac{1}{r^{k+2}}\right\rangle = \frac{1}{2n}\left(\frac{2Z}{na_0}\right)^{k+2} t_k\left(n-l-1, -2l-1\right), \tag{2.68}$$

when $k = 0, 1, \ldots,\ 2l$.

Although a connection of the mean values (2.65) with a family of the hypergeometric polynomials was established by Pasternack [141], the relation with the Chebyshev polynomials of a discrete variable was missing. This is a curious but fruitful case of a "mistaken identity" in the theory of classical orthogonal polynomials. The so-called Hahn polynomials of a discrete variable were originally introduced by Chebyshev [187], they have a discrete measure on the finite equidistant set of points. Bateman, in a series of papers [20–24], and Hardy [89] were the first who studied a continuous measure on the entire real line for the simplest special case of these polynomials of Chebyshev. Pasternack gave an extension of the results of Bateman to a one parameter family of the continuous orthogonal polynomials [142]. After investigation of these Bateman–Pasternack polynomials in the fifties by several authors; see [35, 41, 188, 202] and [42]; Askey and Wilson [11] introduced what nowadays known as the symmetric continuous Hahn polynomials, they have two free parameters—but one parameter had been yet missing! Finally, Suslov [172], Atakishiyev and Suslov [13] and Askey [10] have introduced the continuous Hahn polynomials in their full generality in the mid of eighties. More details on the discovery the continuous Hahn polynomials and their properties are given in [104] among other things.

Indeed, in view of the normalization condition of the Coulomb wave functions (2.63) one gets

$$\begin{aligned}\left\langle r^p\right\rangle &= \int_0^\infty R_{nl}^2(r)\, r^{p+2}\, dr \qquad (2.69)\\ &= \frac{4}{n^4}\left(\frac{na_0}{2Z}\right)^{p+3}\left(\frac{Z}{a_0}\right)^3 \frac{(n-l-1)!}{(n+l)!}\int_0^\infty e^{-\eta}\eta^{p+2l+2}\left(L_{n-l-1}^{2l+1}(\eta)\right)^2\, d\eta\end{aligned}$$

and the last integral can be evaluated with the help of (A.10) or (A.11) giving rise to (2.67) and (2.68), respectively.

A convenient "inversion" relation for the Coulomb matrix elements,

$$\left\langle \frac{1}{r^{k+2}} \right\rangle = \left(\frac{2Z}{na_0} \right)^{2k+1} \frac{(2l-k)!}{(2l+k+1)!} \left\langle r^{k-1} \right\rangle \tag{2.70}$$

with $0 \leq k \leq 2l$, follows directly from (2.67) and (2.68). This relation is contained in an implicit form in [113], it was given explicitly in [141].

2.5.2 Special Cases

The explicit expression (2.67) for the matrix elements $\langle r^p \rangle$ and the familiar three-term recurrence relation for Hahn polynomials $h_k^{(\alpha,\ \beta)}(x, N)$ [138, 139],

$$xh_k^{(\alpha,\ \beta)}(x, N) = \alpha_k h_{k+1}^{(\alpha,\ \beta)}(x, N) + \beta_k h_k^{(\alpha,\ \beta)}(x, N) + \gamma_k h_{k-1}^{(\alpha,\ \beta)}(x, N) \tag{2.71}$$

with

$$\alpha_k = \frac{(n+1)(\alpha+\beta+n+1)}{(\alpha+\beta+2n+1)(\alpha+\beta+2n+2)},$$
$$\beta_k = \frac{\alpha-\beta+2N-2}{4} + \frac{\left(\beta^2-\alpha^2\right)(\alpha+\beta+2N)}{4(\alpha+\beta+2n)(\alpha+\beta+2n+2)},$$
$$\gamma_k = \frac{(\alpha+n)(\beta+n)(\alpha+\beta+N+n)(N-n)}{(\alpha+\beta+2n)(\alpha+\beta+2n+1)},$$

imply the following three-term recurrence relation for the matrix elements (2.65):

$$\begin{aligned} \left\langle r^k \right\rangle = {} & \frac{2n(2k+1)}{k+1} \left(\frac{na_0}{2Z} \right) \left\langle r^{k-1} \right\rangle \\ & - \frac{k\left((2l+1)^2-k^2\right)}{k+1} \left(\frac{na_0}{2Z} \right)^2 \left\langle r^{k-2} \right\rangle \end{aligned} \tag{2.72}$$

with the "initial conditions"

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a_0 n^2}, \qquad \langle 1 \rangle = 1 \tag{2.73}$$

which is convenient for evaluation of the mean values $\left\langle r^k \right\rangle$ for $k \geq 1$ [141]. The inversion relation (2.70) can be used then for all possible negative values of k. One can easily find the following matrix elements

$$\langle r \rangle = \frac{a_0}{2Z}\left(3n^2 - l\,(l+1)\right), \tag{2.74}$$

$$\left\langle r^2 \right\rangle = 2\left(\frac{na_0}{2Z}\right)^2 \left(5n^2 + 1 - 3l\,(l+1)\right), \tag{2.75}$$

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a_0 n^2}, \tag{2.76}$$

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{2Z^2}{a_0^2 n^3\,(2l+1)}, \tag{2.77}$$

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3 n^3\,(l+1)\,(l+1/2)\,l}, \tag{2.78}$$

$$\left\langle \frac{1}{r^4} \right\rangle = \frac{Z^4\left(3n^2 - l\,(l+1)\right)}{2a_0^4 n^5\,(l+3/2)\,(l+1)\,(l+1/2)\,l\,(l-1/2)}, \tag{2.79}$$

which are important in many calculations in quantum mechanics and quantum electrodynamics [2, 25, 28, 96, 194]; see [28] for more examples.

Equations (2.64) and (2.76) show that the total energy of the electron in the hydrogenlike atom is equal to half the average potential energy:

$$\langle U \rangle = -Ze^2 \left\langle \frac{1}{r} \right\rangle = -\frac{Z^2 e^2}{a_0 n^2} = 2E. \tag{2.80}$$

This is the statement of so-called virial theorem in nonrelativistic quantum mechanics; see, for example, [28, p. 165] and [113].

The average distance between the electron and the nucleus $\overline{r} = \langle r \rangle$ is given by (2.74). The mean square deviation of the nucleus-electron separation is

$$\overline{(r-\overline{r})^2} = \overline{r^2} - \overline{r}^{\,2} = \left(\frac{a_0}{2Z}\right)^2 \left(n^2\left(n^2+2\right) - l^2\,(l+1)^2\right). \tag{2.81}$$

The quantum mechanical analogue to Bohr orbits of large eccentricity corresponds to large values of this number (small l).

3 Relativistic Coulomb Problem

A basic problem in quantum theory of the atom is the problem of finding solutions of the nonrelativistic Schrödinger and relativistic Dirac wave equations for the motion of electron in a central attractive force field. The only atom for which these equations can be solved explicitly is the simplest hydrogen atom, or, in general, the one electron hydrogenlike ionized atom with the charge of the nucleus Ze; this is a classical problem in quantum mechanics which is studied in great detail; see, for example, [2, 25, 28, 50, 71, 91, 96, 113, 133, 156] and references therein. Comparison of the results of theoretical calculations with experimental data provides accurate tests of the validity of the quantum electrodynamics [28, 96, 194]. Explicit analytical solutions for hydrogenlike atoms can be useful as the starting point in approximate calculations of more sophisticated quantum-mechanical systems.

However, the relativistic Coulomb wave functions are not well known for a "general audience" and this discussion might be useful for the reader who is not an expert in theoretical physics; this section is written for those who study quantum mechanics and would like to see more details than in the classical textbooks [2, 25, 28, 113]; it is motivated by a course in quantum mechanics which one of the authors (SKS) has been teaching at Arizona State University for more than two decades.

3.1 Dirac Equation

The relativistic wave equation of Dirac [30, 31, 53–55, 70, 71, 111]

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi \tag{3.1}$$

for the electron in an external central field with the potential energy $U(r)$ has the Hamiltonian of the form

$$H = c\boldsymbol{\alpha}\boldsymbol{p} + mc^2\beta + U(r), \tag{3.2}$$

where $\boldsymbol{\alpha}\boldsymbol{p} = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$ with the momentum operator $\boldsymbol{p} = -i\hbar\nabla$ and

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \qquad \psi = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}. \tag{3.3}$$

We use the standard representation of the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.4}$$

and by our definition

$$\mathbf{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The relativistic electron has a four component wave function

$$\psi = \psi\left(\mathbf{r}, t\right) = \begin{pmatrix} \mathbf{u}\left(\mathbf{r}, t\right) \\ \mathbf{v}\left(\mathbf{r}, t\right) \end{pmatrix} = \begin{pmatrix} \psi_1\left(\mathbf{r}, t\right) \\ \psi_2\left(\mathbf{r}, t\right) \\ \psi_3\left(\mathbf{r}, t\right) \\ \psi_4\left(\mathbf{r}, t\right) \end{pmatrix} \tag{3.5}$$

and the Dirac equation (3.1) is a matrix equation that is equivalent to a system of four first order partial differential equations. The inner product for two Dirac (bispinor) wave functions

$$\psi = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \qquad \phi = \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix}$$

is defined as a scalar quantity

$$\begin{aligned} \langle \psi,\ \phi \rangle &= \int_{\mathbb{R}^3} \psi^{\dagger} \phi \, dv = \int_{\mathbb{R}^3} \left(\mathbf{u}_1{}^{\dagger} \mathbf{u}_2 + \mathbf{v}_1{}^{\dagger} \mathbf{v}_2 \right) \, dv \\ &= \int_{\mathbb{R}^3} \left(\overset{*}{\psi}_1 \phi_1 + \overset{*}{\psi}_2 \phi_2 + \overset{*}{\psi}_3 \, \phi_3 + \overset{*}{\psi}_4 \, \phi_4 \right) \, dv \end{aligned} \tag{3.6}$$

with the squared norm

$$\begin{aligned} ||\psi||^2 = \langle \psi,\ \psi \rangle &= \int_{\mathbb{R}^3} \psi^{\dagger} \psi \, dv = \int_{\mathbb{R}^3} \left(\mathbf{u}_1{}^{\dagger} \mathbf{u}_1 + \mathbf{v}_1{}^{\dagger} \mathbf{v}_1 \right) \, dv \\ &= \int_{\mathbb{R}^3} \left(|\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \right) \, dv \end{aligned} \tag{3.7}$$

and the wave functions are usually normalized so that $||\psi|| = \langle \psi,\ \psi \rangle^{1/2} = 1$.

The substitution

$$\psi\left(\mathbf{r}, t\right) = e^{-i(E\ t)/\hbar}\ \psi\left(\mathbf{r}\right), \tag{3.8}$$

gives the stationary Dirac equation

$$H\psi\left(\mathbf{r}\right) = E\psi\left(\mathbf{r}\right), \tag{3.9}$$

where E is the total energy of the electron.

According to Steven Weinberg [194, vol. I, p. 565], physicists learn in kindergarten how to solve problems related to the wave equation of Dirac in the presence of external fields. In Sect. 3.3 of this chapter, for the benefits of the reader who is not an expert in theoretical physics, we outline a procedure of separation of the variables and solve the corresponding first order system of radial equations for the Dirac equation in the Coulomb field $U(r) = -Ze^2/r$. The end results are presented in the next section; see also [2, 25, 28, 71, 96, 133, 138, 140, 156] and references therein for more information.

3.2 Relativistic Coulomb Wave Functions and Discrete Energy Levels

The exact solutions of the stationary Dirac equation

$$H\psi = \left(c\boldsymbol{\alpha p} + mc^2\beta - Ze^2/r\right)\psi = E\psi \tag{3.10}$$

for the Coulomb potential can be obtained in the spherical coordinates. The energy levels were discovered in 1916 by Sommerfeld [168] from the "old" quantum theory and the corresponding (bispinor) Dirac wave functions were found later by Darwin [48] and Gordon [83] at the early age of discovery of the "new" wave mechanics (see also [29] for a modern discussion of "Sommerfeld's puzzle"). These classical results are nowadays included in all textbooks on relativistic quantum mechanics, quantum field theory and advanced texts on mathematical physics (see, for example, [2, 25, 28, 96, 133, 140] and [138]). The end result is

$$\psi = \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix} = \begin{pmatrix} \mathcal{Y}^{\pm}_{jm}(\mathbf{n})\ F(r) \\ i\mathcal{Y}^{\mp}_{jm}(\mathbf{n})\ G(r) \end{pmatrix}, \tag{3.11}$$

where the spinor spherical harmonics $\mathcal{Y}^{\pm}_{jm}(\mathbf{n}) = \mathcal{Y}^{(j\pm 1/2)}_{jm}(\mathbf{n})$ are given explicitly in terms of the ordinary spherical functions $Y_{lm}(\mathbf{n})$, $\mathbf{n} = \mathbf{n}(\theta, \varphi) = \mathbf{r}/r$ and the special Clebsch–Gordan coefficients with the spin $1/2$ as follows [2, 25, 148, 189]:

$$\mathcal{Y}^{\pm}_{jm}(\mathbf{n}) = \begin{pmatrix} \mp\sqrt{\dfrac{(j+1/2)\mp(m-1/2)}{2j+(1\pm 1)}}\ Y_{j\pm 1/2,\ m-1/2}(\mathbf{n}) \\ \sqrt{\dfrac{(j+1/2)\pm(m+1/2)}{2j+(1\pm 1)}}\ Y_{j\pm 1/2,\ m+1/2}(\mathbf{n}) \end{pmatrix} \tag{3.12}$$

with the total angular momentum $j = 1/2, 3/2, 5/2, \ldots$ and its projection $m = -j, -j+1, \ldots, j-1, j$ (see also Sect. 3.3.1 below for the properties of the spinor spherical harmonics).

The radial functions $F(r)$ and $G(r)$ can be presented as [138]

$$\begin{pmatrix} F(r) \\ G(r) \end{pmatrix} = \frac{a^2\beta^{3/2}}{\nu}\sqrt{\frac{(\varepsilon\kappa-\nu)\,n!}{\mu(\kappa-\nu)\,\Gamma(n+2\nu)}}\,\xi^{\nu-1}e^{-\xi/2} \times \begin{pmatrix} f_1 & f_2 \\ g_1 & g_2 \end{pmatrix}\begin{pmatrix} \xi L_{n-1}^{2\nu+1}(\xi) \\ L_n^{2\nu-1}(\xi) \end{pmatrix}. \tag{3.13}$$

Here, $L_k^{\alpha}(\xi)$ are the Laguerre polynomials given by (B.1) and we use the following notations:

$$\kappa = \pm(j+1/2), \qquad \nu = \sqrt{\kappa^2-\mu^2}, \qquad \mu = \alpha Z = Ze^2/\hbar c, \tag{3.14}$$
$$a = \sqrt{1-\varepsilon^2}, \qquad \varepsilon = E/mc^2, \qquad \beta = mc/\hbar = 1/\lambda,$$

and

$$\xi = 2a\beta r = 2\sqrt{1-\varepsilon^2}\,\frac{mc}{\hbar}\,r. \tag{3.15}$$

The elements of 2×2-transition matrix in (3.13) are given by

$$f_1 = \frac{a\mu}{\varepsilon\kappa-\nu}, \quad f_2 = \kappa-\nu, \quad g_1 = \frac{a(\kappa-\nu)}{\varepsilon\kappa-\nu}, \quad g_2 = \mu. \tag{3.16}$$

This particular form of the relativistic radial functions is due to Nikiforov and Uvarov [138]; it is very convenient for taking the nonrelativistic limit $c\to\infty$ (see also [183]).

The relativistic discrete energy levels $\varepsilon = \varepsilon_n = E_n/E_0$ with the rest mass energy $E_0 = mc^2$ are given by the Sommerfeld–Dirac fine structure formula

$$E_n = \frac{mc^2}{\sqrt{1+\mu^2/(n+\nu)^2}}. \tag{3.17}$$

Here, $n = n_r = 0, 1, 2, \ldots$ is the radial quantum number and $\kappa = \pm(j+1/2) = \pm 1, \pm 2, \pm 3, \ldots$. The following identities

$$\varepsilon\mu = a(\nu+n), \quad \varepsilon\mu + a\nu = a(n+2\nu), \quad \varepsilon\mu - a\nu = an, \tag{3.18}$$
$$\varepsilon^2\kappa^2-\nu^2 = a^2 n(n+2\nu) = \mu^2 - a^2\kappa^2$$

are useful in calculation of the matrix elements below.

The familiar recurrence relations for the Laguerre polynomials allow to present the radial functions (3.13) in a traditional form [2, 25, 49, 128, 183] as follows

$$\begin{pmatrix} F(r) \\ G(r) \end{pmatrix} = a^2\beta^{3/2}\sqrt{\frac{n!}{\mu(\kappa-\nu)(\varepsilon\kappa-\nu)\Gamma(n+2\nu)}}\,\xi^{\nu-1}e^{-\xi/2} \times \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix}\begin{pmatrix} L_{n-1}^{2\nu}(\xi) \\ L_n^{2\nu}(\xi) \end{pmatrix}, \tag{3.19}$$

where the coefficients are found in (3.106) and (3.107) below.

We give the explicit form of the radial wave functions (3.13) for the $1s_{1/2}$ state, when $n = n_r = 0,\ l = 0,\ j = 1/2$, and $\kappa = -1$:

$$\begin{pmatrix} F(r) \\ G(r) \end{pmatrix} = \left(\frac{2Z}{a_0}\right)^{3/2}\sqrt{\frac{\nu_1+1}{2\Gamma(2\nu_1+1)}}\begin{pmatrix} -1 \\ \sqrt{\dfrac{1-\nu_1}{1+\nu_1}} \end{pmatrix}\xi_1^{\nu_1-1}e^{-\xi_1/2}. \tag{3.20}$$

Here, $\nu_1 = \sqrt{1-\mu^2} = \varepsilon_1,\ \xi_1 = 2\sqrt{1-\varepsilon_1^2}\beta r = 2Z(r/a_0)$, and $a_0 = \hbar^2/me^2$ is the Bohr radius. One can see also [2, 25, 28, 48, 64, 83, 96, 128, 133], and [156] and references therein for more information on the relativistic Coulomb problem.

In the nonrelativistic limit $c \to \infty$ one can expand the exact Sommerfeld–Dirac formula (3.17) in ascending powers of $\mu^2 = (\alpha Z)^2$, the first terms in this expansion are

$$\frac{E}{mc^2} = 1 - \frac{\mu^2}{2n^2} - \frac{\mu^4}{2n^4}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) + \mathrm{O}\left(\mu^6\right), \quad \mu \to 0. \tag{3.21}$$

Here $n = n_r + j + 1/2$ is the principal quantum number of the nonrelativistic hydrogenlike atom. The first term in this expansion is simply the rest mass energy $E_0 = mc^2$ of the electron, the second term coincides with the energy eigenvalue in the nonrelativistic Schrödinger theory (2.64) and the third term gives the so-called fine structure of the energy levels—the correction obtained for the energy in the Pauli approximation which includes interaction of the spin of the electron with its orbital angular momentum; see [28] and [156] for further discussion of the hydrogenlike energy levels including comparison with the experimental data. One can show that in the same limit $\mu \to 0$ the relativistic Coulomb wave functions (3.11) tend to the nonrelativistic wave functions of the Pauli theory; see, for example, [138] for more details; we shall elaborate more on this limit in Sect. 3.3.4.

3.3 Solution of Dirac Wave Equation for Coulomb Potential

This section is written for the benefits of the reader who is not an expert in relativistic quantum mechanics and quantum field theory. We separate the variables and construct exact solutions of the Dirac equation of in spherical coordinates for the Coulomb field. The corresponding four components (bispinor) wave functions are given explicitly by (3.11)–(3.16). We first construct the angular parts of these solutions in terms of the so-called spinor spherical harmonics or spherical spinors.

3.3.1 The Spinor Spherical Harmonics

The vector addition $\mathbf{j} = \mathbf{l} + \mathbf{s}$ of the orbital $\mathbf{l} = -i\mathbf{r} \times \nabla$ and the spin $\mathbf{s} = \dfrac{1}{2}\boldsymbol{\sigma}$ angular momenta (in the units of $\hbar$) for the electron in the central field gives the eigenfunctions of the total angular momentum $\mathbf{j}$, or the spinor spherical harmonics [189], in the form

$$\mathcal{Y}_{jm}^{(l)}(\mathbf{n}) = \sum_{m_l+m_s=m} C_{lm_l \frac{1}{2} m_s}^{jm} Y_{lm_l}(\mathbf{n})\ \chi_{\frac{1}{2}m_s} \tag{3.22}$$

$$(j = |l - 1/2|\,, l + 1/2; \qquad m = -j, -j+1,\ \ldots,\ j-1, j)$$

where $Y_{lm}(\mathbf{n})$ with $\mathbf{n} = \mathbf{n}(\theta,\varphi) = \mathbf{r}/r$ are the spherical harmonics, $C_{lm_l \frac{1}{2} m_s}^{jm}$ are the special Clebsch–Gordan coefficients, and $\chi_{m_s} = \chi_{\frac{1}{2}m_s}$ are eigenfunctions of the spin $1/2$ operator $\mathbf{s}$:

$$\mathbf{s}^2 \chi_{\frac{1}{2}m_s} = \frac{3}{4}\chi_{\frac{1}{2}m_s}, \qquad s_3 \chi_{\frac{1}{2}m_s} = m_s \chi_{\frac{1}{2}m_s}, \qquad m_s = \pm 1/2 \tag{3.23}$$

given by

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \tag{3.24}$$

see [113, 139, 148, 189]. From (3.22)

$$\begin{aligned} \mathcal{Y}_{jm}^{(l)}(\mathbf{n}) &= \sum_{m_s=-1/2}^{1/2} C_{l,\ m-m_s, \frac{1}{2} m_s}^{jm} Y_{l,\ m-m_s}(\mathbf{n})\ \chi_{m_s} \\ &= C_{l,\ m+\frac{1}{2}, \frac{1}{2},\ -\frac{1}{2}}^{jm} Y_{l,\ m+\frac{1}{2}}(\mathbf{n})\ \chi_{-\frac{1}{2}} + C_{l,\ m-\frac{1}{2}, \frac{1}{2},\ \frac{1}{2}}^{jm} Y_{l,\ m-\frac{1}{2}}(\mathbf{n})\ \chi_{\frac{1}{2}} \\ &= \begin{pmatrix} C_{l,\ m-\frac{1}{2}, \frac{1}{2},\ \frac{1}{2}}^{jm} Y_{l,\ m-\frac{1}{2}}(\mathbf{n}) \\ C_{l,\ m+\frac{1}{2}, \frac{1}{2},\ -\frac{1}{2}}^{jm} Y_{l,\ m+\frac{1}{2}}(\mathbf{n}) \end{pmatrix}, \qquad l = j \pm 1/2. \end{aligned} \tag{3.25}$$

Substituting the special values of the Clebsch–Gordan coefficients [62, 189], we obtain the spinor spherical harmonics $\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = \mathcal{Y}_{jm}^{(j\pm1/2)}(\mathbf{n})$ in the form (3.12). The orthogonality property for the spinor spherical harmonics $\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = \mathcal{Y}_{jm}^{(j\pm1/2)}(\mathbf{n})$ is given by Varshalovich et al. [189]

$$\int_{S^2} \left(\mathcal{Y}_{jm}^{(l)}(\mathbf{n})\right)^{\dagger} \mathcal{Y}_{j'm'}^{(l')}(\mathbf{n})\ d\omega = \delta_{jj'}\delta_{ll'}\delta_{mm'} \tag{3.26}$$

with $d\omega = \sin\theta\ d\theta d\varphi$ and $0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi$. They are common eigenfunctions of the following set of commuting operators

$$\mathbf{j}^2\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = \left(\mathbf{l} + \frac{1}{2}\boldsymbol{\sigma}\right)^2 \mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = j\,(j+1)\,\mathcal{Y}_{jm}^{\pm}(\mathbf{n}), \tag{3.27}$$

$$j_3\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = m\mathcal{Y}_{jm}^{\pm}(\mathbf{n}), \tag{3.28}$$

$$\mathbf{l}^2\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = \left(j \pm \frac{1}{2}\right)\left(j \pm \frac{1}{2} + 1\right)\mathcal{Y}_{jm}^{\pm}(\mathbf{n}), \tag{3.29}$$

$$\boldsymbol{\sigma}^2\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = 3\mathcal{Y}_{jm}^{\pm}(\mathbf{n}). \tag{3.30}$$

However,

$$\mathbf{j}^2 = \left(\mathbf{l} + \frac{1}{2}\boldsymbol{\sigma}\right)^2 = \mathbf{l}^2 + \boldsymbol{\sigma}\cdot\boldsymbol{l} + \frac{3}{4},$$

or

$$\boldsymbol{\sigma}\cdot\boldsymbol{l} = \mathbf{j}^2 - \mathbf{l}^2 - \frac{3}{4}. \tag{3.31}$$

This implies that the spinor spherical harmonics $\mathcal{Y}_{jm}^{\pm}(\mathbf{n})$ are also eigenfunctions of the operator $\boldsymbol{\sigma}\cdot\boldsymbol{l}$:

$$(\boldsymbol{\sigma}\cdot\boldsymbol{l})\,\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = -\left(1 \pm \left(j + \frac{1}{2}\right)\right)\mathcal{Y}_{jm}^{\pm}(\mathbf{n}), \tag{3.32}$$

and it is a custom to write

$$(\boldsymbol{\sigma}\cdot\boldsymbol{l})\,\mathcal{Y}_{jm}^{\pm}(\mathbf{n}) = -\,(1+\kappa)\,\mathcal{Y}_{jm}^{\pm}(\mathbf{n}), \tag{3.33}$$

where the quantum number $\kappa = \kappa_{\pm} = \pm\left(j + \frac{1}{2}\right) = \pm1, \pm2, \pm3, \ldots$ takes all positive and negative integer values with exception of zero: $\kappa \neq 0$.

Finally, the following relation for the spinor spherical harmonics,

$$(\boldsymbol{\sigma} \cdot \mathbf{n})\, \mathcal{Y}^{\pm}_{jm}(\mathbf{n}) = -\mathcal{Y}^{\mp}_{jm}(\mathbf{n}), \tag{3.34}$$

plays an important role in the Dirac theory of relativistic electron. In view of $(\boldsymbol{\sigma} \cdot \mathbf{n})^2 = \mathbf{1}$, it is sufficient to prove only one of these relations, say

$$(\boldsymbol{\sigma} \cdot \mathbf{n})\, \mathcal{Y}^{+}_{jm}(\mathbf{n}) = -\mathcal{Y}^{-}_{jm}(\mathbf{n}),$$

and the second will follow. A direct proof can be given by using the recurrence relations for the spherical harmonics given in the Appendix B, (B.16)–(B.18), or with the help of the Wigner–Eckart theorem; see [148] and [189], the reader can work out the details.

The quadratic forms

$$Q_{jm} = \left(\mathcal{Y}^{(l)}_{jm}(\mathbf{n})\right)^{\dagger}\, \mathcal{Y}^{(l)}_{jm}(\mathbf{n}) \tag{3.35}$$

of the spinor spherical harmonics $\mathcal{Y}^{(l)}_{jm}(\mathbf{n})$ describe the angular distributions of the electron in states with the total angular momentum j, its projection m and the orbital angular momentum l. These forms, given by [189]

$$\begin{aligned} Q_{jm}(\theta) &= \frac{1}{2j}\left((j+m)\left|Y_{j-\frac{1}{2},\, m-\frac{1}{2}}(\mathbf{n})\right|^2 + (j-m)\left|Y_{j-\frac{1}{2},\, m+\frac{1}{2}}(\mathbf{n})\right|^2\right) \\ &= \frac{1}{2j+2}\left((j+m+1)\left|Y_{j+\frac{1}{2},\, m+\frac{1}{2}}(\mathbf{n})\right|^2 + (j-m+1)\left|Y_{j+\frac{1}{2},\, m-\frac{1}{2}}(\mathbf{n})\right|^2\right), \end{aligned} \tag{3.36}$$

are, in fact, independent of l and φ. There is the useful expansion in terms of the Legendre polynomials

$$Q_{jm}(\theta) = \sum_{s=0}^{j-1/2} a_s(j,m)\; P_{2s}(\cos\theta) \tag{3.37}$$

with the coefficients of the form

$$\begin{aligned} a_s(j,m) &= -\frac{(4s+1)\sqrt{2j(2j+1)}}{4\pi}\begin{Bmatrix} j & j & 2s \\ j-\frac{1}{2} & j-\frac{1}{2} & \frac{1}{2}\end{Bmatrix} C^{j-\frac{1}{2},0}_{j-\frac{1}{2},0,2s0}\, C^{jm}_{jm\,2s0} \\ &= (-1)^s\, \frac{4s+1}{4\pi}\sqrt{\frac{(2j+2s+1)(2j-2s)!}{(2j+1)(2j+2s)!}}\; \frac{\left(j+s-\frac{1}{2}\right)!\,(2s)!}{\left(j-s-\frac{1}{2}\right)!\,(s!)^2}\, C^{jm}_{jm\,2s0}. \end{aligned} \tag{3.38}$$

See [189] for more information, including definition of the $6j$-symbol.

3.3.2 Separation of Variables in Spherical Coordinates

Using the explicit form of the $\boldsymbol{\alpha}$ and β matrices

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \tag{3.39}$$

we rewrite the stationary Dirac equation

$$H\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{3.40}$$

in a central field with the Hamiltonian

$$H = c\boldsymbol{\alpha p} + mc^2\beta + U(r) = \begin{pmatrix} U + mc^2 & c\boldsymbol{\sigma p} \\ c\boldsymbol{\sigma p} & U - mc^2 \end{pmatrix} \tag{3.41}$$

and the bispinor wave function

$$\psi = \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix} \tag{3.42}$$

in a block matrix form

$$\begin{pmatrix} U + mc^2 & c\boldsymbol{\sigma p} \\ c\boldsymbol{\sigma p} & U - mc^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix} = E \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix}, \tag{3.43}$$

or

$$c\boldsymbol{\sigma p\varphi} = \left(E + mc^2 - U\right)\boldsymbol{\chi}, \tag{3.44}$$

$$c\boldsymbol{\sigma p\chi} = \left(E - mc^2 - U\right)\boldsymbol{\varphi}. \tag{3.45}$$

Here we shall use the following operator identity

$$\boldsymbol{\sigma} \cdot \nabla = (\boldsymbol{\sigma} \cdot \mathbf{n})\left(\mathbf{n} \cdot \nabla + i\boldsymbol{\sigma} \cdot (\mathbf{n} \times \nabla)\right) \tag{3.46}$$

in the form

$$c\boldsymbol{\sigma p} = \hbar c\,(\boldsymbol{\sigma}\mathbf{n})\left(\frac{1}{i}\mathbf{n}\nabla + \frac{i}{r}\boldsymbol{\sigma l}\right), \tag{3.47}$$

where $\mathbf{l} = -i\mathbf{r} \times \nabla$ is the operator of orbital angular momentum, $\mathbf{n} = \mathbf{r}/r$ and $\boldsymbol{p} = -i\hbar\nabla$. It can be obtained as a consequence of a more general operator identity [53]

$$(\boldsymbol{\sigma} \cdot \mathbf{A})\,(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\,\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B})\,, \tag{3.48}$$

which is valid for any vector operators $\mathbf{A}$ and $\mathbf{B}$ commuting with the Pauli $\boldsymbol{\sigma}$-matrices; it is not required that $\mathbf{A}$ and $\mathbf{B}$ commute. The proof uses a familiar property of the Pauli matrices

$$\sigma_i\ \sigma_k = i e_{ikl}\ \sigma_l + \delta_{ik}, \tag{3.49}$$

where e_{ikl} is the completely antisymmetric Levi-Civita symbol, δ_{ik} is the symmetric Kronecker delta symbol and we use Einstein's summation rule over the repeating indices; it is understood that a summation is to be taken over the three values of $l = 1, 2, 3$. Thus

$$\begin{aligned}(\boldsymbol{\sigma} \cdot \mathbf{A})\,(\boldsymbol{\sigma} \cdot \mathbf{B}) &= (\sigma_i\ A_i)\,(\sigma_k\ B_k)\\ &= (\sigma_i\ \sigma_k)\ A_i\ B_k = i\sigma_l\ e_{ikl}\ A_i\ B_k + \delta_{ik}\ A_i\ B_k\\ &= i\sigma_l\ (\mathbf{A} \times \mathbf{B})_l + A_k\ B_k = i\,\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B}) + \mathbf{A} \cdot \mathbf{B},\end{aligned}$$

where $(\mathbf{A} \times \mathbf{B})_l = e_{lik}\ A_i\ B_k = e_{ikl}\ A_i\ B_k$ in view of antisymmetry of the Levi-Civita symbol: $e_{ikl} = -e_{ilk} = e_{lik}$, and $\delta_{ik}\ A_i = A_k$.

If $\mathbf{A} = \mathbf{B}$, Eq. (3.48) implies $(\boldsymbol{\sigma} \cdot \mathbf{A})^2 = \mathbf{A}^2$. In particular, $(\boldsymbol{\sigma}\,\mathbf{n})^2 = \mathbf{n}^2 = \mathbf{1}$, and the proof of the "gradient" formula (3.46) is

$$\begin{aligned}\boldsymbol{\sigma} \cdot \nabla &= (\boldsymbol{\sigma} \cdot \mathbf{n})^2\,(\boldsymbol{\sigma} \cdot \nabla) = (\boldsymbol{\sigma} \cdot \mathbf{n})\,((\boldsymbol{\sigma} \cdot \mathbf{n})\,(\boldsymbol{\sigma} \cdot \nabla))\\ &= (\boldsymbol{\sigma} \cdot \mathbf{n})\,(\mathbf{n} \cdot \nabla + i\boldsymbol{\sigma} \cdot (\mathbf{n} \times \nabla))\end{aligned}$$

by (3.48) with $\mathbf{A} = \mathbf{n}$ and $\mathbf{B} = \nabla$. In a similar fashion, one can derive the following "anticommutation" relation,

$$(\boldsymbol{\sigma}\mathbf{n})\,(\boldsymbol{\sigma}\boldsymbol{l}) + (\boldsymbol{\sigma}\boldsymbol{l})\,(\boldsymbol{\sigma}\mathbf{n}) = -2\,(\boldsymbol{\sigma}\mathbf{n})\,, \qquad \mathbf{n} = \mathbf{r}/r, \tag{3.50}$$

we leave details to the reader.

The structure of operator $\boldsymbol{\sigma}\boldsymbol{p}$ in (3.47) suggests to look for solutions of the Dirac system (3.44) and (3.45) in spherical coordinates $\mathbf{r} = r\ \mathbf{n}\,(\theta, \varphi)$ in the form of the Ansatz:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}\,(\mathbf{r}) = \mathcal{Y}\,(\mathbf{n})\ \ F\,(r)\,, \tag{3.51}$$

$$\boldsymbol{\chi} = \boldsymbol{\chi}\,(\mathbf{r}) = -i\,((\boldsymbol{\sigma}\mathbf{n})\,\mathcal{Y}\,(\mathbf{n}))\ \ G\,(r)\,, \tag{3.52}$$

where $\mathcal{Y} = \mathcal{Y}_{jm}^{\pm}(\mathbf{n})$ are the spinor spherical harmonics given by (3.12). This substitution preserves the symmetry properties of the wave functions under inversion $\mathbf{r} \to -\mathbf{r}$. Then the radial functions $F(r)$ and $G(r)$ satisfy the system of two first order ordinary differential equations

$$\frac{dF}{dr} + \frac{1+\kappa}{r} F = \frac{mc^2 + E - U(r)}{\hbar c} G, \tag{3.53}$$

$$\frac{dG}{dr} + \frac{1-\kappa}{r} G = \frac{mc^2 - E + U(r)}{\hbar c} F, \tag{3.54}$$

where $\kappa = \kappa_{\pm} = \pm(j+1/2) = \pm 1, \pm 2, \pm 3, \ \dots$, respectively.

If $f = f(\mathbf{r}) = f(r\mathbf{n})$, then

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial \mathbf{r}} \frac{\partial \mathbf{r}}{\partial r} = \mathbf{n} \nabla f$$

and in spherical coordinates Eq. (3.47) becomes

$$c\boldsymbol{\sigma}\boldsymbol{p} = \hbar c(\boldsymbol{\sigma}\mathbf{n}) \left(\frac{1}{i} \frac{\partial}{\partial r} + \frac{i}{r} \boldsymbol{\sigma}\boldsymbol{l} \right). \tag{3.55}$$

Thus

$$\begin{aligned} c\boldsymbol{\sigma}\boldsymbol{p}\,\boldsymbol{\varphi} &= \hbar c(\boldsymbol{\sigma}\mathbf{n}) \left(\frac{1}{i} \frac{\partial}{\partial r} + \frac{i}{r} \boldsymbol{\sigma}\boldsymbol{l} \right) \mathcal{Y} F \\ &= \hbar c(\boldsymbol{\sigma}\mathbf{n}) \left(\frac{1}{i} \mathcal{Y} \frac{dF}{dr} + \frac{i}{r} (\boldsymbol{\sigma}\boldsymbol{l}\mathcal{Y}) F \right) \\ &= -i\hbar c(\boldsymbol{\sigma}\mathbf{n}\,\mathcal{Y}) \left(\frac{dF}{dr} + \frac{1+\kappa}{r} F \right) \end{aligned}$$

by (3.33), and we arrive at (3.53) in view of (3.44) and (3.52). Equation (3.54) can be verified in a similar fashion with the help of (3.50) or (3.34).

Equations (3.53) and (3.54) hold in any central field with the potential energy $U = U(r)$. For states with discrete spectra the radial functions $rF(r)$ and $rG(r)$ should be bounded as $r \to 0$ and satisfy the normalization condition

$$\int_{\mathbb{R}^3} \psi^{\dagger}\,\psi\,dv = \int_0^{\infty} r^2 \left(F^2(r) + G^2(r) \right)\,dr = 1 \tag{3.56}$$

in view of

$$||\psi||^2 = \langle \psi, \ \psi \rangle = \int_{\mathbb{R}^3} \psi^\dagger \psi \, dv = \int_{\mathbb{R}^3} \left(\mathbf{u}_1{}^\dagger \mathbf{u}_1 + \mathbf{v}_1{}^\dagger \mathbf{v}_1 \right) \, dv \qquad (3.57)$$
$$= \int_{\mathbb{R}^3} \left(|\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \right) \, dv.$$

And Eqs. (3.51), (3.52) and (3.26).

3.3.3 Solution of Radial Equations

For the relativistic Coulomb problem, $U = -Ze^2/r$, we introduce the dimensionless quantities

$$\varepsilon = \frac{E}{mc^2}, \qquad x = \beta r = \frac{mc}{\hbar} r, \qquad \mu = \frac{Ze^2}{\hbar c} \qquad (3.58)$$

and the radial functions

$$f(x) = F(r), \qquad g(x) = G(r). \qquad (3.59)$$

The system (3.53) and (3.54) becomes

$$\frac{df}{dx} + \frac{1+\kappa}{x} f = \left(1 + \varepsilon + \frac{\mu}{x} \right) g, \qquad (3.60)$$
$$\frac{dg}{dx} + \frac{1-\kappa}{x} g = \left(1 - \varepsilon - \frac{\mu}{x} \right) f. \qquad (3.61)$$

We shall see later that in nonrelativistic limit $c \to \infty$ the following estimate holds $|f(x)| \gg |g(x)|$.

We follow [138] with somewhat different details. Let us rewrite the system (3.60) and (3.61) in matrix form [83]. If

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} xf(x) \\ xg(x) \end{pmatrix}, \qquad u' = \begin{pmatrix} u_1' \\ u_2' \end{pmatrix}. \qquad (3.62)$$

Then

$$u' = Au, \qquad (3.63)$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -\dfrac{\kappa}{x} & 1 + \varepsilon + \dfrac{\mu}{x} \\ 1 - \varepsilon - \dfrac{\mu}{x} & \dfrac{\kappa}{x} \end{pmatrix}. \tag{3.64}$$

To find $u_1(x)$, we eliminate $u_2(x)$ from the system (3.63), obtaining a second order differential equation

$$\begin{aligned} u_1'' &- \left(a_{11} + a_{22} + \frac{a_{12}'}{a_{12}} \right) u_1' \\ &+ \left(a_{11}a_{22} - a_{12}a_{21} - a_{11}' + \frac{a_{12}'}{a_{12}}\, a_{11} \right) u_1 = 0. \end{aligned} \tag{3.65}$$

Similarly, eliminating $u_1(x)$, one gets equation for $u_2(x)$:

$$\begin{aligned} u_2'' &- \left(a_{11} + a_{22} + \frac{a_{21}'}{a_{21}} \right) u_2' \\ &+ \left(a_{11}a_{22} - a_{12}a_{21} - a_{22}' + \frac{a_{21}'}{a_{21}}\, a_{22} \right) u_2 = 0. \end{aligned} \tag{3.66}$$

The components of matrix A have the form

$$a_{ik} = b_{ik} + c_{ik}/x, \tag{3.67}$$

where b_{ik} and c_{ik} are constants. Equations (3.65) and (3.66) are not generalized equations of hypergeometric type (2.1). Indeed,

$$\frac{a_{12}'}{a_{12}} = -\frac{c_{12}}{c_{12}x + b_{12}x^2},$$

and the coefficients of $u_1'(x)$ and $u_1(x)$ in (3.65) are

$$\begin{aligned} a_{11} + a_{22} + \frac{a_{12}'}{a_{12}} &= \frac{p_1(x)}{x} - \frac{c_{12}}{c_{12}x + b_{12}x^2}, \\ a_{11}a_{22} - a_{12}a_{21} - a_{11}' + \frac{a_{12}'}{a_{12}} a_{11} &= \frac{p_2(x)}{x^2} - \frac{c_{12}(c_{11} + b_{11}x)}{(c_{12} + b_{12}x)\, x^2}, \end{aligned}$$

where $p_1(x)$ and $p_2(x)$ are polynomials of degrees at most one and two, respectively. Equation (3.65) will become a generalized equation of hypergeometric type (2.1) with $\sigma(x) = x$ if either $b_{12} = 0$ or $c_{12} = 0$. The following consideration helps. By a linear transformation

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = C \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{3.68}$$

with a nonsingular matrix C that is independent of x we transform the original system (3.63) to a similar one

$$v' = \widetilde{A} v, \tag{3.69}$$

where

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \qquad \widetilde{A} = CAC^{-1} = \begin{pmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} \\ \widetilde{a}_{21} & \widetilde{a}_{22} \end{pmatrix}.$$

The new coefficients $\widetilde{a}_{ik}$ are linear combinations of the original ones a_{ik}. Hence they have a similar form

$$\widetilde{a}_{ik} = \widetilde{b}_{ik} + \widetilde{c}_{ik}/x, \tag{3.70}$$

where $\widetilde{b}_{ik}$ and $\widetilde{c}_{ik}$ are constants.

The equations for $v_1(x)$ and $v_2(x)$ are similar to (3.65) and (3.66):

$$v_1'' - \left(\widetilde{a}_{11} + \widetilde{a}_{22} + \frac{\widetilde{a}_{12}'}{\widetilde{a}_{12}} \right) v_1' + \left(\widetilde{a}_{11}\widetilde{a}_{22} - \widetilde{a}_{12}\widetilde{a}_{21} - \widetilde{a}_{11}' + \frac{\widetilde{a}_{12}'}{\widetilde{a}_{12}}\, \widetilde{a}_{11} \right) v_1 = 0, \tag{3.71}$$

$$v_2'' - \left(\widetilde{a}_{11} + \widetilde{a}_{22} + \frac{\widetilde{a}_{21}'}{\widetilde{a}_{21}} \right) v_2' + \left(\widetilde{a}_{11}\widetilde{a}_{22} - \widetilde{a}_{12}\widetilde{a}_{21} - \widetilde{a}_{22}' + \frac{\widetilde{a}_{21}'}{\widetilde{a}_{21}}\, \widetilde{a}_{22} \right) v_2 = 0. \tag{3.72}$$

The calculation of the coefficients in (3.71) and (3.72) is facilitated by a similarity of the matrices A and $\widetilde{A}$:

$$\widetilde{a}_{11} + \widetilde{a}_{22} = a_{11} + a_{22}, \qquad \widetilde{a}_{11}\widetilde{a}_{22} - \widetilde{a}_{12}\widetilde{a}_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

By a previous consideration, in order for (3.71) to be an equation of hypergeometric type, it is sufficient to choose either $\widetilde{b}_{12} = 0$ or $\widetilde{c}_{12} = 0$. Similarly, for (3.72): either

$\widetilde{b}_{21} = 0$ or $\widetilde{c}_{21} = 0$. These conditions impose certain restrictions on our choice of the transformation matrix C. Let

$$C = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{3.73}$$

Then

$$C^{-1} = \frac{1}{\Delta} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}, \qquad \Delta = \det C = \alpha\delta - \beta\gamma,$$

and

$$\begin{aligned} \widetilde{A} &= CAC^{-1} \qquad (3.74) \\ &= \frac{1}{\Delta} \begin{pmatrix} (a_{11}\delta - a_{12}\gamma)\alpha + (a_{21}\beta\delta - a_{22}\gamma)\beta & (\alpha^2 - \beta^2)a_{12} + (a_{22} - a_{11})\,\alpha\beta \\ a_{21}\delta^2 - a_{12}\gamma^2 + (a_{11} - a_{22})\,\gamma\delta & a_{12}\alpha\gamma - a_{11}\beta\gamma + a_{22}\alpha\delta - a_{21}\beta\delta \end{pmatrix}. \end{aligned}$$

For the Dirac system (3.63) and (3.64):

$$\begin{aligned} a_{11} &= -\frac{\kappa}{x}, \qquad a_{12} = 1 + \varepsilon + \frac{\mu}{x}, \\ a_{21} &= 1 - \varepsilon - \frac{\mu}{x}, \qquad a_{22} = \frac{\kappa}{x} \end{aligned}$$

and

$$\Delta\ \widetilde{a}_{12} = \alpha^2 - \beta^2 + \left(\alpha^2 + \beta^2\right)\varepsilon + \frac{\left(\alpha^2 + \beta^2\right)\mu + 2\alpha\beta\kappa}{x}, \tag{3.75}$$

$$\Delta\ \widetilde{a}_{21} = \delta^2 - \gamma^2 - \left(\delta^2 + \gamma^2\right)\varepsilon - \frac{\left(\delta^2 + \gamma^2\right)\mu + 2\gamma\delta\kappa}{x}. \tag{3.76}$$

The condition	$\widetilde{b}_{12} = 0$	yields	$(1+\varepsilon)\,\alpha^2 - (1-\varepsilon)\,\beta^2 = 0,$
” ”	$\widetilde{c}_{12} = 0$	”	$\left(\alpha^2 + \beta^2\right)\mu + 2\alpha\beta\kappa = 0,$
” ”	$\widetilde{b}_{21} = 0$	”	$(1+\varepsilon)\,\gamma^2 - (1-\varepsilon)\,\delta^2 = 0,$
” ”	$\widetilde{c}_{21} = 0$	”	$\left(\delta^2 + \gamma^2\right)\mu + 2\gamma\delta\kappa = 0.$

We see that there are several possibilities to choose the elements $\alpha,\ \beta,\ \gamma,\ \delta$ of the transition matrix C. All quantum mechanics textbooks use the original one, namely, $\widetilde{b}_{12} = 0$ and $\widetilde{b}_{21} = 0$, due to Darwin [48] and Gordon [83]; cf. Eqs. (3.106) and (3.107) below. Nikiforov and Uvarov [138] take another path, they choose $\widetilde{c}_{12} = 0$

and $\widetilde{c}_{21} = 0$ and show that it is more convenient for taking the nonrelativistic limit $c \to \infty$. These conditions are satisfied if

$$C = \begin{pmatrix} \mu & \nu - \kappa \\ \nu - \kappa & \mu \end{pmatrix}, \tag{3.77}$$

where $\nu = \sqrt{\kappa^2 - \mu^2}$, and we finally arrive at the following system of the first order equations for $v_1(x)$ and $v_2(x)$:

$$v_1' = \left(\frac{\varepsilon\mu}{\nu} - \frac{\nu}{x}\right) v_1 + \left(1 + \frac{\varepsilon\kappa}{\nu}\right) v_2, \tag{3.78}$$

$$v_2' = \left(1 - \frac{\varepsilon\kappa}{\nu}\right) v_1 + \left(\frac{\nu}{x} - \frac{\varepsilon\mu}{\nu}\right) v_2. \tag{3.79}$$

Here

$$\operatorname{Tr} \widetilde{A} = \widetilde{a}_{11} + \widetilde{a}_{22} = 0, \quad \det \widetilde{A} = \varepsilon^2 - 1 + \frac{2\varepsilon\mu}{x} - \frac{\nu^2}{x^2}, \quad \nu^2 = \kappa^2 - \mu^2, \tag{3.80}$$

which is simpler than the original choice in [138, 140]. The corresponding second order differential equations (3.71)–(3.72) become

$$v_1'' + \frac{\left(\varepsilon^2 - 1\right) x^2 + 2\varepsilon\mu x - \nu(\nu + 1)}{x^2} v_1 = 0, \tag{3.81}$$

$$v_2'' + \frac{\left(\varepsilon^2 - 1\right) x^2 + 2\varepsilon\mu x - \nu(\nu - 1)}{x^2} v_2 = 0. \tag{3.82}$$

They are the generalized equations of hypergeometric type (2.1) of a simplest form $\widetilde{\tau} = 0$, thus resembling the one-dimensional Schrödinder equation; the second equation can be obtained from the first one by replacing $\nu \to -\nu$.

Let $1 + \varepsilon\kappa/\nu = 0$, then $\varepsilon = -\nu/\kappa$ that is possible only if $\kappa < 0$, since $\nu > 0$ and $\varepsilon > 0$. The corresponding solution of (3.78),

$$v_1(x) = C_1 x^{-\nu} e^{(\varepsilon\mu\, x)/\nu},$$

satisfies the conditions of the problem only if $C_1 = 0$. Then from (3.79)

$$v_2(x) = C_2 x^{\nu} e^{-(\varepsilon\mu\, x)/\nu},$$

which does satisfy the condition of the problem with $C_2 \neq 0$.

Let us analyze the behavior of the solutions of (3.81) as $x \to 0$. Since

$$\left|\left(\varepsilon^2 - 1\right) x^2 + 2\varepsilon\mu x\right| \ll \nu(\nu + 1)$$

as $x \to 0$, one can approximate this equation in the neighborhood of $x = 0$ by the corresponding Euler equation

$$x^2 v_1'' - \nu(\nu+1) v_1 = 0,$$

whose solutions are

$$v_1(x) = C_1 x^{\nu+1} + C_2 x^{-\nu}, \qquad C_2 = 0.$$

Thus $v_1 \to C_1 x^{\nu+1}$ as $x \to 0$. The results for (3.82) are similar: $v_2 \to C_2 x^{\nu}$ as $x \to 0$; one can use the symmetry $\nu \to -\nu$.

Equation (3.81) is the generalized equation of hypergeometric type (2.1) with

$$\sigma(x) = x, \qquad \widetilde{\tau}(x) = 0,$$
$$\widetilde{\sigma}(x) = \left(\varepsilon^2 - 1\right) x^2 + 2\varepsilon\mu x - \nu(\nu+1).$$

The substitution

$$v_1 = \varphi(x)\, y(x), \qquad \frac{\varphi'}{\varphi} = \frac{\pi(x)}{\sigma(x)}, \tag{3.83}$$

where

$$\pi(x) = \frac{\sigma' - \widetilde{\tau}}{2} \pm \sqrt{\left(\frac{\sigma' - \widetilde{\tau}}{2}\right)^2 - \widetilde{\sigma} + k\sigma} \tag{3.84}$$

with $k = \lambda - \pi'$ and $\tau(x) = \widetilde{\tau}(x) + 2\pi(x)$, results in the equation of hypergeometric type

$$\sigma(x)\, y'' + \tau(x)\, y' + \lambda y = 0 \tag{3.85}$$

by the method of [138]; see also Sect. 2.1. From the four possible forms of $\pi(x)$:

$$\pi(x) = \frac{1}{2} \pm \left(\sqrt{1-\varepsilon^2}\, x \pm \left(\nu + \frac{1}{2}\right)\right), \tag{3.86}$$

corresponding to the values of k determined by the condition of the zero discriminant of the quadratic polynomial under the square root sign in (3.84):

$$k - 2\varepsilon\mu = \pm\sqrt{1-\varepsilon^2}\, (2\nu+1), \tag{3.87}$$

we select the one when the function $\tau(x)$ has a negative derivative and a zero on $(0, +\infty)$. This is true if one chooses

$$\begin{aligned} k &= 2\varepsilon\mu - a(2\nu+1), \\ \pi(x) &= \nu + 1 - ax, \\ \tau(x) &= 2\pi(x) = 2(\nu+1-ax), \\ \lambda &= k + \pi' = 2(\varepsilon\mu - a(\nu+1)) \end{aligned}$$

and

$$\varphi(x) = x^{\nu+1}e^{-ax}, \qquad \rho(x) = x^{2\nu+1}e^{-2ax},$$

where $a = \sqrt{1-\varepsilon^2}$ and $\nu = \sqrt{\kappa^2 - \mu^2}$. The analysis for (3.82) is similar, one can use the symmetry $\nu \to -\nu$ in (3.86) and (3.87).

From (3.56) and (3.62)

$$\int_0^\infty r^2 \left(F^2(r) + G^2(r)\right) dr = \beta^{-3}\int_0^\infty \left(u_1^2(x) + u_2^2(x)\right) dx = 1. \tag{3.88}$$

It requires by (3.68) the square integrability of $v_1(x)$ and $v_2(x)$. Their boundness at $x = 0$ follows from the asymptotic behavior as $x \to 0$. So

$$\int_0^\infty v_1^2(x)\ dx = \int_0^\infty \varphi^2(x)\, y^2(x)\ dx = \int_0^\infty x\ y^2(x)\,\rho(x)\,dx < \infty. \tag{3.89}$$

For the time being, we replace this condition by

$$\int_0^\infty y^2(x)\,\rho(x)\,dx < \infty \tag{3.90}$$

in order to apply Theorem 2.1, and will verify the normalization condition (3.88) later. Then the corresponding energy levels $\varepsilon = \varepsilon_n$ are determined by

$$\lambda + n\tau' + \frac{1}{2}n(n-1)\sigma'' = 0 \qquad (n = 0, 1, 2, \ \ldots), \tag{3.91}$$

whence

$$\varepsilon\mu = a(\nu + n + 1), \tag{3.92}$$

and the eigenfunctions are given by the Rodrigues formula

$$y_n(x) = \frac{C_n}{\rho(x)}\left(\sigma^n(x)\,\rho(x)\right)^{(n)} = C_n\ x^{-2\nu-1}e^{2ax}\frac{d^n}{dx^n}\left(x^{2\nu+n+1}e^{-2ax}\right). \tag{3.93}$$

The functions $y_n(x)$ are, up to certain constants, Laguerre polynomials $L_n^{2\nu+1}(\xi)$ with $\xi = 2ax$.

The previously found eigenvalue $\varepsilon = -\nu/\kappa$ satisfies (3.92) with $n = -1$. Consequently it is natural to replace n by $n - 1$ in (3.92)–(3.93) and define the eigenvalues by

$$\varepsilon\mu = a(\nu + n), \qquad a = \sqrt{1-\varepsilon^2} \qquad (n = 0, 1, 2, \ \ldots). \tag{3.94}$$

The corresponding eigenfunctions have the form

$$v_1(x) = \begin{cases} 0, & n = 0, \\ A_n \xi^{\nu+1} e^{-\xi/2} L_{n-1}^{2\nu+1}(\xi), & n = 1, 2, 3, \ \ldots . \end{cases} \tag{3.95}$$

They are square integrable functions on $(0, \infty)$. The counterparts are

$$v_2(x) = B_n \xi^{\nu} e^{-\xi/2} L_n^{2\nu-1}(\xi), \qquad n = 0, 1, 2, \ \ldots . \tag{3.96}$$

It is easily seen that our previous solution for $\varepsilon = -\nu/\kappa$ is included in this formula when $n = 0$. By Eq. (3.78) the other solutions can be obtained as

$$v_2(x) = \frac{1}{1 + \kappa\varepsilon/\nu} \left(v_1'(x) + \left(\frac{\nu}{x} - \frac{\varepsilon\mu}{\nu} \right) v_1(x) \right)$$

and substituting $v_1(x)$ from (3.95) one gets

$$v_2(x) = \xi^{\nu} e^{-\xi/2} Y(\xi),$$

where $Y(\xi)$ is a polynomial of degree n. But function $v_2(x)$ satisfies (3.82). By the previous consideration the substitution

$$v_2(x) = x^{\nu} e^{-ax} y(x)$$

gives

$$xy'' + (2\nu - 2ax) y' + 2any = 0,$$

in view of the quantization rule (3.94). The change of the variable $y(x) = Y(\xi)$ with $\xi = 2ax$ results in

$$\xi Y'' + (2\nu - \xi) Y' + nY = 0 \tag{3.97}$$

and the only polynomial solutions are the Laguerre polynomials $L_n^{2\nu-1}(\xi)$, whence (3.96) is correct. Solutions $v_2(x)$ are square integrable functions on $(0, \infty)$.

To find the relations between the coefficients A_n and B_n in (3.95) and (3.96) we take the limit $x \to 0$ in (3.78) with the help of the following properties of the Laguerre polynomials [138, 139, 184]:

$$\frac{d}{d\xi} L_n^{\alpha}(\xi) = -L_{n-1}^{\alpha+1}(\xi), \qquad L_n^{\alpha}(0) = \frac{\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+1)}. \tag{3.98}$$

The result is

$$2a(\nu+1) A_n L_{n-1}^{2\nu+1}(0) = -2a\nu A_n L_{n-1}^{2\nu+1}(0) + \left(1 + \frac{\varepsilon\kappa}{\nu}\right) B_n L_n^{2\nu-1}(0),$$

whence

$$A_n = \frac{\nu + \varepsilon\kappa}{an(n+2\nu)} B_n \qquad (n = 1, 2, 3, \ \dots).$$

Since

$$\begin{aligned} a^2 n(n+2\nu) &= a^2\left((n+\nu)^2 - \nu^2\right) = \mu^2\varepsilon^2 - a^2\nu^2 \\ &= \mu^2\varepsilon^2 - \left(1-\varepsilon^2\right)\nu^2 = \kappa^2\varepsilon^2 - \nu^2, \end{aligned}$$

we have proved the useful identity

$$a^2 n(n+2\nu) = \varepsilon^2\kappa^2 - \nu^2, \tag{3.99}$$

and the final relation is

$$A_n = \frac{a}{\kappa\varepsilon - \nu} B_n. \tag{3.100}$$

By (3.68) and (3.77) we find

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = C^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \qquad C^{-1} = \frac{1}{2\nu(\kappa-\nu)} \begin{pmatrix} \mu & \kappa-\nu \\ \kappa-\nu & \mu \end{pmatrix}.$$

Therefore

$$xf(x) = \frac{B_n}{2\nu(\kappa-\nu)} \xi^{\nu} e^{-\xi/2} \left(f_1 \xi L_{n-1}^{2\nu+1}(\xi) + f_2 L_n^{2\nu-1}(\xi)\right), \tag{3.101}$$

$$xg(x) = \frac{B_n}{2\nu(\kappa-\nu)} \xi^{\nu} e^{-\xi/2} \left(g_1 \xi L_{n-1}^{2\nu+1}(\xi) + g_2 L_n^{2\nu-1}(\xi)\right), \tag{3.102}$$

where

$$f_1 = \frac{a\mu}{\varepsilon\kappa - \nu}, \quad f_2 = \kappa - \nu, \quad g_1 = \frac{a\left(\kappa - \nu\right)}{\varepsilon\kappa - \nu}, \quad g_2 = \mu. \tag{3.103}$$

These formulas remain valid for $n = 0$; in this case the terms containing $L_{n-1}^{2\nu+1}(\xi)$ have to be taken to be zero. Thus we derive the representation for the radial functions (3.13) up to the constant B_n. The normalization condition (3.88) gives the value of this constant as

$$B_n = a\beta^{3/2}\sqrt{\frac{\left(\kappa - \nu\right)\left(\varepsilon\kappa - \nu\right)n!}{\mu\Gamma\left(n + 2\nu\right)}}. \tag{3.104}$$

This will be verified in Sect. 5.4. Observe that Eq. (3.104) applies when $n = 0$.

The familiar recurrence relations for the Laguerre polynomials (B.4) and (B.5) allow to present the radial functions (3.13) in a traditional form [2, 25, 49] as

$$\begin{aligned} \begin{pmatrix} F\left(r\right) \\ G\left(r\right) \end{pmatrix} &= a^2\beta^{3/2}\sqrt{\frac{n!}{\mu\left(\kappa - \nu\right)\left(\varepsilon\kappa - \nu\right)\Gamma\left(n + 2\nu\right)}}\,\xi^{\nu-1}e^{-\xi/2} \\ &\quad \times \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} L_{n-1}^{2\nu}\left(\xi\right) \\ L_n^{2\nu}\left(\xi\right) \end{pmatrix}, \end{aligned} \tag{3.105}$$

where

$$\begin{aligned} \alpha_1 &= \sqrt{1+\varepsilon}\left(\left(\kappa - \nu\right)\sqrt{1+\varepsilon} + \mu\sqrt{1-\varepsilon}\right), \\ \alpha_2 &= -\sqrt{1+\varepsilon}\left(\left(\kappa - \nu\right)\sqrt{1+\varepsilon} - \mu\sqrt{1-\varepsilon}\right) \end{aligned} \tag{3.106}$$

and

$$\begin{aligned} \beta_1 &= \sqrt{1-\varepsilon}\left(\left(\kappa - \nu\right)\sqrt{1+\varepsilon} + \mu\sqrt{1-\varepsilon}\right), \\ \beta_2 &= \sqrt{1-\varepsilon}\left(\left(\kappa - \nu\right)\sqrt{1+\varepsilon} - \mu\sqrt{1-\varepsilon}\right). \end{aligned} \tag{3.107}$$

A convenient identity holds

$$\left(\left(\kappa - \nu\right)\sqrt{1+\varepsilon} \pm \mu\sqrt{1-\varepsilon}\right)^2 = 2\left(\kappa - \nu\right)\left(\kappa - \nu\varepsilon \pm a\mu\right). \tag{3.108}$$

One can rewrite this representation in terms of the confluent hypergeometric functions (B.1).

3.3.4 Nonrelativistic Limit of the Wave Functions

Throughout this section, we have always used the notation $n = n_r$ for the radial quantum number, which determines the number of zeros of the radial functions in the relativistic Coulomb problem; see (3.13). For the sake of passing to the limit $c \to \infty$ in this section, let us introduce the principal quantum number of the nonrelativistic hydrogen atom as $n = n_r + |\kappa| = n_r + j + 1/2$ and temporarily consider $N = n_r + \nu$ as its "relativistic analog". As $c \to \infty$ one gets

$$\nu = \sqrt{\kappa^2 - \mu^2} = |\kappa| - \frac{\mu^2}{2|\kappa|} - \frac{\mu^4}{8|\kappa|^3} + \mathrm{O}\left(\mu^6\right), \tag{3.109}$$

$$N = n_r + \nu = n_r + |\kappa| - \frac{\mu^2}{2|\kappa|} - \frac{\mu^4}{8|\kappa|^3} + \mathrm{O}\left(\mu^6\right) \tag{3.110}$$

as $\mu = Ze^2/\hbar c \to 0$. As a result, for the discrete energy levels

$$\varepsilon = \left(1 + \frac{\mu^2}{N^2}\right)^{-1/2} \tag{3.111}$$

we arrive at the expansion

$$\frac{E}{mc^2} = 1 - \frac{\mu^2}{2n^2} - \frac{\mu^4}{2n^4}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) + \mathrm{O}\left(\mu^6\right), \quad \mu \to 0. \tag{3.112}$$

in the nonrelativistic limit $c \to \infty$.

In a similar fashion,

$$a = \sqrt{1-\varepsilon^2} = \frac{\mu}{n_r + |\kappa|}\left(1 + \frac{n_r\,\mu^2}{2|\kappa|\left(n_r + |\kappa|\right)^2} + \mathrm{O}\left(\mu^4\right)\right), \tag{3.113}$$

$$\xi = \xi(c) = 2a\,\frac{mc}{\hbar}\,r = \frac{2Ze^2 m}{n\hbar^2} r\left(1 + \mathrm{O}\left(\mu^2\right)\right) \tag{3.114}$$

as $\mu \to 0$, thus giving

$$\lim_{c\to\infty} \xi(c) = \eta = \frac{2Z}{n}\left(\frac{r}{a_0}\right), \qquad a_0 = \frac{\hbar^2}{me^2}. \tag{3.115}$$

Also

$$\kappa - \nu = (\kappa - |\kappa|) + \frac{\mu^2}{2|\kappa|} + \mathrm{O}\left(\mu^4\right), \tag{3.116}$$

$$\varepsilon\kappa - \nu = (\kappa - |\kappa|) + \frac{(n_r + |\kappa|)^2 - \kappa\,|\kappa|}{2\,|\kappa|\,(n_r + |\kappa|)^2}\,\mu^2 + \mathrm{O}\left(\mu^4\right) \tag{3.117}$$

as $\mu \to 0$. This allows to evaluate the nonrelativistic limit of the transition matrix:

$$S = \begin{pmatrix} f_1 & f_2 \\ g_1 & g_2 \end{pmatrix} = \begin{pmatrix} \dfrac{a\mu}{\varepsilon\kappa - \nu} & \kappa - \nu \\ \dfrac{a\,(\kappa - \nu)}{\varepsilon\kappa - \nu} & \mu \end{pmatrix}. \tag{3.118}$$

There are two distinct cases with the end result

$$\psi_\pm = \begin{pmatrix} \mathcal{Y}^\pm F \\ i\mathcal{Y}^\mp G \end{pmatrix} \to \begin{pmatrix} \pm\mathcal{Y}^\pm R \\ 0 \end{pmatrix}, \qquad \mu \to 0. \tag{3.119}$$

Here $R = R_{nl}\,(r)$ are the nonrelativistic radial functions

$$R\,(r) = R_{nl}\,(r) = \frac{2}{n^2}\left(\frac{Z}{a_0}\right)^{3/2}\sqrt{\frac{(n-l-1)!}{(n+l)!}}\,e^{-\eta/2}\eta^l\,L_{n-l-1}^{2l+1}\,(\eta) \tag{3.120}$$

with

$$\eta = \frac{2Z}{n}\left(\frac{r}{a_0}\right), \qquad a_0 = \frac{\hbar^2}{me^2} \tag{3.121}$$

and $\mathcal{Y}^\pm = \mathcal{Y}_{jm}^{(j\pm 1/2)}\,(\mathbf{n})$ are the spinor spherical harmonics (3.12).

Indeed, if $\kappa = |\kappa| = j + 1/2 = l$,

$$S = S_+\,(\mu) = \begin{pmatrix} \dfrac{2\kappa\,(n_r + \kappa)}{n_r\,(n_r + 2\kappa)} + \mathrm{O}\left(\mu^2\right) & \dfrac{\mu^2}{2\,|\kappa|} + \mathrm{O}\left(\mu^4\right) \\ \dfrac{(n_r + \kappa)\ \mu}{n_r\,(n_r + 2\kappa)} + \mathrm{O}\left(\mu^3\right) & \mu \end{pmatrix} \sim \begin{pmatrix} 1 & \mu^2 \\ \mu & \mu \end{pmatrix}$$

as $\mu \to 0$ or

$$\lim_{\mu\to 0} S_+\,(\mu) = \frac{2nl}{n^2 - l^2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.122}$$

In this case $\nu \to l$ and, therefore,

$$\begin{pmatrix} F\,(r) \\ G\,(r) \end{pmatrix} \to \left(\frac{Ze^2m}{\hbar^2}\right)^{3/2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{(n+l)!}}\,\eta^l e^{-\eta/2}\ \eta L_{n-l-1}^{2l+1}\,(\eta) \tag{3.123}$$

in the limit $c \to \infty$ thus giving

$$\psi_+ = \begin{pmatrix} \mathcal{Y}^+ F \\ i\mathcal{Y}^- G \end{pmatrix} \to \begin{pmatrix} \mathcal{Y}^+ R \\ 0 \end{pmatrix}, \qquad \mu \to 0. \tag{3.124}$$

In a similar fashion, when $\kappa = -|\kappa| = -(j + 1/2) = -l - 1$ one gets

$$\psi_- = \begin{pmatrix} \mathcal{Y}^- F \\ i\mathcal{Y}^+ G \end{pmatrix} \to \begin{pmatrix} -\mathcal{Y}^- R \\ 0 \end{pmatrix}, \qquad \mu \to 0 \tag{3.125}$$

due to the corresponding asymptotic form of the transition matrix $S = S_-(\mu)$:

$$S_-(\mu) = \begin{pmatrix} -\dfrac{\mu^2}{2|\kappa|(n_r + |\kappa|)} + \mathrm{O}(\mu^4) & -2|\kappa| + \mathrm{O}(\mu^2) \\ \dfrac{\mu}{n_r + |\kappa|} + \mathrm{O}(\mu^3) & \mu \end{pmatrix} \sim \begin{pmatrix} \mu^2 & 1 \\ \mu & \mu \end{pmatrix} \tag{3.126}$$

as $\mu \to 0$ [138]. This completes the proof of (3.119).

The representation of the radial functions in the form (3.13), due to Nikiforov and Uvarov [138], is well adapted for passing to the nonrelativistic limit since one coefficient of the transition matrix S is much larger than the others as $\mu \to 0$. In the traditional form (3.106) and (3.107), however, there is an overlap of the orders of these coefficients and one has to use the recurrence relations (B.4) and (B.5) in order to obtain the nonrelativistic wave functions as a limiting case of relativistic ones.

4 Symmetry of Quantum Harmonic Oscillators

In this section, we elaborate on a "missing" class of solutions to the time-dependent Schrödinger equation for the simple harmonic oscillator in one dimension [121, 122, 127] and provide an interesting computer-animated feature of these solutions—the phase space oscillations of the electron density and the corresponding probability distribution of the particle linear momentum. As a result, a dynamic visualization of the fundamental Heisenberg Uncertainty Principle [92] is given, for better understanding of quantum mechanics [123, 182].

4.1 Symmetry and "Hidden" Solutions

The time-dependent Schrödinger equation for the linear harmonic oscillator,

$$2i\psi_t + \psi_{xx} - x^2\psi = 0, \tag{4.1}$$

in dimensionless units, has the following multiparameter family of square integrable solutions

$$\psi_n(x,t) = \frac{e^{i\left(\alpha(t)x^2+\delta(t)x+\kappa(t)\right)+i(2n+1)\gamma(t)}}{\sqrt{2^n n!\mu(t)\sqrt{\pi}}}\, e^{-(\beta(t)x+\varepsilon(t))^2/2}\, H_n(\beta(t)\,x+\varepsilon(t)), \tag{4.2}$$

where $H_n(x)$ are the Hermite polynomials and the periodic time-dependent parameters are given by

$$\mu(t) = \mu_0\sqrt{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}, \tag{4.3}$$

$$\alpha(t) = \frac{\alpha_0\cos 2t + \sin 2t\,\left(\beta_0^4 + 4\alpha_0^2 - 1\right)/4}{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}, \tag{4.4}$$

$$\beta(t) = \frac{\beta_0}{\sqrt{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}} = \frac{\beta_0\mu_0}{\mu(t)}, \tag{4.5}$$

$$\gamma(t) = \gamma_0 - \frac{1}{2}\arctan\frac{\beta_0^2\sin t}{2\alpha_0\sin t + \cos t}, \tag{4.6}$$

$$\delta(t) = \frac{\delta_0(2\alpha_0\sin t + \cos t) + \varepsilon_0\beta_0^3\sin t}{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}, \tag{4.7}$$

$$\varepsilon(t) = \frac{\varepsilon_0(2\alpha_0\sin t + \cos t) - \beta_0\delta_0\sin t}{\sqrt{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}}, \tag{4.8}$$

$$\kappa(t) = \kappa_0 + \sin^2 t\,\frac{\varepsilon_0\beta_0^2(\alpha_0\varepsilon_0 - \beta_0\delta_0) - \alpha_0\delta_0^2}{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2} + \frac{1}{4}\sin 2t\,\frac{\varepsilon_0^2\beta_0^2 - \delta_0^2}{\beta_0^4\sin^2 t + (2\alpha_0\sin t + \cos t)^2}. \tag{4.9}$$

(Here, $\mu_0 > 0$, α_0, $\beta_0 \neq 0$, γ_0, δ_0, ε_0, κ_0 are real-valued initial data.) These "missing" solutions can be derived analytically in a unified approach to generalized harmonic oscillators (see, for example, [44, 45, 108, 115, 131] and references therein). They are also verified by a direct substitution with the aid of `Mathematica` computer algebra system [106, 123, 182]. (The simplest special case $\mu_0 = \beta_0 = 1$ and $\alpha_0 = \gamma_0 = \delta_0 = \varepsilon_0 = \kappa_0 = 0$ reproduces the textbook solution obtained by the separation of variables [73, 82, 113, 132]; see also the original Schrödinger papers [157, 158]; and the shape-preserving oscillator evolutions occur when $\alpha_0 = 0$ and $\beta_0 = 1$. More details on the derivation of these formulas can be found in [122, 127, 137]; see also references therein and Sects. 4.4–4.5 below.)

On the other hand, the "dynamic harmonic oscillator states" (4.2)–(4.9) are eigenfunctions,

$$E(t)\,\psi_n(x,t) = \left(n+\frac{1}{2}\right)\psi_n(x,t)\,, \tag{4.10}$$

of the time-dependent quadratic invariant,

$$\begin{aligned} E(t) &= \frac{1}{2}\left[\frac{(p-2\alpha x-\delta)^2}{\beta^2} + (\beta x+\varepsilon)^2\right] \\ &= \frac{1}{2}\left[\widehat{a}(t)\,\widehat{a}^{\dagger}(t) + \widehat{a}^{\dagger}(t)\,\widehat{a}(t)\right], \qquad \frac{d}{dt}\langle E\rangle = 0, \end{aligned} \tag{4.11}$$

with the required operator identity [57, 154]:

$$\frac{\partial E}{\partial t} + i^{-1}[E,H] = 0, \qquad H = \frac{1}{2}\left(p^2+x^2\right). \tag{4.12}$$

Here, the time-dependent annihilation $\widehat{a}(t)$ and creation $\widehat{a}^{\dagger}(t)$ operators are explicitly given by

$$\widehat{a}(t) = \frac{1}{\sqrt{2}}\left(\beta x+\varepsilon+i\frac{p-2\alpha x-\delta}{\beta}\right), \quad \widehat{a}^{\dagger}(t) = \frac{1}{\sqrt{2}}\left(\beta x+\varepsilon-i\frac{p-2\alpha x-\delta}{\beta}\right)$$

with $p = i^{-1}\partial/\partial x$ in terms of the periodic functions (4.4)–(4.9). These operators satisfy the canonical commutation relation,

$$\widehat{a}(t)\,\widehat{a}^{\dagger}(t) - \widehat{a}^{\dagger}(t)\,\widehat{a}(t) = 1, \tag{4.13}$$

and the oscillator-type spectrum (4.10) of the dynamic invariant E can be obtained in a standard way by using the Heisenberg–Weyl algebra of the raising and lowering operators (a "second quantization" [2, 119], the Fock states):

$$\widehat{a}(t)\,\Psi_n(x,t) = \sqrt{n}\,\Psi_{n-1}(x,t)\,, \quad \widehat{a}^{\dagger}(t)\,\Psi_n(x,t) = \sqrt{n+1}\,\Psi_{n+1}(x,t)\,. \tag{4.14}$$

Here,

$$\psi_n(x,t) = e^{i(2n+1)\gamma(t)}\,\Psi_n(x,t) \tag{4.15}$$

is the relation to the wave functions (4.2) with $\varphi_n(t) = -(2n+1)\,\gamma(t)$ being the nontrivial Lewis phase [119, 154].

This quadratic dynamic invariant and the corresponding creation and annihilation operators for the generalized harmonic oscillators have been introduced in [154] (see also [44, 57, 177] and references therein for important special cases). Applications to electromagnetic-field quantization is discussed in [109, 110, 112].

The key ingredients, the maximum kinematical invariance groups of the free particle and harmonic oscillator, were introduced in [4, 5, 87, 97, 136] and [137] (see also [34, 98, 134, 151, 170, 171, 191] and references therein). We establish a (hidden symmetry revealing) connection with certain Ermakov-type system which allows us to bypass a complexity of the traditional Lie algebra approach [122] (see [67, 116] and references therein regarding the Ermakov equation). (A general procedure of obtaining new solutions by acting on any set of given ones by enveloping algebra of generators of the Heisenberg–Weyl group is described in [57].) In addition, the maximal invariance group of the generalized driven harmonic oscillators is shown to be isomorphic to the Schrödinger group of the free particle and the simple harmonic oscillator [122, 136, 137].

4.2 Computer Animations

`Mathematica` source CODE lines are taken from [123].

Example 1 Animation of the dynamic ground state $n = 0$, using $\alpha_0 = \gamma_0 = \varepsilon_0 = 0$, $\beta_0 = 2/3$, $\delta_0 = 1$:

$$\text{In[1]:= } \textbf{Animate}\left[\text{Plot}\left[\left\{\frac{9\sqrt{2}\, e^{-\frac{72\left(x-\text{Sin}\left[\frac{1}{500}\pi(-1+T)\right]\right)^2}{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}}{\sqrt{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}},\ e^{-\frac{4x^2}{9}}\right\}, \{x, -3.5, 3.5\},\right.\right.$$

$$\text{AxesLabel -> } \{x, (\text{Abs}[\psi])\text{^}2\},\ \text{PlotRange} \to \{0, 2.3\},\ \text{Filling} \to \{1 \to \text{Bottom}\},\ \left.\text{PlotStyle}\to \{\text{Thick, Blue}\}\right], \left.\{T, 1001\}\right]$$

Out[1]= Fig. 1

Example 2 The following animation is for the first excited dynamic state $n = 1$, using $\alpha_0 = \gamma_0 = \varepsilon_0 = 0$, $\beta_0 = 2/3$, $\delta_0 = 1$:

$$\text{In[2]:= } \textbf{Animate}\left[\text{Plot}\left[\left\{\left(1296\sqrt{2}\, e^{-\frac{72\left(x-\text{Sin}\left[\frac{1}{500}\pi(-1+T)\right]\right)^2}{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}\right.\right.\right.$$

$$\left.\left(x - \text{Sin}\left[\tfrac{1}{500}\pi(-1+T)\right]\right)^2\right) / \left(97 + 65\,\text{Cos}\left[\tfrac{1}{250}\pi(-1+T)\right]\right)^{3/2},\ \tfrac{8}{9}\, e^{-\frac{4x^2}{9}} x^2\Big\},$$

$$\{x, -4.5, 4.5\},\ \text{AxesLabel ->}\{x, (\text{Abs}[\psi])\text{^}2\},\ \text{PlotRange} \to \{0, 1.67\},$$

$$\left.\text{Filling} \to \{1 \to \text{Bottom}\},\ \text{PlotStyle} \to \{\text{Thick, Blue}\}\right], \left.\{T, 1001\}\right]$$

Out[2]= Fig. 2

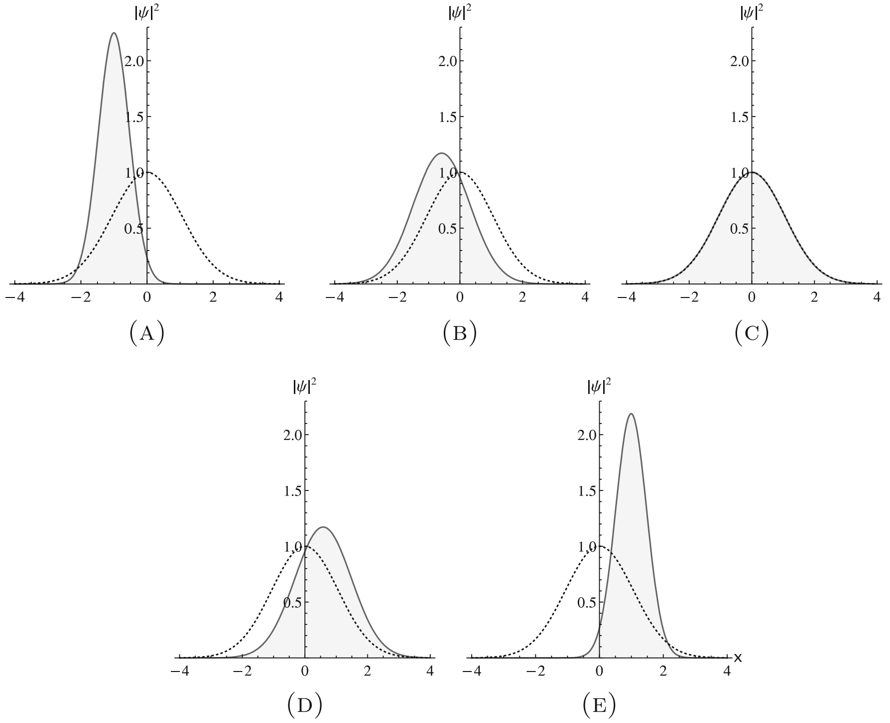


Fig. 1 Subfigures (**a**)–(**e**) are a few *stills* taken from the mathematica movie animation [123]. Starting with (**a**) they denote the oscillating electron density (shaded) of the ground "dynamic harmonic state" of the time-dependent Schrödinger equation (4.1). These space oscillations complement (with the help of `Mathematica`) the corresponding "static" textbook solution (dashed)[157]

Example 3 The following animations simultaneously show the phase space oscillations of the electron density and the momentum probability distribution, according to the Heisenberg Uncertainty Principle, for the dynamic ground state $n = 0$ with parameters $\alpha_0 = \gamma_0 = \varepsilon_0 = \kappa_0 = 0$, $\beta_0 = 2/3$ and $\delta_0 = 3/2$:

$$\text{In[3]:=}\ \mathbf{Animate}\left[\text{Plot}\left[\left\{\frac{9\sqrt{2}\,e^{-\frac{18\left(2x+3\,\text{Cos}\left[\frac{1}{500}\pi(249+T)\right]\right)^2}{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}}{\sqrt{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}},\ \frac{9\sqrt{2}\,e^{\frac{18\left(-2x+3\,\text{Cos}\left[\frac{1}{500}\pi(-1+T)\right]\right)^2\sqrt{97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}{-97+65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}}{\sqrt{97-65\,\text{Cos}\left[\frac{1}{250}\pi(-1+T)\right]}}\right\},\ \{x,-4.5,4.5\},\right.\right.$$

$$\text{AxesLabel -> }\{\{x,p\},\{(\text{Abs}[\psi])\text{^}2,(\text{Abs}[a])\text{^}2\}\},\ \text{PlotRange}\to\{0,2.3\},$$

$$\left.\left.\text{Filling}\to\{1\to\text{Bottom}\},\ \text{PlotStyle}\to\{\text{Thick},\text{Blue}\}\right],\{T,1001\}\right]$$

Out[3]= Fig. 3

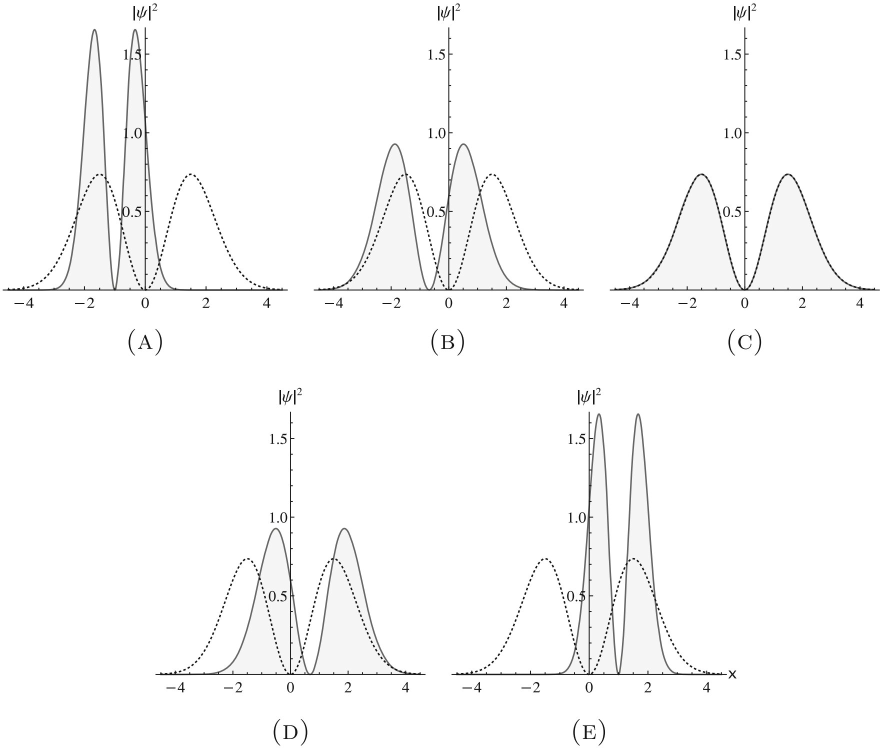


Fig. 2 Subfigures (**a**)–(**e**) are a few *stills* taken from the mathematica movie animation [123]. Starting with (**a**) they denote the oscillating electron density (shaded) of the first excited "dynamic harmonic state" of the time-dependent Schrödinger equation (4.1). These space oscillations complement (with the help of `Mathematica`) the corresponding "static" textbook solution (dashed)[157]

One immediately recognizes from these animations that the particle is the most localized at the turning points when its linear momentum is the least precisely determined, as required by the fundamental Heisenberg Uncertainty Principle [92]—The more precisely the position is determined, the less precisely the momentum is known in this instant, and vice versa (see also [43]). In the creator own words—"If the classical motion of the system is periodic, it may happen that the size of the wave packet at first undergoes only periodic changes" (see [92, p. 38]).

According to the time-dependent form of creation and annihilation operators and (4.14), the corresponding expectation values are given by

$$\langle x \rangle = -\frac{1}{\beta_0}\left[(2\alpha_0\varepsilon_0 - \beta_0\delta_0)\sin t + \varepsilon_0 \cos t\right], \qquad \frac{d}{dt}\langle x \rangle = \langle p \rangle, \tag{4.16}$$

$$\langle p \rangle = -\frac{1}{\beta_0}\left[(2\alpha_0\varepsilon_0 - \beta_0\delta_0)\cos t - \varepsilon_0 \sin t\right], \qquad \frac{d}{dt}\langle p \rangle = -\langle x \rangle \tag{4.17}$$

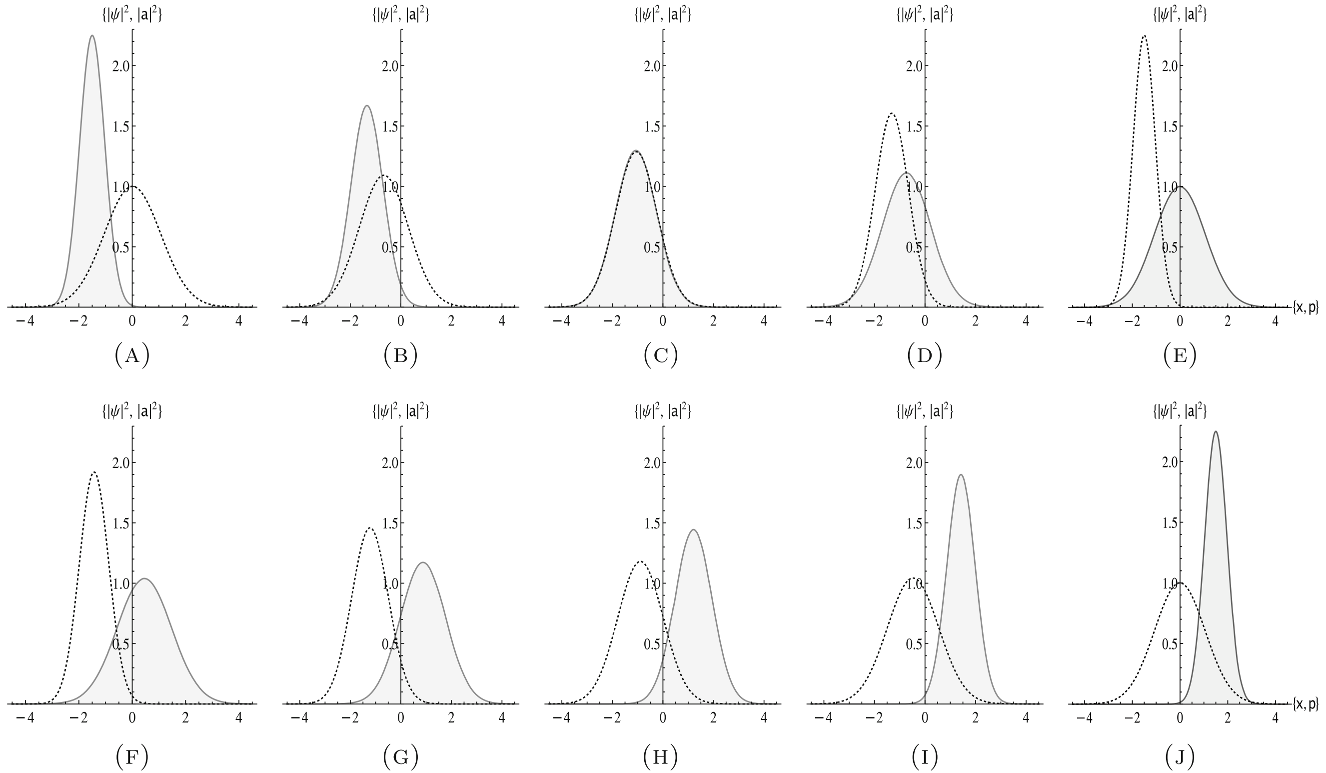


Fig. 3 Subfigures (**a**)–(**j**) are a few *stills* taken from the mathematica movie animation [123]. Starting with (**a**) they denote simultaneous oscillations of electron density (shaded) and probability distribution of momentum (dashed) for the ground "dynamic harmonic state" of the time-dependent Schrödinger equation (4.1). These phase space oscillations complement (with the help of `Mathematica`) the corresponding "static" textbook solutions [82, 157]

with the initial data $\langle x\rangle|_{t=0} = -\varepsilon_0/\beta_0$ and $\langle p\rangle|_{t=0} = -\left(2\alpha_0\varepsilon_0 - \beta_0\delta_0\right)/\beta_0$. This provides a classical interpretation of our "hidden" parameters.

The expectation values $\langle x\rangle$ and $\langle p\rangle$ satisfy the classical equation for harmonic motion, $y'' + y = 0$, with the total mechanical energy

$$\frac{1}{2}\left[\langle p\rangle^2 + \langle x\rangle^2\right] = \frac{(2\alpha_0\varepsilon_0 - \beta_0\delta_0)^2 + \varepsilon_0^2}{2\beta_0^2} = \frac{1}{2}\left[\langle p\rangle^2 + \langle x\rangle^2\right]\Big|_{t=0}. \tag{4.18}$$

For the standard deviations,

$$\langle(\Delta p)^2\rangle = \left(n + \tfrac{1}{2}\right)\frac{1 + 4\alpha_0^2 + \beta_0^4 + \left(4\alpha_0^2 + \beta_0^4 - 1\right)\cos 2t - 4\alpha_0 \sin 2t}{2\beta_0^2}, \tag{4.19}$$

$$\langle(\Delta x)^2\rangle = \left(n + \tfrac{1}{2}\right)\frac{1 + 4\alpha_0^2 + \beta_0^4 - \left(4\alpha_0^2 + \beta_0^4 - 1\right)\cos 2t + 4\alpha_0 \sin 2t}{2\beta_0^2}, \tag{4.20}$$

one gets

$$\begin{aligned}\langle(\Delta p)^2\rangle\langle(\Delta x)^2\rangle = \left(n + \frac{1}{2}\right)^2 \frac{1}{4\beta_0^4}\Big[&\left(1 + 4\alpha_0^2 + \beta_0^4\right)^2 \\ &- \left(\left(4\alpha_0^2 + \beta_0^4 - 1\right)\cos 2t - 4\alpha_0 \sin 2t\right)^2\Big].\end{aligned} \tag{4.21}$$

In the case of the Schrödinger solution [157, 158] when $\alpha_0 = \delta_0 = \varepsilon_0 = 0$ and $\beta_0 = 1$, we arrive at $\langle x\rangle = \langle p\rangle \equiv 0$ and

$$\langle(\Delta p)^2\rangle = \langle(\Delta x)^2\rangle = n + \frac{1}{2} \tag{4.22}$$

as presented in the textbooks [73, 82, 84, 94, 113, 132]. The dependence on the quantum number n, which disappears from the Ehrenfest theorem [63, 92], is coming back at the level of the higher moments of the distribution.

According to (4.21),

$$\langle(\Delta p)^2\rangle\langle(\Delta x)^2\rangle = \left(n + \frac{1}{2}\right)^2 \frac{1 - 4\alpha_0^2 \sin^2 2t}{\beta_0^4}, \tag{4.23}$$

provided that $4\alpha_0^2 + \beta_0^4 = 1$, and the product is equal to $1/4$, if $n = 0$ and $\sin^2 2t = 1$. These are conditions for the minimum-uncertainty squeezed states of the simple harmonic oscillator (see, for example, [94, 110]). For the coherent states $\alpha_0 = 0$ and $\beta_0 = 1$, which describes a two-parameter family with the initial data $\langle x\rangle|_{t=0} = -\varepsilon_0$ and $\langle p\rangle|_{t=0} = \delta_0$.

The corresponding wave functions in the momentum representation are derived by the (inverse) Fourier transform of our solutions (4.2) and (4.3)–(4.9). Moreover, below we explicitly present the action of the Schrödinger group on the wave functions of harmonic oscillators and elaborate on the corresponding eigenfunction expansion for the sake of 'completeness'[110]. More examples are available in the authors' websites.

4.3 The Momentum Representation

For the wave functions in the momentum representation,

$$a_n(p,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ipx} \psi_n(x,t)\ dx, \tag{4.24}$$

the integral evaluation is similar to Ref. [115]. As a result, the functions $a_n(p,t)$ are of the same form (4.2)–(4.9), if $\psi_n \to a_n$ and $x \to p$, with the initial data

$$\alpha_1 = -\frac{\alpha_0}{4\alpha_0^2+\beta_0^2}, \qquad \beta_1 = \frac{\beta_0}{\sqrt{4\alpha_0^2+\beta_0^2}}, \tag{4.25}$$

$$\gamma_1 = \gamma_0 + \frac{1}{2}\operatorname{arccot}\frac{\beta_0^2}{2\alpha_0}, \quad \mu_1 = \mu_0\sqrt{4\alpha_0^2+\beta_0^2}, \tag{4.26}$$

$$\delta_1 = \frac{2\alpha_0\delta_0+\beta_0^3\varepsilon_0}{4\alpha_0^2+\beta_0^2}, \qquad \varepsilon_1 = \frac{2\alpha_0\varepsilon_0-\beta_0\delta_0}{\sqrt{4\alpha_0^2+\beta_0^2}}, \tag{4.27}$$

$$\kappa_1 = \kappa_0 + \frac{\alpha_0\left(\beta_0^2\varepsilon_0^2-\delta_0^2\right)+\beta_0^3\delta_0\varepsilon_0}{4\alpha_0^2+\beta_0^2}. \tag{4.28}$$

The calculation details are left to the reader [182] (see, for example, [82] for the classical case).

4.4 The Schrödinger Group for Simple Harmonic Oscillators

The following substitution

$$\psi(x,t) = \frac{e^{i\left(\alpha(t)x^2+\delta(t)x+\kappa(t)\right)}}{\sqrt{\mu(t)}}\,\chi(\xi,\tau), \tag{4.29}$$

where relations (4.3)–(4.9) hold, transforms the time-dependent Schrödinger equation (4.1) into itself with respect to the new variables $\xi = \beta(t)\,x + \varepsilon(t)$ and

$\tau = -\gamma(t)$ [137] (see also [122] and references therein). A `Mathematica` verification can be found in [106] and [182].

The eigenfunction expansion of the “dynamic harmonic states” with respect to the standard “static” ones can be obtain in an obvious way (see, for example, [114, 120] for similar integral evaluations and [110] for the details). The corresponding matrix elements define the representation of the Schrödinger group acting on the oscillator wave functions. (The structure of the Schrödinger group in two-dimensional space-time as a semidirect product of $SL(2, \mathbb{R})$ and Weyl $W(1)$ groups is discussed, for example, in Refs. [34, 98] and [134].)

An explicit time evolution of the (bosonic field) creation and annihilation operators for the “dynamic harmonic (Fock) states” (with the embedded hidden Schrödinger group symmetry) can be easily derived from (4.1), (4.14) and (4.15). Applications to the quantization of electromagnetic fields are discussed in [109, 110, 112].

4.5 A Complex Parametrization of the Schrödinger Group

The Ansatz

$$\psi(x,t) = \sqrt{\beta(t)}e^{iS(x,t)}\chi(\xi,\tau), \qquad S = \alpha(t)x^2 + \delta(t)x + \kappa(t), \tag{4.30}$$

where relations (4.4)–(4.9) hold, transforms the time-dependent Schrödinger equation (4.1) into itself:

$$2i\psi_t + \psi_{xx} - x^2\psi = e^{iS}\beta^{5/2}\left(2i\chi_\tau + \chi_{\xi\xi} - \xi^2\chi\right) = 0 \tag{4.31}$$

with respect to the new variables $\xi = \beta(t)x + \varepsilon(t)$ and $\tau = -\gamma(t)$. This transformation is known as the Schrödinger group for linear harmonic oscillator [137] (see also the previous subsection).

Let us introduce the following complex-valued function:

$$z = c_1e^{it} + c_2e^{-it}, \qquad z'' + z = 0, \tag{4.32}$$

where by definition

$$c_1 = \left(1+\beta_0^2\right)/2 - i\alpha_0, \qquad c_2 = \left(1-\beta_0^2\right)/2 + i\alpha_0 \tag{4.33}$$

$$\left(c_1 + c_2 = 1, \qquad |c_1|^2 - |c_2|^2 = \beta_0^2\right),$$

and

$$c_3 = \frac{\delta_0}{\beta_0} - i\varepsilon_0. \tag{4.34}$$

Then Eqs. (4.4)–(4.9) can be rewritten in a compact form in terms of our complex-valued parameters $c_1,\ c_2$, and c_3. Indeed, with the help of identities

$$\frac{\delta}{\beta}+i\varepsilon=\left(\frac{\delta_0}{\beta_0}+i\varepsilon_0\right)e^{2i(\gamma-\gamma_0)}, \tag{4.35}$$

$$\delta-\frac{2\alpha\varepsilon}{\beta}+i\frac{\varepsilon}{\beta}=\left(\delta_0-\frac{2\alpha_0\varepsilon_0}{\beta_0}+i\frac{\varepsilon_0}{\beta_0}\right)e^{-it}, \tag{4.36}$$

$$\frac{1-\beta^2}{2}+i\alpha=e^{-it}\left(\frac{1-\beta_0^2}{2}+i\alpha_0\right)/\left(2\alpha_0\sin t+\cos t+i\beta_0^2\sin t\right), \tag{4.37}$$

$$\frac{1+\beta^2}{2}-i\alpha=e^{it}\left(\frac{1+\beta_0^2}{2}-i\alpha_0\right)/\left(2\alpha_0\sin t+\cos t+i\beta_0^2\sin t\right) \tag{4.38}$$

one gets

$$|z|=\left(|c_1|^2+c_1c_2^*e^{2it}+c_1^*c_2e^{-2it}+|c_2|^2\right)^{1/2} \tag{4.39}$$

and

$$\alpha=i\frac{c_1c_2^*e^{2it}-c_1^*c_2e^{-2it}}{2\,|z|^2}, \tag{4.40}$$

$$\beta=\frac{\beta_0}{|z|}=\pm\frac{\sqrt{|c_1|^2-|c_2|^2}}{|z|}, \tag{4.41}$$

$$\gamma=\gamma_0-\frac{1}{2}\arg z, \tag{4.42}$$

$$\delta=\frac{\beta_0}{2\,|z|}\left(c_3e^{i\arg z}+c_3^*e^{-i\arg z}\right), \tag{4.43}$$

$$\varepsilon=\frac{i}{2}\left(c_3e^{i\arg z}-c_3^*e^{-i\arg z}\right), \tag{4.44}$$

$$\kappa=\kappa_0-\frac{i}{8}\left[c_3^2\left(1-e^{2i\arg z}\right)-c_3^{*\,2}\left(1-e^{-2i\arg z}\right)\right]. \tag{4.45}$$

The inverse relations between the essential, real and complex, parameters are given by

$$\alpha_0=\frac{i}{2}\left(c_1c_2^*-c_1^*c_2\right),\qquad \beta_0=\pm\sqrt{|c_1|^2-|c_2|^2}, \tag{4.46}$$

$$\delta_0=\pm\frac{1}{2}\sqrt{|c_1|^2-|c_2|^2}\left(c_3+c_3^*\right),\qquad \varepsilon_0=\frac{i}{2}\left(c_3-c_3^*\right). \tag{4.47}$$

These formulas (4.40)–(4.45) provide a complex parametrization of the Schrödinger group for the simple harmonic oscillator originally found in [137] (see also [121, 122] and references therein; see [88] and [109] for an extension to generalized harmonic oscillators).

4.6 Discussion

Quantum systems with quadratic Hamiltonians (see, for example, [3, 26, 27, 44, 57, 58, 68, 72, 88, 124, 200, 203, 206, 207] and references therein) have attracted substantial attention over the years because of their great importance in many advanced quantum problems. Examples are coherent and squeezed states, uncertainty relations, Berry's phase, quantization of mechanical systems and Hamiltonian cosmology. More applications include, but are not limited to charged particle traps and motion in uniform magnetic fields, molecular spectroscopy and polyatomic molecules in varying external fields, crystals through which an electron is passing and exciting the oscillator modes, and other mode interactions with external fields. Quadratic Hamiltonians have particular applications in quantum electrodynamics because the electromagnetic field can be represented as a set of generalized driven harmonic oscillators [59, 72].

The maximal kinematical invariance group of the simple harmonic oscillator [137] provides the multiparameter family of solutions, namely (4.2) and (4.3)–(4.9), for an arbitrary choice of the initial data (of the corresponding Ermakov-type system [67, 115, 116, 122]). These "hidden parameters", that are explicitly visible in the wave function, usually disappear after evaluation of matrix elements from the spectrum. How to distinguish between these "new dynamic" and the "standard static" harmonic oscillator states (and which of them is realized in a particular measurement) is thus a fundamental problem. However, these oscillator states can be detected in cold ion trap experiments [90, 117] and in the measuring the quantum states of light [118]. A similar effect of proton beam (super)focusing in a thin monocrystal film was predicted in [51, 52]; see also [110].

At the same time, the probability density $|\psi(x,t)|^2$ of the solution (4.2) is obviously moving with time, somewhat contradicting to the standard textbooks [73, 82, 113, 132, 157, 158]—an elementary `Mathematica` simulation reveals such space oscillations for the simplest "dynamic oscillator states" [122, 123]. The same is true for the probability distribution of the particle linear momentum due to the Heisenberg Uncertainty Principle [92]. These effects, quite possibly, can be observed experimentally, say in Bose condensates, if the nonlinearity of the Gross–Pitaevskii equation is turned off by the Feshbach resonance [46, 69, 102, 144, 169, 180]. A more elementary example is an electron moving in a uniform magnetic field. By slowly changing the magnetic field, say, from an initially occupied Landau level with the standard solution [113, 120], one may continuously follow the initial wave function evolution (with the quadratic invariant) until the magnetic field becomes a constant once again (a parametric excitation; see, for example, [45, 57, 114, 124] and the references therein). The terminal state will have, in general, the initial conditions that are required for the "dynamic harmonic states" (4.2)–(4.9) and the probability density should oscillate on the corresponding Landau level just as our solution predicts. However, it is still not clear how to observe this effect experimentally (but these "dynamic harmonic states" will have a nontrivial Berry's phase [26, 27, 154, 181, 182]).

One may imagine other possible applications, for example, in molecular spectroscopy [124], theory of crystals, quantum optics [84, 206], and cavity quantum electrodynamics [56, 59, 77, 109, 205]. We believe in a dynamic character of the nature [93]. All of that puts the consideration of this section into a much broader mathematical and physical context—This may help better understand some intriguing features of quantum motion and will be useful for pedagogy.

5 Expectation Values in Relativistic Coulomb Problems

Recent experimental progress has renewed interest in quantum electrodynamics of atomic hydrogenlike systems. Experimentalists and theorists in atomic and particle physics are discovering problems of common interest with new ideas and methods. A current account of the status of this fundamental area of quantum physics, which is more than a century old, is given in [99, 100, 135, 165]. Exciting research topics vary from experimental testing of Quantum Electrodynamics (QED) to fruitful training models for the bound-state Quantum Chromodynamics and Bose-Einstein Condensation [47, 99–101, 135, 165, 204].

The highly charged ions are an ideal testing ground for the strong-field bound-state QED. They posses a strong static Coulomb field of the nucleus and a simple electronic structure which can be accurately computed from first principles. It is possible nowadays to make massive highly charged ions with a strong nuclear charge and only one electron through the periodic table up to uranium, the most highly charged ion [85, 86]. These systems are truly relativistic and require the Dirac wave equation as a starting point in a detailed investigation of their spectra [135, 163]. The binding energy of a single K-shell electron in the electric field of a uranium nucleus corresponds to roughly one third of the electron rest mass. For the simple hydrogen atom the nonrelativistic Schrödinger approximation can be used [28].

In the last decade, the two-time Green's function method of deriving formal expressions for the energy shift of a bound-state level of high-Z few-electron systems was developed [163] and numerical calculations of QED effects in heavy ions were performed with excellent agreement to current experimental data [85, 86] (see [161, 162, 165–167, 204] and references therein for more details). These advances motivate, among other technical things, evaluation of the expectation values $\langle Or^p \rangle$ for the standard Dirac matrix operators $O = \{1, \beta, i\boldsymbol{\alpha}\mathbf{n}\beta\}$ between the bound-state relativistic Coulomb wave functions. Special cases appear in calculations of the magnetic dipole hyperfine splitting, the electric quadrupole hyperfine splitting, the anomalous Zeeman effect, and the relativistic recoil corrections in hydrogenlike ions (see, for example, [1, 162, 164] and references therein). We discuss convenient closed forms of these integrals in general and derive matrix symmetry relations among them which can be useful in the theory of relativistic Coulomb systems.

In this section, we evaluate the matrix elements $\langle Or^{p}\rangle$, where $O=\{1,\beta,i\boldsymbol{\alpha}\mathbf{n}\beta\}$ are the standard Dirac matrix operators and the angular brackets denote the quantum-mechanical average for the relativistic Coulomb problem, in terms of generalized hypergeometric functions ${}_3F_2\left(1\right)$ for all suitable powers. Their connections with the Chebyshev and Hahn polynomials of a discrete variable are emphasized. As a result, we derive two sets of Pasternack-type matrix identities for these integrals, when $p\to -p-1$ and $p\to -p-3$, respectively [176].

5.1 Evaluation of the Matrix Elements

We evaluate the following integrals of the radial functions:

$$A_{p}=\int_{0}^{\infty}r^{p+2}\left(F^{2}\left(r\right)+G^{2}\left(r\right)\right)\,dr, \tag{5.1}$$

$$B_{p}=\int_{0}^{\infty}r^{p+2}\left(F^{2}\left(r\right)-G^{2}\left(r\right)\right)\,dr, \tag{5.2}$$

$$C_{p}=\int_{0}^{\infty}r^{p+2}F\left(r\right)G\left(r\right)\,dr \tag{5.3}$$

in terms of generalized hypergeometric series. (Equation (C.4) of the Appendix C establish their relations with the expectation values $\langle Or^{p}\rangle$, where $O=\{1,\beta,i\boldsymbol{\alpha}\mathbf{n}\beta\}$, respectively.) The final results with the notations from Sect. 3 can be presented in two different closed forms. Use of the traditional radial functions (3.19) results in:

$$\begin{aligned}2\mu\left(2a\beta\right)^{p}\,\frac{\Gamma\left(2\nu+1\right)}{\Gamma\left(2\nu+p+1\right)}\,A_{p}&=2p\varepsilon an\;{}_3F_2\left(\begin{matrix}1-n,\;-p,\;p+1\\ 2\nu+1,\quad 2\end{matrix}\right)\\&+\left(\mu+a\kappa\right)\;{}_3F_2\left(\begin{matrix}1-n,\;-p,\;p+1\\ 2\nu+1,\quad 1\end{matrix}\right)+\;\left(\mu-a\kappa\right)\;{}_3F_2\left(\begin{matrix}-n,\;-p,\;p+1\\ 2\nu+1,\quad 1\end{matrix}\right),\end{aligned} \tag{5.4}$$

$$\begin{aligned}2\mu\left(2a\beta\right)^{p}\,\frac{\Gamma\left(2\nu+1\right)}{\Gamma\left(2\nu+p+1\right)}\,B_{p}&=2pan\;{}_3F_2\left(\begin{matrix}1-n,\;-p,\;p+1\\ 2\nu+1,\quad 2\end{matrix}\right)\\&+\varepsilon\left(\mu+a\kappa\right)\;{}_3F_2\left(\begin{matrix}1-n,\;-p,\;p+1\\ 2\nu+1,\quad 1\end{matrix}\right)+\;\varepsilon\left(\mu-a\kappa\right)\;{}_3F_2\left(\begin{matrix}-n,\;-p,\;p+1\\ 2\nu+1,\quad 1\end{matrix}\right),\end{aligned} \tag{5.5}$$

$$4\mu\left(2a\beta\right)^{p}\ \frac{\Gamma\left(2\nu+1\right)}{\Gamma\left(2\nu+p+1\right)}\ C_{p} \tag{5.6}$$
$$= a\left(\mu+a\kappa\right)\ {}_3F_2\left(\begin{array}{c} 1-n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right) - a\left(\mu-a\kappa\right)\ {}_3F_2\left(\begin{array}{c} -n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right).$$

Nikiforov and Uvarov's form (3.13) gives the following result:

$$4\mu\nu^{2}\left(2a\beta\right)^{p}\ A_{p} \tag{5.7}$$
$$= a\kappa\left(\varepsilon\kappa+\nu\right)\frac{\Gamma\left(2\nu+p+3\right)}{\Gamma\left(2\nu+2\right)}\ {}_3F_2\left(\begin{array}{c} 1-n,\ p+2,\ -p-1 \\ 2\nu+2,\quad 1 \end{array}\right)$$
$$-2\left(p+2\right)a^{2}\mu n\frac{\Gamma\left(2\nu+p+2\right)}{\Gamma\left(2\nu+1\right)}\ {}_3F_2\left(\begin{array}{c} 1-n,\ p+2,\ -p-1 \\ 2\nu+1,\quad 2 \end{array}\right)$$
$$+a\kappa\left(\varepsilon\kappa-\nu\right)\frac{\Gamma\left(2\nu+p+1\right)}{\Gamma\left(2\nu\right)}\ {}_3F_2\left(\begin{array}{c} -n,\ p+2,\ -p-1 \\ 2\nu,\quad 1 \end{array}\right),$$

$$4\mu\nu\left(2a\beta\right)^{p}\ B_{p} \tag{5.8}$$
$$= a\left(\varepsilon\kappa+\nu\right)\frac{\Gamma\left(2\nu+p+3\right)}{\Gamma\left(2\nu+2\right)}\ {}_3F_2\left(\begin{array}{c} 1-n,\ p+2,\ -p-1 \\ 2\nu+2,\quad 1 \end{array}\right)$$
$$-a\left(\varepsilon\kappa-\nu\right)\frac{\Gamma\left(2\nu+p+1\right)}{\Gamma\left(2\nu\right)}\ {}_3F_2\left(\begin{array}{c} -n,\ p+2,\ -p-1 \\ 2\nu,\quad 1 \end{array}\right),$$

$$8\mu\nu^{2}\left(2a\beta\right)^{p}\ C_{p} \tag{5.9}$$
$$= a\mu\left(\varepsilon\kappa+\nu\right)\frac{\Gamma\left(2\nu+p+3\right)}{\Gamma\left(2\nu+2\right)}\ {}_3F_2\left(\begin{array}{c} 1-n,\ p+2,\ -p-1 \\ 2\nu+2,\quad 1 \end{array}\right)$$
$$-2\left(p+2\right)a^{2}\kappa n\frac{\Gamma\left(2\nu+p+2\right)}{\Gamma\left(2\nu+1\right)}\ {}_3F_2\left(\begin{array}{c} 1-n,\ p+2,\ -p-1 \\ 2\nu+1,\quad 2 \end{array}\right)$$
$$+a\mu\left(\varepsilon\kappa-\nu\right)\frac{\Gamma\left(2\nu+p+1\right)}{\Gamma\left(2\nu\right)}\ {}_3F_2\left(\begin{array}{c} -n,\ p+2,\ -p-1 \\ 2\nu,\quad 1 \end{array}\right).$$

Here, the terminating generalized hypergeometric series ${}_3F_2\left(1\right)$ are related to the Hahn and Chebyshev polynomials of a discrete variable [139, 183]. (See Eq. (5.11) below, we usually omit the argument of the hypergeometric series ${}_3F_2$ if it is equal to 1.) Two more forms occur if one takes one of the radial wave functions from (3.13) and another one from (3.19). We leave the details to the reader.

The averages of r^{p} for the relativistic hydrogen atom were evaluated by Davis [49] in a form which is slightly different from our Eqs. (5.4) and (5.7); see also [6] and [183] for a simple proof of the second formula including evaluation of the corresponding integral of the product of two Laguerre polynomials (Appendix A below):

$$\int_0^\infty e^{-x} x^{\alpha+s}\, L_n^{\alpha}(x)\, L_m^{\beta}(x)\; dx \tag{5.10}$$
$$= (-1)^{n-m} \frac{\Gamma(\alpha+s+1)\,\Gamma(\beta+m+1)\,\Gamma(s+1)}{m!\,(n-m)!\,\Gamma(\beta+1)\,\Gamma(s-n+m+1)}$$
$$\times\ {}_3F_2\left(\begin{array}{c} -m,\ s+1,\ \beta-\alpha-s \\ \beta+1,\quad n-m+1 \end{array}\right), \qquad n \geq m.$$

(The limit $c \to \infty$ of the integral A_p is discussed in [183].) Equations (5.5)–(5.6) and (5.8)–(5.9), which we have not been able to find in the available literature, can be derived in a similar fashion. It does not appear to have been noticed that the corresponding ${}_3F_2$ functions can be expressed in terms of Hahn polynomials:

$$h_n^{(\alpha,\,\beta)}(x, N) = (-1)^n \frac{\Gamma(N)\,(\beta+1)_n}{n!\,\Gamma(N-n)}\ {}_3F_2\left(\begin{array}{c} -n,\ \alpha+\beta+n+1,\ -x \\ \beta+1,\quad 1-N \end{array}\right). \tag{5.11}$$

The ease of handling of these matrix elements for the discrete levels is greatly increased if use is made of the known properties of these polynomials [66, 138, 139].

For example, the difference-differentiation formulas (4.34) and (4.35) of Ref. [183] (see also (B.22) below) take the following convenient form

$$\frac{p(p+1)}{n+2\nu}\ {}_3F_2\left(\begin{array}{c} 1-n,\ -p,\ p+1 \\ 2\nu+1,\quad 2 \end{array}\right) = \frac{p(p+1)}{2\nu+1}\ {}_3F_2\left(\begin{array}{c} 1-n,\ 1-p,\ p+2 \\ 2\nu+2,\quad 2 \end{array}\right)$$
$$= {}_3F_2\left(\begin{array}{c} -n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right) - {}_3F_2\left(\begin{array}{c} 1-n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right) \tag{5.12}$$

in terms of the generalized hypergeometric functions. (Another proof of these identities is given in the Appendix B.) As a result, the linear relation holds [1, 160]

$$2\kappa\left(A_p - \varepsilon B_p\right) - (p+1)\left(B_p - \varepsilon A_p\right) = 4\mu C_p, \tag{5.13}$$

and we can rewrite (5.4) and (5.5) in the following matrix form

$$2(p+1)\, a\mu\, (2a\beta)^p\ \frac{\Gamma(2\nu+1)}{\Gamma(2\nu+p+1)} \begin{pmatrix} A_p \\ B_p \end{pmatrix} \tag{5.14}$$
$$= \begin{pmatrix} \gamma_1 & \gamma_2 \\ \delta_1 & \delta_2 \end{pmatrix} \begin{pmatrix} {}_3F_2\left(\begin{array}{c} 1-n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right) \\ {}_3F_2\left(\begin{array}{c} -n,\ -p,\ p+1 \\ 2\nu+1,\quad 1 \end{array}\right) \end{pmatrix} \qquad (p \neq -1),$$

where

$$\gamma_1 = (\mu + a\kappa)\,(a\,(2\varepsilon\kappa + p + 1) - 2\varepsilon\mu)\,,$$
$$\gamma_2 = (\mu - a\kappa)\,(a\,(2\varepsilon\kappa + p + 1) + 2\varepsilon\mu) \tag{5.15}$$

and

$$\delta_1 = (\mu + a\kappa)\,(a\,(2\kappa + \varepsilon\,(p + 1)) - 2\mu)\,,$$
$$\delta_2 = (\mu - a\kappa)\,(a\,(2\kappa + \varepsilon\,(p + 1)) + 2\mu)\,. \tag{5.16}$$

This representation of integrals A_p and B_p involves the Chebyshev polynomials of a discrete variable $h_p^{(0,\,0)}\,(x, -2\nu)$ at $x = n,\ n - 1$ only; see also Eq. (5.6) for C_p. The corresponding dual Hahn polynomials [139] may be considered as difference analogs of the Laguerre polynomials in Eq. (3.19) for the relativistic radial functions [178].

5.2 Inversion Formulas

Due to the symmetry of the hypergeometric functions in (5.4)–(5.6) under the transformation $p \to -p - 1$, one gets [176]

$$A_{-p-1} = (2a\beta)^{2p+1}\,\frac{\Gamma\,(2\nu - p)}{\Gamma\,(2\nu + p + 1)}\,\frac{\left(\left(1 + \varepsilon^2\right)p + \varepsilon^2\right)A_p - (2p + 1)\,\varepsilon B_p}{\left(1 - \varepsilon^2\right)p}, \tag{5.17}$$

$$B_{-p-1} = (2a\beta)^{2p+1}\,\frac{\Gamma\,(2\nu - p)}{\Gamma\,(2\nu + p + 1)}\,\frac{(2p + 1)\,\varepsilon A_p - \left(\left(1 + \varepsilon^2\right)p + 1\right)B_p}{\left(1 - \varepsilon^2\right)p}, \tag{5.18}$$

$$C_{-p-1} = (2a\beta)^{2p+1}\,\frac{\Gamma\,(2\nu - p)}{\Gamma\,(2\nu + p + 1)}\,C_p. \tag{5.19}$$

(These relations allow us to evaluate all the convergent integrals with $p \le -2$.) Indeed,

$$A_{-p-1} - \varepsilon B_{-p-1} = (2a\beta)^{2p+1}\,\frac{\Gamma\,(2\nu - p)}{\Gamma\,(2\nu + p + 1)}\left(A_p - \varepsilon B_p\right), \tag{5.20}$$

$$B_{-p-1} - \varepsilon A_{-p-1} = -\frac{p + 1}{p}\,(2a\beta)^{2p+1}\,\frac{\Gamma\,(2\nu - p)}{\Gamma\,(2\nu + p + 1)}\left(B_p - \varepsilon A_p\right), \tag{5.21}$$

which gives the first two equations, if $B_p \neq \varepsilon A_p$ and $p \neq 0, -1$. The last one follows from (5.6). Special cases $p = 0, -1$ of (5.20)–(5.21) are simply identity (5.46) and Fock's virial theorem (5.43), respectively. In view of our formulas (5.4)–(5.5), equation $B_p = \varepsilon A_p$ occurs only when $p = 0$ or $n = 0$.

The symmetry of the hypergeometric functions in (5.7)–(5.9) under another reflection $p \to -p-3$ gives

$$A_{-p-3} = (2a\beta)^{2p+3} \ \frac{\Gamma(2\nu - p - 2)}{\Gamma(2\nu + p + 3)} \times \left(\frac{4\mu^2 (2p+3) + (p+2)\left(4\nu^2 + (p+1)(p+2)\right)}{p+2} \ A_p - 2\kappa (2p+3) \ B_p - 8\kappa\mu \frac{2p+3}{p+2} \ C_p \right), \tag{5.22}$$

$$B_{-p-3} = (2a\beta)^{2p+3} \ \frac{\Gamma(2\nu - p - 2)}{\Gamma(2\nu + p + 3)} \times \left(-2\kappa (2p+3) \ A_p + \left(4\nu^2 + (p+1)(p+2)\right) \ B_p + 4\mu (2p+3) \ C_p \right), \tag{5.23}$$

$$C_{-p-3} = (2a\beta)^{2p+3} \ \frac{\Gamma(2\nu - p - 2)}{\Gamma(2\nu + p + 3)} \times \left(2\kappa\mu \frac{2p+3}{p+2} \ A_p - \mu (2p+3) \ B_p - \frac{4\mu^2 (2p+3) + (p+1)\left(4\nu^2 - (p+2)^2\right)}{p+2} \ C_p \right) \tag{5.24}$$

as a result of elementary matrix multiplications. These relations can be used for all the convergent integrals with $p \le -3$. Further details are left to the reader [176].

The corresponding single two-term nonrelativistic relation was found by Pasternack [141, 142] (see also subsection 2.5.1, [160] and references therein). We have been unable to find the relativistic matrix identities (5.17)–(5.19) and (5.22)–(5.24) in the available literature (see Eq. (18) of [6] as the closest analog).

5.3 Recurrence Relations

A set of useful recurrence relations between the relativistic matrix elements was derived by Shabaev [160] (see also [1, 65, 164, 192]) on the basis of a hypervirial theorem:

$$2\kappa A_p - (p+1) B_p = 4\mu C_p + 4\beta\varepsilon C_{p+1}, \tag{5.25}$$

$$2\kappa B_p - (p+1)\,A_p = 4\beta C_{p+1}, \tag{5.26}$$

$$\mu B_p - (p+1)\,C_p = \beta\left(A_{p+1} - \varepsilon B_{p+1}\right). \tag{5.27}$$

(Their computer algebra derivation is presented in [143].) Linear relation (5.13) and convenient recurrence formulas

$$A_{p+1} = -(p+1)\,\frac{4\nu^2\varepsilon + 2\kappa\,(p+2) + \varepsilon\,(p+1)\,(2\kappa\varepsilon + p + 2)}{4\left(1-\varepsilon^2\right)(p+2)\,\beta\mu}\,A_p \tag{5.28}$$
$$+\frac{4\mu^2\,(p+2) + (p+1)\,(2\kappa\varepsilon + p + 1)\,(2\kappa\varepsilon + p + 2)}{4\left(1-\varepsilon^2\right)(p+2)\,\beta\mu}\,B_p,$$

$$B_{p+1} = -(p+1)\,\frac{4\nu^2 + 2\kappa\varepsilon\,(2p+3) + \varepsilon^2\,(p+1)\,(p+2)}{4\left(1-\varepsilon^2\right)(p+2)\,\beta\mu}\,A_p \tag{5.29}$$
$$+\frac{4\mu^2\varepsilon\,(p+2) + (p+1)\,(2\kappa\varepsilon + p + 1)\,(2\kappa + \varepsilon\,(p+2))}{4\left(1-\varepsilon^2\right)(p+2)\,\beta\mu}\,B_p,$$
$$C_{p+1} = \frac{1}{4\mu}\,(2\kappa + \varepsilon\,(p+2))\;A_{p+1} - \frac{1}{4\mu}\,(2\kappa\varepsilon + p + 2)\;B_{p+1} \tag{5.30}$$

are obtained from these equations (see [1, 160, 164] for more details). Their connections with the theory of generalized hypergeometric functions will be discussed elsewhere.

5.4 Special Expectation Values and Their Applications

The Sommerfeld–Dirac formula (3.17) is derived for a point charge atomic nucleus with infinite mass and no internal structure (electron moving in static Coulomb field). In reality, the electron's mass is not negligibly small compared with the nuclear mass and one has to consider the effect of nuclear motion on the energy levels. Actual nuclei have a finite size and possess some internal structure, such as an internal angular momentum or spin, a magnetic dipole moment, and a small electric quadrupole moment associated with the spin, which also affect the energy levels. Radiative corrections are introduced by the quantization of the electromagnetic radiation field. (See [28, 37, 161–163, 165, 166, 204] and references therein for more details.) Calculations of the real energy levels of the high-Z one-electron systems with the help of the perturbation theory require special relativistic matrix elements.

From the explicit expressions (5.4)–(5.9) one can derive the following special matrix elements:

$$A_2 = \left\langle r^2 \right\rangle = \frac{5n\,(n+2\nu) + 4\nu^2 + 1 - \varepsilon\kappa\,(2\varepsilon\kappa + 3)}{2\,(a\beta)^2} \tag{5.31}$$

$$= \frac{2\kappa^2\varepsilon^4 + 3\kappa\varepsilon^3 + \left(3\mu^2 - \nu^2 - 1\right)\varepsilon^2 - 3\kappa\varepsilon - \nu^2 + 1}{2\beta^2\left(1-\varepsilon^2\right)^2},$$

$$A_1 = \langle r \rangle = \frac{3\varepsilon\mu^2 - \kappa\left(1-\varepsilon^2\right)(1+\varepsilon\kappa)}{2\beta\mu\left(1-\varepsilon^2\right)}, \tag{5.32}$$

$$A_0 = \langle 1 \rangle = 1, \tag{5.33}$$

$$A_{-1} = \left\langle \frac{1}{r} \right\rangle = \frac{\beta}{\mu\nu}\left(1-\varepsilon^2\right)\left(\varepsilon\nu + \mu\sqrt{1-\varepsilon^2}\right) \tag{5.34}$$

$$= \frac{m^2c^4 - E^2}{m^2c^4}\left(\frac{E}{Ze^2} + \sqrt{\frac{m^2c^4 - E^2}{\hbar^2c^2\kappa^2 - Z^2e^4}}\right),$$

$$A_{-2} = \left\langle \frac{1}{r^2} \right\rangle = \frac{2a^3\beta^2\kappa\,(2\varepsilon\kappa - 1)}{\mu\nu\left(4\nu^2 - 1\right)}, \tag{5.35}$$

$$A_{-3} = \left\langle \frac{1}{r^3} \right\rangle = 2\,(a\beta)^3\,\frac{3\varepsilon^2\kappa^2 - 3\varepsilon\kappa - \nu^2 + 1}{\nu\left(\nu^2 - 1\right)\left(4\nu^2 - 1\right)}. \tag{5.36}$$

(Note that A_{-3} exists only if $|\kappa| \geq 2$ [160].) The average distance between the electron and the nucleus $\overline{r} = \langle r \rangle$ is given by A_1. The mean square deviation of the nucleus-electron separation is $\overline{(r - \overline{r})^2} = A_2 - (A_1)^2$. The energy eigenvalue $\langle E \rangle$, mean radius $\langle r \rangle$ and mean square radius $\left\langle r^2 \right\rangle$ are frequently used when making comparisons of wave functions computed by different approximation methods. The integrals A_1 and A_2 have been evaluated in [40, 78, 146, 183] (see also Ref. [6] for closed-form expressions for $\left\{A_p\right\}_{p=-6}^{5}$). Matrix element A_{-3} appears in calculation of the electric quadrupole hyperfine splitting [145, 161, 164]. Integrals A_p are also part of the expression for the effective electrostatic potential for the relativistic hydrogenlike atom [183].

$$B_2 = \left\langle \beta r^2 \right\rangle = \frac{\varepsilon}{2\,(a\beta)^2}\left(5n\,(n+2\nu) + 2\nu^2 + 1 - 3\varepsilon\kappa\right) \tag{5.37}$$

$$= \varepsilon\,\frac{3\kappa\varepsilon^3 + \left(5\mu^2 + 3\nu^2 - 1\right)\varepsilon^2 - 3\kappa\varepsilon - 3\nu^2 + 1}{2\beta^2\left(1-\varepsilon^2\right)^2},$$

$$B_1 = \langle \beta r \rangle = \frac{3\varepsilon^2\mu^2 - \left(1-\varepsilon^2\right)\left(\varepsilon\kappa + \nu^2\right)}{2\beta\mu\left(1-\varepsilon^2\right)}, \tag{5.38}$$

$$B_0 = \langle \beta \rangle = \varepsilon = \frac{E}{mc^2}, \tag{5.39}$$

$$B_{-1} = \left\langle \frac{\beta}{r} \right\rangle = \frac{\beta a^2}{\mu} = \frac{m^2 c^4 - E^2}{Z e^2 m c^2}, \tag{5.40}$$

$$B_{-2} = \left\langle \frac{\beta}{r^2} \right\rangle = \frac{2a^3 \beta^2 \left(2\nu^2 - \varepsilon\kappa\right)}{\mu\nu \left(4\nu^2 - 1\right)}, \tag{5.41}$$

$$B_{-3} = \left\langle \frac{\beta}{r^3} \right\rangle = 2\,(a\beta)^3\, \varepsilon \frac{1 + 2\nu^2 - 3\varepsilon\kappa}{\nu \left(\nu^2 - 1\right)\left(4\nu^2 - 1\right)}. \tag{5.42}$$

The integral B_0 appears in the virial theorem for the Dirac equation in a Coulomb field,

$$E = mc^2 \left\langle \beta \right\rangle, \tag{5.43}$$

established by Fock [74] and then developed by many authors (see [37, 39, 60, 65, 76, 81, 125, 130, 149, 150, 152, 155, 160, 164] and references therein). Relation (5.43) can also be obtained with the help of the Hellmann–Feynman theorem,

$$\frac{\partial E}{\partial \lambda} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \tag{5.44}$$

(see [17, 18, 65, 130] and references therein), if applied to the mass parameter [1, 164]. This theorem implies two more relations

$$\frac{\partial E}{\partial Z} = -e^2 \left\langle \frac{1}{r} \right\rangle = -e^2 A_{-1}, \qquad \frac{\partial E}{\partial \kappa} = 2\hbar c C_{-1}. \tag{5.45}$$

The following identities hold

$$A_{-1} - \varepsilon B_{-1} = \frac{a^3 \beta}{\nu} = \frac{1}{\beta}\left(\mu B_{-2} + C_{-2}\right) \tag{5.46}$$

by (5.27). The integral B_{-1} is evaluated in [37] and $A_{-1},\ A_{-2},\ B_{-2},\ C_{-2}$, and A_{-3} are given in [160] (see also [164]).

The relativistic recoil corrections to the energy levels, when nuclear motion is taken into consideration, require matrix elements $A_{-2},\ B_{-1}$ and C_{-2} (see [1, 37, 162, 166] and references therein).

$$\begin{aligned} C_2 &= \frac{\kappa a^2 \left(3n\,(n + 2\nu) + 2\nu^2 + 1\right) - 3\mu^2 \varepsilon}{4\mu\,(a\beta)^2} \\ &= \frac{\kappa \left(1 - \varepsilon^2\right)\left(1 - \nu^2\right) + 3\varepsilon\mu^2\,(\varepsilon\kappa - 1)}{4\mu\beta^2 \left(1 - \varepsilon^2\right)}, \end{aligned} \tag{5.47}$$

$$C_1 = \frac{2\varepsilon\kappa - 1}{4\beta} = \frac{\hbar}{4m^2c^3}\left(2\kappa E - mc^2\right), \tag{5.48}$$

$$C_0 = \frac{\kappa}{2\mu}\left(1-\varepsilon^2\right) = \frac{\hbar c\kappa}{2Ze^2}\frac{m^2c^4 - E^2}{m^2c^4}, \tag{5.49}$$

$$C_{-1} = \frac{\kappa}{2\mu\nu}a^3\beta = \frac{a\beta}{\nu}C_0 \tag{5.50}$$

$$= \frac{\hbar\kappa}{2Ze^2m^2c^3}\frac{\left(m^2c^4 - E^2\right)^{3/2}}{\left(\hbar^2c^2\kappa^2 - Z^2e^4\right)^{1/2}},$$

$$C_{-2} = \frac{a^3\beta^2\left(2\varepsilon\kappa - 1\right)}{\nu\left(4\nu^2-1\right)} = \frac{4\left(a\beta\right)^3 C_1}{\nu\left(4\nu^2-1\right)}, \tag{5.51}$$

$$C_{-3} = \left(a\beta\right)^3 \frac{\kappa\left(1-\varepsilon^2\right)\left(1-\nu^2\right) + 3\varepsilon\mu^2\left(\varepsilon\kappa - 1\right)}{\mu\nu\left(\nu^2-1\right)\left(4\nu^2-1\right)} = \frac{4\left(a\beta\right)^5 C_2}{\nu\left(\nu^2-1\right)\left(4\nu^2-1\right)}. \tag{5.52}$$

The integrals C_0, C_1, and B_{-1} are computed in [81]. In view of (5.40) and (5.49), respectively (5.35) and (5.51), the following simple relations hold

$$C_0 = \frac{\kappa}{2\beta}B_{-1}, \qquad A_{-2} = \frac{2\kappa}{\mu}C_{-2} = \frac{8\left(a\beta\right)^3\kappa}{\mu\nu\left(4\nu^2-1\right)}C_1. \tag{5.53}$$

The last but one was originally found in [39].

The integral C_1 occurs in calculations of the bound-electron g factor (the anomalous Zeeman effect in the presence of an external homogeneous static magnetic field) [126, 149, 164, 167, 201]. The matrix element C_{-1} has also been found by the Hellmann–Feynman theorem (5.45). The integral C_{-2} appears in calculation of the magnetic dipole hyperfine splitting [36, 39, 76, 129, 145, 149, 161].

We hope that the rest of matrix elements will also be useful in the current theory of hydrogenlike heavy ions and other exotic relativistic Coulomb systems. Professor Shabaev kindly pointed out that the formulas derived in this section can be used in calculations with hydrogenlike wave functions where a high precision is required (see also [105] for the most accurate up-to-date electron mass measurements, where the relativistic mean electric fields are estimated with the help of one of our integrals).

In Table 1, we list the expectation values for the $1s_{1/2}$ state, when $n = n_r = 0$, $l = 0$, $j = 1/2$, and $\kappa = -1$. The corresponding radial wave functions are given by Eq. (3.20).

Table 1 Expectation values for the $1s_{1/2}$ state

p	A_p	B_p	C_p
2	$\frac{1}{2}\left(\frac{a_0}{Z}\right)^2(\nu_1+1)(2\nu_1+1)$	$\frac{1}{2}\left(\frac{a_0}{Z}\right)^2\nu_1(\nu_1+1)(2\nu_1+1)$	$-\frac{\lambda a_0}{4Z}(\nu_1+1)(2\nu_1+1)$
1	$\frac{a_0}{2Z}(2\nu_1+1)$	$\frac{a_0}{2Z}\nu_1(2\nu_1+1)$	$-\frac{\lambda}{4}(2\nu_1+1)$
0	1	ν_1	$-\frac{\lambda Z}{2a_0}$
−1	$\frac{Z}{a_0\nu_1}$	$\frac{Z}{a_0}$	$-\left(\frac{Z}{a_0}\right)^2\frac{\lambda}{2\nu_1}$
−2	$\left(\frac{Z}{a_0}\right)^2\frac{2}{\nu_1(2\nu_1-1)}$	$\left(\frac{Z}{a_0}\right)^2\frac{2}{(2\nu_1-1)}$	$-\left(\frac{Z}{a_0}\right)^3\frac{\lambda}{\nu_1(2\nu_1-1)}$
−3	$\left(\frac{Z}{a_0}\right)^3\frac{2}{\nu_1(\nu_1-1)(2\nu_1-1)}$	$\left(\frac{Z}{a_0}\right)^3\frac{2}{(\nu_1-1)(2\nu_1-1)}$	$-\left(\frac{Z}{a_0}\right)^4\frac{\lambda}{\nu_1(\nu_1-1)(2\nu_1-1)}$

In the table, $\varepsilon_1=\nu_1=\sqrt{1-\mu^2}=\sqrt{1-(\alpha Z)^2}$, $\alpha=e^2/\hbar c$ is the Sommerfeld fine structure constant, $a_0=\hbar^2/me^2$ is the Bohr radius, and $\lambda=\hbar/mc$ is the Compton wavelength. The relations

$$B_p=\varepsilon_1\,A_p,\qquad C_p=-\frac{\lambda Z}{2a_0}\,A_p,\qquad A_p=\left(\frac{a_0}{2Z}\right)^p\frac{\Gamma(2\nu_1+p+1)}{\Gamma(2\nu_1+1)} \tag{5.54}$$

(for all the suitable integers $p>-2\nu_1-1>-3$) follow directly from (5.4), (5.5) and (5.7), (5.9). (The formal expressions for $A_{-3},\ B_{-3}$, and C_{-3}, when the integrals diverge, are included into the table for "completeness"; see Ref. [1] for more details.) The reflection relation (5.19) holds for all the convergent integrals $A_p,\ B_p$, and C_p.

5.5 *Three-Term Recurrence Relations and Computer Algebra Methods*

The following three-term recurrence relations for the relativistic matrix elements have been found in [178]:

$$\begin{aligned}A_{p+1}=&\ \frac{\mu P(p)}{a^2\beta\left(4\mu^2(p+1)+p(2\varepsilon\kappa+p)(2\varepsilon\kappa+p+1)\right)(p+2)}\,A_p \\ &-\frac{\left(4\nu^2-p^2\right)\left(4\mu^2(p+2)+(p+1)(2\varepsilon\kappa+p+1)(2\varepsilon\kappa+p+2)\right)p}{(2a\beta)^2\left(4\mu^2(p+1)+p(2\varepsilon\kappa+p)(2\varepsilon\kappa+p+1)\right)(p+2)}\,A_{p-1},\end{aligned} \tag{5.55}$$

$$B_{p+1}=\frac{\varepsilon\mu Q\left(p\right)}{a^2\beta\left(4\nu^2+2\varepsilon\kappa\left(2p+1\right)+\varepsilon^2p\left(p+1\right)\right)\left(p+2\right)}\,B_p \tag{5.56}$$
$$-\frac{\left(4\nu^2-p^2\right)\left(4\nu^2+2\varepsilon\kappa\left(2p+3\right)+\varepsilon^2\left(p+1\right)\left(p+2\right)\right)\left(p+1\right)}{\left(2a\beta\right)^2\left(4\nu^2+2\varepsilon\kappa\left(2p+1\right)+\varepsilon^2p\left(p+1\right)\right)\left(p+2\right)}\,B_{p-1},$$

where

$$\begin{aligned}P\left(p\right)=&\ 2\varepsilon p\left(p+2\right)\left(2\varepsilon\kappa+p\right)\left(2\varepsilon\kappa+p+1\right)\\&+\varepsilon\left(4\left(\varepsilon^2\kappa^2-\nu^2\right)-p\left(4\varepsilon^2\kappa^2+p\left(p+1\right)\right)\right)\\&+\left(2p+1\right)\left(4\varepsilon^2\kappa+2\left(p+2\right)\left(2\varepsilon\mu^2-\kappa\right)\right),\end{aligned} \tag{5.57}$$

$$\begin{aligned}Q\left(p\right)=&\left(2p+3\right)\left(4\nu^2+2\varepsilon\kappa\left(2p+1\right)+p\left(p+1\right)\right)\\&-a^2\left(2p+1\right)\left(p+1\right)\left(p+2\right).\end{aligned} \tag{5.58}$$

In comparison with other papers (see [1, 6, 159, 160, 176, 179] and references therein), our consideration provides an alternative way of the recursive evaluation of the special values A_p and B_p, when we deal separately with one of these integrals only. The corresponding initial data $A_0=1,\ B_{-1}=a^2\beta/\mu$ can be found in [176].

These three-term recurrence relations are investigated by advanced computer algebra methods in [107, 143] among other things. For instance, their direct derivations from the integrals of Laguerre polynomials (5.1) and (5.2) are given [107].

Acknowledgements We are very grateful to the organizers of the AIMS-Volkswagen Stiftung Workshop on Introduction to Orthogonal Polynomials and Applications held in Douala, Cameroon, in October 2018, without their help this publication would be impossible. Sincere thanks to our co-authors Christoph Koutschan, Sergey Kryuchkov, Raquel M. López, Peter Paule and Erwin Suazo; some of our joint results were included in this chapter. We are very grateful to the referee for an outstanding job, her/his numerous comments helped us to improve the manuscript. Special thanks are directed to Jeremy Alm, Valentin Andreev, Al Boggess and Michael Laidacker for support and continuous encouragement. This work was partially supported by the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health (NIH; Grant # K12GM102778) and by the Lamar University Research Enhancement Grant (REG # 420266).

Appendix A: Evaluation of an Integral

Let us compute the following integral

$$J_{nms}^{\alpha\beta}=\int_0^\infty e^{-x}x^{\alpha+s}L_n^{\alpha}\left(x\right)L_m^{\beta}\left(x\right)\,dx, \tag{A.1}$$

where $n \geq m$ and $\alpha - \beta = 0, \pm 1, \pm 2, \dots$. Similar integrals were evaluated in [28, 49] and [113], see also references therein, but an important relation with the Hahn polynomials seems to be missing.

It is convenient to assume at the beginning that parameter s takes some continuous values such that $\alpha + s > -1$ for convergence of the integral. Using the Rodrigues formula for the Laguerre polynomials [138, 139, 184]

$$L_n^{\alpha}(x) = \frac{1}{n!} e^x x^{-\alpha} \left(x^{\alpha+n} e^{-x}\right)^{(n)}, \tag{A.2}$$

see the proof in Sect. 1 of the present chapter, and integrating by parts

$$\begin{aligned} J_{nms}^{\alpha\beta} &= \frac{1}{n!} \int_0^{\infty} \left(x^{\alpha+n} e^{-x}\right)^{(n)} \left(x^s L_m^{\beta}(x)\right) dx \\ &= \frac{1}{n!} \left(\left(x^{\alpha+n} e^{-x}\right)^{(n-1)} \left(x^s L_m^{\beta}(x)\right) \right) \Big|_0^{\infty} \\ &\quad - \frac{1}{n!} \int_0^{\infty} \left(x^{\alpha+n} e^{-x}\right)^{(n-1)} \left(x^s L_m^{\beta}(x)\right)' dx \\ &\quad \vdots \\ &= \frac{(-1)^n}{n!} \int_0^{\infty} \left(x^{\alpha+n} e^{-x}\right) \left(x^s L_m^{\beta}(x)\right)^{(n)} dx. \end{aligned}$$

However, in view of (B.1),

$$\begin{aligned} \left(x^s L_m^{\beta}(x)\right)^{(n)} &= \frac{\Gamma(\beta+m+1)}{m!\,\Gamma(\beta+1)} \sum_k \frac{(-m)_k}{k!\,(\beta+1)_k} \left(x^{k+s}\right)^{(n)} \\ &= \frac{\Gamma(\beta+m+1)\,\Gamma(s+1)}{m!\,\Gamma(\beta+1)\,\Gamma(s-n+1)} \sum_k \frac{(-m)_k\,(s+1)_k}{k!\,(\beta+1)_k\,(s-n+1)_k}\, x^{k+s-n} \end{aligned} \tag{A.3}$$

and with the help of Euler's integral representation for the gamma function [8, 138]

$$\int_0^{\infty} x^{\alpha+k+s} e^{-x}\, dx = \Gamma(\alpha+k+s+1) = (\alpha+s+1)_k\, \Gamma(\alpha+s+1),$$

see also (B.10) below, one gets

$$\begin{aligned} J_{nms}^{\alpha\beta} &= (-1)^n \frac{\Gamma(\alpha+s+1)\,\Gamma(\beta+m+1)\,\Gamma(s+1)}{n!\,m!\,\Gamma(\beta+1)\,\Gamma(s-n+1)} \\ &\quad \times {}_3F_2\left(\begin{matrix} -m, & s+1, & \alpha+s+1 \\ \beta+1, & & s-n+1 \end{matrix}\right). \end{aligned} \tag{A.4}$$

See [16] or Eq. (1.12) for the definition of the generalized hypergeometric series ${}_3F_2\,(1)$. Thomae's transformation (B.8), see also [16] or [79], results in [183]

$$J_{nms}^{\alpha\beta} = \int_0^\infty e^{-x} x^{\alpha+s}\, L_n^{\alpha}(x)\, L_m^{\beta}(x)\; dx \tag{A.5}$$
$$= (-1)^{n-m} \frac{\Gamma(\alpha+s+1)\,\Gamma(\beta+m+1)\,\Gamma(s+1)}{m!\,(n-m)!\,\Gamma(\beta+1)\,\Gamma(s-n+m+1)}$$
$$\times\ {}_3F_2\left(\begin{matrix}-m,\ s+1,\ \beta-\alpha-s\\ \beta+1,\quad n-m+1\end{matrix}\right), \quad n\geq m,$$

where parameter s may take some integer values. This establishes a connection with the Hahn polynomials given by Eq. (B.6) below; one can also rewrite this integral in terms of the dual Hahn polynomials [139].

Letting $s=0$ and $\alpha=\beta$ in (A.5) results in the orthogonality relation for the Laguerre polynomials. Two special cases

$$I_1 = J_{nn1}^{\alpha\alpha} = \int_0^\infty e^{-x} x^{\alpha+1} \left(L_n^{\alpha}(x)\right)^2\; dx = (\alpha+2n+1)\,\frac{\Gamma(\alpha+n+1)}{n!} \tag{A.6}$$

and

$$I_2 = J_{n,\ n-1,\ 2}^{\alpha-2,\ \alpha} = \int_0^\infty e^{-x} x^{\alpha} L_{n-1}^{\alpha}(x)\, L_n^{\alpha-2}(x)\; dx = -2\frac{\Gamma(\alpha+n)}{(n-1)!} \tag{A.7}$$

are convenient for normalization of the wave functions of the discrete spectra in the nonrelativistic and relativistic Coulomb problems [28, 138].

Two other special cases of a particular interest in this chapter are

$$J_k = J_{nnk}^{\alpha\alpha} = \int_0^\infty e^{-x} x^{\alpha+k}\ \left(L_n^{\alpha}(x)\right)^2\; dx \tag{A.8}$$
$$= \frac{\Gamma(\alpha+k+1)\,\Gamma(\alpha+n+1)}{n!\,\Gamma(\alpha+1)}\ {}_3F_2\left(\begin{matrix}-k,\ k+1,\ -n\\ 1,\quad \alpha+1\end{matrix}\right)$$

and

$$J_{-k-1} = J_{nn,\ -k-1}^{\alpha\alpha} = \int_0^\infty e^{-x} x^{\alpha-k-1}\ \left(L_n^{\alpha}(x)\right)^2\; dx \tag{A.9}$$
$$= \frac{\Gamma(\alpha-k)\,\Gamma(\alpha+n+1)}{n!\,\Gamma(\alpha+1)}\ {}_3F_2\left(\begin{matrix}-k,\ k+1,\ -n\\ 1,\quad \alpha+1\end{matrix}\right).$$

The Chebyshev polynomials of a discrete variable $t_k(x)$ are special case of the Hahn polynomials $t_k(x, N) = h_k^{(0,\,0)}(x, N)$ [185, 186] and [187]. Thus from (A.8)–(A.9) and (B.6) one finally gets

$$J_k = J_{nnk}^{\alpha\alpha} = \int_0^\infty e^{-x} x^{\alpha+k} \left(L_n^\alpha(x)\right)^2 \, dx \tag{A.10}$$
$$= \frac{\Gamma(\alpha+n+1)}{n!} t_k(n, -\alpha)$$

and

$$J_{-k-1} = J_{nn,\,-k-1}^{\alpha\alpha} = \int_0^\infty e^{-x} x^{\alpha-k-1} \left(L_n^\alpha(x)\right)^2 \, dx \tag{A.11}$$
$$= \frac{\Gamma(\alpha-k)\,\Gamma(\alpha+n+1)}{n!\,\Gamma(\alpha+k+1)} t_k(n, -\alpha)$$

for $0 \le k < \alpha$. One can see that the positivity of these integrals is related to a nonstandard orthogonality relation for the corresponding Chebyshev polynomials of a discrete variable $t_k(x, N)$ when the parameter takes negative integer values $N = -\alpha$. Indeed, according to the method of [139] and [138], these polynomials are orthogonal with the discrete uniform distribution on the interval $[-\alpha, -1]$ which contains all their zeros and, therefore, they are positive for all nonnegative values of their argument. The explicit representation (B.6) gives also a positive sum for all positive x and negative N.

Appendix B: Hypergeometric Series, Discrete Orthogonal Polynomials, and Useful Relations

This section contains some relations involving the generalized hypergeometric series, the Laguerre and Hahn polynomials, the spherical harmonics and Clebsch–Gordan coefficients, which are used throughout the paper.

The Laguerre polynomials are defined as [8, 138, 139, 184]

$$L_n^\alpha(x) = \frac{\Gamma(\alpha+n+1)}{n!\,\Gamma(\alpha+1)} \, {}_1F_1\left(\begin{matrix} -n \\ \alpha+1 \end{matrix}; \, x\right). \tag{B.1}$$

(It is a consequence of Theorem 1.1.) The differentiation formulas [138, 139]

$$\frac{d}{dx} L_n^\alpha(x) = -L_{n-1}^{\alpha+1}(x), \tag{B.2}$$

$$x\frac{d}{dx}L_n^{\alpha}(x) = nL_n^{\alpha}(x) - (\alpha+n)\,L_{n-1}^{\alpha}(x) \tag{B.3}$$

imply a recurrence relation

$$xL_{n-1}^{\alpha+1}(x) = (\alpha+n)\,L_{n-1}^{\alpha}(x) - nL_n^{\alpha}(x)\,. \tag{B.4}$$

The simplest case of the connecting relation (c.f. [8] and [9]) is

$$L_n^{\alpha}(x) = L_n^{\alpha+1}(x) - L_{n-1}^{\alpha+1}(x)\,. \tag{B.5}$$

The Hahn polynomials are [138, 139]

$$h_n^{(\alpha,\,\beta)}(x, N) = (-1)^n \frac{\Gamma(N)\,(\beta+1)_n}{n!\,\Gamma(N-n)}\ {}_3F_2\left(\begin{matrix}-n,\ \alpha+\beta+n+1,\ -x\\ \beta+1,\quad 1-N\end{matrix};\ 1\right). \tag{B.6}$$

(We usually omit the argument of the hypergeometric series ${}_3F_2$ if it is equal to one.) An asymptotic relation with the Jacobi polynomials is

$$\frac{1}{\widetilde{N}^{\,n}}\,h_n^{(\alpha,\,\beta)}\left(\frac{\widetilde{N}}{2}(1+s) - \frac{\beta+1}{2},\ N\right) = P_n^{(\alpha,\,\beta)}(s) + \mathrm{O}\left(\frac{1}{\widetilde{N}^{\,2}}\right), \tag{B.7}$$

where $\widetilde{N} = N + (\alpha+\beta)/2$ and $N \to \infty$; see [139] for more details.

Thomae's transformation [16, 79] is

$${}_3F_2\left(\begin{matrix}-n,\ a,\ b\\ c,\ d\end{matrix};\ 1\right) = \frac{(d-b)_n}{(d)_n}\ {}_3F_2\left(\begin{matrix}-n,\ c-a,\ b\\ c,\ b-d-n+1\end{matrix};\ 1\right) \tag{B.8}$$

with $n = 0, 1, 2, \ \ldots\ .$

The summation formula of Gauss [8, 16, 79]

$${}_2F_1\left(\begin{matrix}a,\ b\\ c\end{matrix};\ 1\right) = \frac{\Gamma(c)\,\Gamma(c-a-b)}{\Gamma(c-a)\,\Gamma(c-b)}, \qquad \mathrm{Re}\,(c-a-b) > 0. \tag{B.9}$$

The gamma function is defined as [8, 66, 138]

$$\Gamma(z) = \int_0^{\infty} e^{-t}t^{z-1}\,dt, \qquad \mathrm{Re}\ z > 0. \tag{B.10}$$

It can be continued analytically over the whole complex plane except the points $z = 0, -1, -2, \ldots$ at which it has simple poles. Functional equations are

$$\Gamma(z+1) = z\Gamma(z)\,, \tag{B.11}$$

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin \pi z}, \tag{B.12}$$

$$2^{2z-1}\Gamma(z)\,\Gamma(z+1/2) = \sqrt{\pi}\,\Gamma(2z)\,. \tag{B.13}$$

The generating function for the Legendre polynomials and the addition theorem for spherical harmonics give rise to the following expansion formula [138, 189]

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{4\pi}{2l+1} \frac{r_<^l}{r_>^{l+1}}\, Y_{lm}(\theta_1, \varphi_1)\, Y_{lm}^*(\theta_2, \varphi_2)\,, \tag{B.14}$$

where $r_< = \min(r_1, r_2)$ and $r_> = \max(r_1, r_2)\,.$

The Clebsch–Gordan series for the spherical harmonics is [139, 148, 189]

$$\begin{aligned} Y_{l_1 m_1}(\theta, \varphi)\; Y_{l_2 m_2}(\theta, \varphi) = &\sum_{l=|l_1-l_2|}^{l_1+l_2} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi\,(2l+1)}} \\ &\times C_{l_1 m_1 l_2 m_2}^{l,\; m_1+m_2}\, C_{l_1 0 l_2 0}^{l,\;0}\, Y_{l, m_1+m_2}(\theta, \varphi)\,, \end{aligned} \tag{B.15}$$

where $C_{l_1 m_1 l_2 m_2}^{lm}$ are the Clebsch–Gordan coefficients. The special case $l_2 = 1$ reads [71],

$$-\sin\theta e^{i\varphi}\; Y_{l,\;m-1} = \sqrt{\tfrac{(l+m)(l+m+1)}{(2l+1)(2l+3)}}\, Y_{l+1,\;m} - \sqrt{\tfrac{(l-m)(l-m-1)}{(2l+1)(2l-1)}}\, Y_{l-1,\;m}, \tag{B.16}$$

$$\sin\theta e^{-i\varphi}\; Y_{l,\;m+1} = \sqrt{\tfrac{(l-m)(l-m+1)}{(2l+1)(2l+3)}}\, Y_{l+1,\;m} - \sqrt{\tfrac{(l+m)(l+m+1)}{(2l+1)(2l-1)}}\, Y_{l-1,\;m}, \tag{B.17}$$

$$\cos\theta\; Y_{lm} = \sqrt{\tfrac{(l+1)^2-m^2}{(2l+1)(2l+3)}}\, Y_{l+1,\;m} + \sqrt{\tfrac{l^2-m^2}{(2l-1)(2l+1)}}\, Y_{l-1,\;m}, \tag{B.18}$$

where

$$\sqrt{\frac{8\pi}{3}}\, Y_{1,\;\pm 1} = \mp \sin\theta e^{\pm i\varphi}, \qquad \sqrt{\frac{4\pi}{3}}\, Y_{10} = \cos\theta. \tag{B.19}$$

These relations allow to prove (3.34) by a direct calculation.

The required identity (5.12) can be derived from the theory of classical polynomials in the following fashion. Let us start from the difference equation for the Hahn polynomials $y_m = h_m^{(\alpha,\;\beta)}(x, N)$ [139]:

$$(\sigma(x)\,\nabla + \tau(x))\,\Delta y_m + \lambda_m y_m = 0, \tag{B.20}$$

where $\Delta f(x) = \nabla f(x+1) = f(x+1) - f(x)$ and

$$\begin{aligned} \sigma(x) &= x(\alpha + N - x), \\ \tau(x) &= (\beta+1)(N-1) - (\alpha+\beta+2)x, \\ \lambda_m &= m(\alpha+\beta+m+1), \end{aligned} \tag{B.21}$$

and use the familiar difference-differentiation formula:

$$\Delta h_m^{(\alpha,\,\beta)}(x, N) = (\alpha+\beta+m+1)\, h_{m-1}^{(\alpha+1,\,\beta+1)}(x, N-1). \tag{B.22}$$

As a result,

$$(\sigma(x)\nabla + \tau(x))\, h_{m-1}^{(\alpha+1,\,\beta+1)}(x, N-1) + m h_m^{(\alpha,\,\beta)}(x, N) = 0. \tag{B.23}$$

Letting $\alpha = \beta$ and $\beta \to -1$, one gets

$$\begin{aligned} x(N-x-1)\nabla h_{m-1}^{(0,\,0)}(x, N-1) &= -m \lim_{\beta\to -1} h_m^{(\beta,\,\beta)}(x, N) \\ &= (-1)^m m(m-1)\frac{\Gamma(N-1)}{\Gamma(N-m)} x\, {}_3F_2\left(\begin{matrix} 1-m,\ m,\ 1-x \\ 2,\quad 2-N \end{matrix}\right) \end{aligned} \tag{B.24}$$

by (5.11). The last identity takes the form (5.12), if the Chebyshev polynomials of a discrete variable $h_{m-1}^{(0,\,0)}(x, N-1)$ are replaced by the corresponding generalized hypergeometric functions. (Use of (B.22) in (B.24) gives the special ${}_3F_2$ transformation.)

Appendix C: Dirac Matrices and Inner Product

We use the standard representations of the Dirac and Pauli matrices (3.3) and (3.4). The inner product of two Dirac (bispinor) wave functions

$$\psi = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{v}_1 \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \qquad \phi = \begin{pmatrix} \boldsymbol{u}_2 \\ \boldsymbol{v}_2 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} \tag{C.1}$$

is defined as a scalar quantity

$$\langle \psi,\ \phi \rangle = \int_{\mathbb{R}^3} \psi^\dagger \phi \, dv = \int_{\mathbb{R}^3} \left(\boldsymbol{u}_1^\dagger \boldsymbol{u}_2 + \boldsymbol{v}_1^\dagger \boldsymbol{v}_2 \right) \, dv$$
$$= \int_{\mathbb{R}^3} \left(\psi_1^* \phi_1 + \psi_2^* \phi_2 + \psi_3^* \phi_3 + \psi_4^* \phi_4 \right) \, dv \tag{C.2}$$

and the raised asterisk is used to denote the complex conjugate. The corresponding expectation values of a matrix operator A are given by

$$\langle A \rangle = \langle \psi,\ A\psi \rangle \,. \tag{C.3}$$

From this definition one gets

$$\langle r^p \rangle = A_p, \qquad \langle \beta r^p \rangle = B_p, \qquad \langle i \boldsymbol{\alpha} \mathbf{n} \beta r^p \rangle = -2C_p, \tag{C.4}$$

where the integrals $A_p,\ B_p$, and C_p are given by (5.1)–(5.3), respectively.

Indeed, the first relation is derived, for example, in [183] and the second one can be obtained by integrating the identity

$$r^p \psi^\dagger \beta \psi = r^p \left(\boldsymbol{\varphi}^\dagger,\ \boldsymbol{\chi}^\dagger \right) \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix} = r^p \left(\boldsymbol{\varphi}^\dagger,\ \boldsymbol{\chi}^\dagger \right) \begin{pmatrix} \boldsymbol{\varphi} \\ -\boldsymbol{\chi} \end{pmatrix}$$
$$= r^p \left(\boldsymbol{\varphi}^\dagger \boldsymbol{\varphi} - \boldsymbol{\chi}^\dagger \boldsymbol{\chi} \right) = r^p \left(\mathcal{Y}^\dagger \mathcal{Y} \right) \left(F^2 - G^2 \right) \tag{C.5}$$

(we leave details to the reader) in a similar fashion.

In the last case, we start from the matrix identity

$$(\boldsymbol{\alpha} \mathbf{n}) \, \beta \psi = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \mathbf{n} \\ \boldsymbol{\sigma} \mathbf{n} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi} \\ -\boldsymbol{\chi} \end{pmatrix} = \begin{pmatrix} -(\boldsymbol{\sigma} \mathbf{n}) \ \boldsymbol{\chi} \\ (\boldsymbol{\sigma} \mathbf{n}) \ \boldsymbol{\varphi} \end{pmatrix} \tag{C.6}$$

and use the Ansatz [183]

$$\boldsymbol{\varphi} = \boldsymbol{\varphi} (\mathbf{r}) = \mathcal{Y} (\mathbf{n}) \ F(r) \,, \qquad \boldsymbol{\chi} = \boldsymbol{\chi} (\mathbf{r}) = -i \left((\boldsymbol{\sigma} \mathbf{n}) \, \mathcal{Y} (\mathbf{n}) \right) \ G(r) \,, \tag{C.7}$$

where $\mathbf{n} = \mathbf{r}/r$ and $\mathcal{Y} = \mathcal{Y}_{jm}^{\pm}(\mathbf{n})$ are the spinor spherical harmonics given by (3.12). As a result,

$$\begin{aligned} ir^p\psi^\dagger\left((\boldsymbol{\alpha}\mathbf{n})\,\beta\psi\right) &= ir^p\left(\boldsymbol{\varphi}^\dagger,\ \boldsymbol{\chi}^\dagger\right)\begin{pmatrix} -(\boldsymbol{\sigma}\mathbf{n})\ \boldsymbol{\chi} \\ (\boldsymbol{\sigma}\mathbf{n})\ \boldsymbol{\varphi} \end{pmatrix} \\ &= ir^p\left(F\mathcal{Y}^\dagger,\ iG\mathcal{Y}^\dagger(\boldsymbol{\sigma}\mathbf{n})\right)\begin{pmatrix} i\mathcal{Y}G \\ (\boldsymbol{\sigma}\mathbf{n})\,\mathcal{Y}F \end{pmatrix} \\ &= -r^p\left(\mathcal{Y}^\dagger\mathcal{Y}\right)FG - r^p\left(\mathcal{Y}^\dagger(\boldsymbol{\sigma}\mathbf{n})^2\mathcal{Y}\right)FG \\ &= -2r^p\left(\mathcal{Y}^\dagger\mathcal{Y}\right)FG \end{aligned} \tag{C.8}$$

with the help of the familiar identity $(\boldsymbol{\sigma}\mathbf{n})^2 = \mathbf{n}^2 = \mathbf{1}$. Integration over $\mathbb{R}^3$ in the spherical coordinates completes the proof.

References

1. G.S. Adkins, Dirac-Coulomb energy levels and expectation values. Am. J. Phys. **76**(6), 579–584 (2008). https://doi.org/10.1119/1.2830535
2. A.I. Akhiezer, V.B. Berestetskii, *Quantum Electrodynamics*. Interscience Monographs and Texts in Physics and Astronomy, vol. 11 (Interscience Publishers, New York, 1965). Translated from the Russian edition (Moscow, ed. 2, 1959) by A. Sen and R. N. Sen. Israel Program for Scientific Translations, Jerusalem
3. V. Aldaya, F. Cossío, J. Guerrero, F.F. López-Ruiz, The quantum Arnold transformation. J. Phys. A Math. Theor. **44**(6), 065302 (2011). https://doi.org/10.1088/1751-8113/44/6/065302
4. R.L. Anderson, S. Kumei, C.E. Wulfman, Invariants of the equations of wave mechanics I. Rev. Mex. Fís. **21**(1), 1–33 (1972)
5. R.L. Anderson, S. Kumei, C.E. Wulfman, Invariants of the equations of wave mechanics II. One-particle Schrödinger equations. Rev. Mex. Fís. **21**(1), 35–57 (1972)
6. D. Andrae, Recursive evaluation of expectation values $\langle r^k \rangle$ for arbitrary states of the relativistic one–electron atom. J. Phys. B Atom. Mol. Phys. **30**(20), 4435–4451 (1997). https://doi.org/10.1088/0953-4075/30/20/008
7. G.E. Andrews, R. Askey, Classical orthogonal polynomials, in *Polynômes Orthogonaux et Applications*, ed. by C. Brezinski, A. Draux, A.P. Magnus, P.M. Ronveaux. Lecture Notes in Mathematics, vol. 1171 (Springer, Berlin, 1985), pp. 36–62. https://doi.org/10.1007/BFb0076530
8. G.E. Andrews, R. Askey, R. Roy, *Special Functions*. Encyclopedia of Mathematics and Its Applications, vol. 71 (Cambridge University Press, Cambridge, 1999). https://doi.org/10.1017/CBO9781107325937
9. R. Askey, *Orthogonal Polynomials and Special Functions*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. CB21 (Society for Industrial and Applied Mathematics, Philadelphia, 1975). https://doi.org/10.1137/1.9781611970470
10. R. Askey, Continuous Hahn polynomials. J. Phys. A Math. Gen. **18**(16), L1017–L1019 (1985). https://doi.org/10.1088/0305-4470/18/16/004
11. R. Askey, J. Wilson, A set of hypergeometric orthogonal polynomials. SIAM J. Math. Anal. **13**(4), 651–655 (1982). https://doi.org/10.1137/0513043

12. R. Askey, J. Wilson, Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials. Mem. Am. Math. Soc. **54**(319), iv+55 (1985). https://doi.org/10.1090/memo/0319
13. N.M. Atakishiyev, S.K. Suslov, The Hahn and Meixner polynomials of an imaginary argument and some of their applications. J. Phys. A Math. Gen. **18**(10), 1583–1596 (1985). https://doi.org/10.1088/0305-4470/18/10/014
14. N.M. Atakishiyev, S.K. Suslov, On the moments of classical and related polynomials. Rev. Mex. Fís. **34**(2), 147–151 (1988)
15. N.M. Atakishiyev, S.K. Suslov, Difference hypergeometric functions, in *Progress in Approximation Theory*, ed. by A.A. Gonchar, E.B. Saff. Computational Mathematics Book Series, vol. 19 (Springer, New York, 1992), pp. 1–35. https://doi.org/10.1007/978-1-4612-2966-7_1
16. W.N. Bailey, *Generalized Hypergeometric Series*. Cambridge Tracts in Mathematics and Mathematical Physics, vol. 32 (Cambridge University Press, London, 1935)
17. S. Balasubramanian, Note on Feynman's theorem. Am. J. Phys. **52**(12), 1143–1144 (1984). https://doi.org/10.1119/1.13746
18. S. Balasubramanian, A note on the generalized Hellmann-Feynman theorem. Am. J. Phys. **58**(12), 1204–1205 (1990). https://doi.org/10.1119/1.16254
19. V. Bargmann, Zur Theorie des Wasserstoffatoms. Z. Phys. **98**(7), 576–582 (1936). https://doi.org/10.1007/BF01338811
20. H. Bateman, Some properties of a certain set of polynomials. Tôhoku Math. J. First Ser. **37**, 23–38 (1933)
21. H. Bateman, Functions orthogonal in the Hermitian sense. A new application of basic numbers. Proc. Natl. Acad. Sci. **20**(1), 63–66 (1934). https://doi.org/10.1073/pnas.20.1.63
22. H. Bateman, The polynomial $F_n(x)$. Ann. Math. **35**, 767–775 (1934). https://doi.org/10.2307/1968493
23. H. Bateman, The polynomial $F_n(x)$ and its relation to other functions. Ann. Math. **38**(2), 303–310 (1937). https://doi.org/10.2307/1968555
24. H. Bateman, An orthogonal property of the hypergeometric polynomial. Proc. Natl. Acad. Sci. **28**(9), 374–377 (1942). https://doi.org/10.1073/pnas.28.9.374
25. V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, Relativistic quantum theory: part 1, in *Course of Theoretical Physics*, vol. 4, 1st edn. (Pergamon Press, Oxford, 1971). Translated from the Russian by J. B. Sykes and J. S. Bell
26. M.V. Berry, Quantal phase factors accompanying adiabatic changes. Proc. R. Soc. Lond. A Math. Phys. Sci. **392**(1802), 45–57 (1984). https://doi.org/10.1098/rspa.1984.0023
27. M.V. Berry, Classical adiabatic angles and quantal adiabatic phase. J. Phys. A Math. Gen. **18**(1), 15–27 (1985). https://doi.org/10.1088/0305-4470/18/1/012
28. H.A. Bethe, E.E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, Berlin, 1957). https://doi.org/10.1007/978-3-662-12869-5
29. L.C. Biedenharn, The "Sommerfeld Puzzle" revisited and resolved. Found. Phys. **13**(1), 13–34 (1983). https://doi.org/10.1007/BF01889408
30. J.D. Bjorken, S.D. Drell, *Relativistic Quantum Mechanics*. International Series in Pure and Applied Physics (McGraw-Hill Book, New York, 1964)
31. N.N. Bogolíùbov, D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, 3rd edn. (Wiley, New York, 1980)
32. G. Boole, *A Treatise on the Calculus of Finite Differences*, ed. by J.F. Moulton. Dover Books on Advanced Mathematics (MacMillan and Company, New York, 1872)
33. G. Boole, *A Treatise on Differential Equations*, 5th edn. (Chelsea Publishing Company, New York, 1959)
34. C.P. Boyer, R.T. Sharp, P. Winternitz, Symmetry breaking interactions for the time dependent Schrödinger equation. J. Math. Phys. **17**(8), 1439–1451 (1976). https://doi.org/10.1063/1.523068
35. F. Brafman, On Touchard polynomials. Can. J. Math. **9**, 191–193 (1957). https://doi.org/10.4153/CJM-1957-022-6

36. G. Breit, Possible effects of nuclear spin on X-ray terms. Phys. Rev. **35**(12), 1447–1451 (1930). https://doi.org/10.1103/PhysRev.35.1447
37. G. Breit, G.E. Brown, Effect of nuclear motion on the fine structure of hydrogen. Phys. Rev. **74**(10), 1278–1284 (1948). https://doi.org/10.1103/PhysRev.74.1278
38. T.J.I. Bromwich, *An Introduction to the Theory of Infinite Series*, 2nd edn. (MacMillan, London, 1965)
39. G.E. Brown, Note on a relation in Dirac's theory of the electron. Proc. Natl. Acad. Sci. **36**(1), 15–17 (1950). https://doi.org/10.1073/pnas.36.1.15
40. V.M. Burke, I.P. Grant, The effect of relativity on atomic wave functions. Proc. Phys. Soc. **90**(2), 297–314 (1967). https://doi.org/10.1088/0370-1328/90/2/301
41. L. Carlitz, Some polynomials of Touchard connected with the Bernoulli numbers. Can. J. Math. **9**, 188–190 (1957). https://doi.org/10.4153/CJM-1957-021-9
42. L. Carlitz, Bernoulli and Euler numbers and orthogonal polynomials. Duke Math. J. **26**(1), 1–15 (1959). https://doi.org/10.1215/s0012-7094-59-02601-8
43. D.C. Cassidy, *Uncertainty: The Life and Science of Werner Heisenberg* (W. H. Freeman, New York, 1991). http://www.aip.org/history/heisenberg/p08.htm
44. R. Cordero-Soto, E. Suazo, S.K. Suslov, Quantum integrals of motion for variable quadratic Hamiltonians. Ann. Phys. **325**(9), 1884–1912 (1981). https://doi.org/10.1016/j.aop.2010.02.020
45. R. Cordero-Soto, R.M. López, E. Suazo, S.K. Suslov, Propagator of a charged particle with a spin in uniform magnetic and perpendicular electric fields. Lett. Math. Phys. **84**(2–3), 159–178 (2008). https://doi.org/10.1007/s11005-008-0239-6
46. E.A. Cornell, C.E. Wieman, Nobel lecture: Bose-Einstein condensation in a dilute gas, the first 70 years and some recent experiments. Rev. Mod. Phys. **74**, 875–893 (2002). https://doi.org/10.1103/RevModPhys.74.875
47. F. Dalfovo, S. Giorgini, L.P. Pitaevskii, S. Stringari, Theory of Bose-Einstein condensation in trapped gases. Rev. Mod. Phys. **71**, 463–512 (1999). https://doi.org/10.1103/RevModPhys.71.463
48. C.G. Darwin, The wave equations of the electron. Proc. R. Soc. A Math. Phys. Eng. Sci. **118**(780), 654–680 (1928). https://doi.org/10.1098/rspa.1928.0076
49. L. Davis, A note on the wave functions of the relativistic hydrogenic atom. Phys. Rev. **56**, 186–187 (1939). https://doi.org/10.1103/PhysRev.56.186
50. A.S. Davydov, *Quantum Mechanics* (Pergamon Press, Oxford, 1965)
51. Y.N. Demkov, Channeling, superfocusing, and nuclear reactions. Phys. At. Nucl. **72**(5), 779–785 (2009). https://doi.org/10.1134/s1063778809050056
52. Y.N. Demkov, J.D. Meyer, A sub-atomic microscope, superfocusing in channeling and close encounter atomic and nuclear reactions. Eur. Phys. J. B **42**(3), 361–365 (2004). https://doi.org/10.1140/epjb/e2004-00391-6
53. P.A.M. Dirac, The quantum theory of the electron. Proc. R. Soc. Lond. A **117**(778), 610–624 (1928). https://doi.org/10.1098/rspa.1928.0023
54. P.A.M. Dirac, The quantum theory of the electron. Part II. Proc. R. Soc. Lond. **118**(779), 351–361 (1928). https://doi.org/10.1098/rspa.1928.0056
55. P.A.M. Dirac, *The Principles of Quantum Mechanics*. International Series of Monographs on Physics, vol. 27, 3rd edn. (Clarendon Press, Oxford, 1947)
56. V.V. Dodonov, Current status of the dynamical Casimir effect. Phys. Scr. **82**(3), 38105 (2010). https://doi.org/10.1088/0031-8949/82/03/038105
57. V.V. Dodonov, V.I. Man'ko, Invariants and the evolution of nonstationary quantum system (in Russian), in *Proceedings of Lebedev Physics Institute*, vol. 183 (Nauka, Moscow, 1989), pp. 71–181; English translation published by Nova Science, Commack, New York, 1989, pp. 103–261
58. V.V. Dodonov, I.A. Malkin, V.I. Man'ko, Integrals of the motion, Green functions, and coherent states of dynamical systems. Int. J. Theor. Phys. **14**(1), 37–54 (1975). https://doi.org/10.1007/BF01807990

59. V.V. Dodonov, A.B. Klimov, D.E. Nikonov, Quantum phenomena in nonstationary media. Phys. Rev. A **47**(5), 4422–4429 (1993). https://doi.org/10.1103/PhysRevA.47.4422
60. J. Dorling, Energy levels of the hydrogen atom as a relativistic clock-retardation effect? Am. J. Phys. **38**(4), 510–512 (1970). https://doi.org/10.1119/1.1976376
61. V.A. Dulock, H.V. McIntosh, On the degeneracy of the Kepler problem. Pac. J. Math. **19**(1), 39–55 (1966). https://doi.org/10.2140/pjm.1966.19.39
62. A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, 1953)
63. P. Ehrenfest, Bemerkung über die angenä herte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik (in German). Z. Phys. **45**(7), 455–457 (1927). https://doi.org/10.1007/BF01329203
64. S.T. Epstein, A differential equation for the energy eigenvalues of relativistic hydrogenic atoms, and its solution. Am. J. Phys. **44**(3), 251–252 (1976). https://doi.org/10.1119/1.10466
65. J.H. Epstein, S.T. Epstein, Some applications of hypervirial theorems to the calculation of average values. Am. J. Phys. **30**(4), 266–268 (1962). https://doi.org/10.1119/1.1941987
66. A. Erdélyi, W. Magnus, F. Oberhettinger, F.G. Tricomi, Higher Transcendental Functions, Vols. I–III. Robert E. Krieger, Melbourne (1981). Based on notes left by Harry Bateman, With a preface by Mina Rees, With a foreword by E. C. Watson, Reprint of the 1953 original
67. V.P. Ermakov, Second-order differential equations: conditions of complete integrability. Appl. Anal. Discrete Math. **2**(2), 123–145 (2008). https://doi.org/10.2298/AADM0802123E
68. L.D. Faddeev, The Feynman integral for singular Lagrangians (in Russian). Theor. Math. Phys. **1**(1), 1–13 (1969). https://doi.org/10.1007/BF01028566
69. P.O. Fedichev, Y. Kagan, G.V. Shlyapnikov, J.T.M. Walraven, Influence of nearly resonant light on the scattering length in low-temperature atomic gases. Phys. Rev. Lett. **77**, 2913–2916 (1996). https://doi.org/10.1103/PhysRevLett.77.2913
70. E. Fermi, Quantum theory of radiation. Rev. Mod. Phys. **4**, 87–132 (1932). https://doi.org/10.1103/RevModPhys.4.87
71. E. Fermi, *Notes on Quantum Mechanics*. Phoenix Science, vol. 512 (University of Chicago Press, Chicago, 1961)
72. R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals*. International Series in Pure and Applied Physics (McGraw–Hill Book, New York, 1984)
73. S. Flügge, *Practical Quantum Mechanics*. Classics in Mathematics, vol. 177 (Springer, Berlin, 1999). https://doi.org/10.1007/978-3-642-61995-3
74. V. Fock, Bemerkung zum Virialsatz. Z. Phys. **63**(11), 855–858 (1930). https://doi.org/10.1007/BF01339281
75. V. Fock, Zur theorie des Wasserstoffatoms. Z. Phys. **98**(3), 145–154 (1935). https://doi.org/10.1007/BF01336904
76. J.L. Friar, J.W. Negele, Hypervirial theorems for the Dirac equation. Phys. Rev. C **13**, 1338–1340 (1976). https://doi.org/10.1103/PhysRevC.13.1338
77. T. Fujii, S. Matsuo, N. Hatakenaka, S. Kurihara, A. Zeilinger, Quantum circuit analog of the dynamical Casimir effect. Phys. Rev. B **84**, 174521 (2011). https://doi.org/10.1103/PhysRevB.84.174521
78. R.H. Garstang, D.F. Mayers, Screening constants for relativistic wave functions. Math. Proc. Camb. Philos. Soc. **62**(4), 777–782 (1966). https://doi.org/10.1017/S0305004100040482
79. G. Gasper, M. Rahman, Basic hypergeometric series, in *Encyclopedia of Mathematics and its Applications*, vol. 96, 2nd edn. (Cambridge University Press, Cambridge, 2004). https://doi.org/10.1017/CBO9780511526251
80. J.L. Geronimus, Orthogonal polynomials, in *Two Papers on Special Functions*, ed. by J.L. Geronimus, G. Szegö. American Mathematical Society Translations, vol. 108, chap. 3 (American Mathematical Society, Providence, 1977), pp. 37–130. https://doi.org/10.1090/trans2/108
81. S.P. Goldman, G.W.F. Drake, Relativistic sum rules and integral properties of the Dirac equation. Phys. Rev. A **25**, 2877–2881 (1982). https://doi.org/10.1103/PhysRevA.25.2877

82. I.I. Gol'dman, V.D. Krivchenkov, B.T. Geĭlikman, *Problems in Quantum Mechanics* (Dover Publications, New York, 1993)
83. W. Gordon, Die Energieniveaus des Wasserstoffatoms nach der Diracshen Quantentheorie des Elektrons (in German). Z. Phys. **48**(1), 11–14 (1928). https://doi.org/10.1007/BF01351570
84. J. Guerrero, F.F. López-Ruiz, V. Aldaya, F. Cossío, Harmonic states for the free particle. J. Phys. A Math. Theor. **44**(44), 445307 (2011). https://doi.org/10.1088/1751-8113/44/44/445307
85. A. Gumberidze, T. Stöhlker, D. Banaś, K. Beckert, P. Beller, H.F. Beyer, F. Bosch, S. Hagmann, C. Kozhuharov, D. Liesen, F. Nolden, X. Ma, P.H. Mokler, M. Steck, D. Sierpowski, S. Tashenov, Quantum electrodynamics in strong electric fields: the ground-state Lamb shift in hydrogenlike uranium. Phys. Rev. Lett. **94**, 223001 (2005). https://doi.org/10.1103/PhysRevLett.94.223001
86. A. Gumberidze, T. Stöhlker, D. Banaś, K. Beckert, P. Beller, H.F. Beyer, F. Bosch, X. Cai, S. Hagmann, C. Kozhuharov, D. Liesen, F. Nolden, X. Ma, P.H. Mokler, M. Steck, D. Sierpowski, S. Tashenov, A. Warczak, Y. Zou, Precision tests of QED in strong fields: experiments on hydrogen- and helium-like uranium. J. Phys. Conf. Ser. **58**, 87–92 (2007). https://doi.org/10.1088/1742-6596/58/1/013
87. C.R. Hagen, Scale and conformal transformations in Galilean-covariant field theory. Phys. Rev. D **5**, 377–388 (1972). https://doi.org/10.1103/PhysRevD.5.377
88. G. Harari, Y. Ben-Aryeh, A. Mann, Propagator for the general time-dependent harmonic oscillator with application to an ion trap. Phys. Rev. A **84**, 062104 (2011). https://doi.org/10.1103/PhysRevA.84.062104
89. G.H. Hardy, Notes on special systems of orthogonal functions (III): a system of orthogonal polynomials. Math. Proc. Camb. Philos. Soc. **36**(1), 1–8 (1940). https://doi.org/10.1017/S0305004100016947
90. S. Haroche, J.M. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, Oxford, 2006). https://doi.org/10.1093/acprof:oso/9780198509141.001.0001
91. D.R. Hartree, *The Calculation of Atomic Structures*. Structure of Matter Series (Wiley, New York, 1957)
92. W. Heisenberg, *The Physical Principles of the Quantum Theory*. Dover Books on Physics and Chemistry (University of Chicago Press/Dover Publications, Chicago/New York 1930/1940)
93. W. Heisenberg, *Physics and Philosophy: The Revolution in Modern Science*. Lectures Delivered at University of St. Andrews, Scotland, Winter 1955–1956 (Harper and Row, New York, 1958). http://www.aip.org/history/heisenberg/p13e.htm.
94. R.W. Henry, S.C. Glotzer, A squeezed-state primer. Am. J. Phys. **56**(4), 318–328 (1988). https://doi.org/10.1119/1.15631
95. M.E.H. Ismail, D.R. Masson, M. Rahman, Complex weight functions for classical orthogonal polynomials. Can. J. Math. **43**(6), 1294–1308 (1991). https://doi.org/10.4153/CJM-1991-074-8
96. C. Itzykson, J. Zuber, *Quantum Field Theory*. Dover Books on Physics (Dover Publications, New York, 2005)
97. R. Jackiw, Dynamical symmetry of the magnetic monopole. Ann. Phys. **129**(1), 183–200 (1980). https://doi.org/10.1016/0003-4916(80)90295-X
98. E.G. Kalnins, W. Miller, Lie theory and separation of variables. 5. The equations $iU_t + U_{xx} = 0$ and $iU_t + U_{xx} - c/x^2 U = 0$. J. Math. Phys. **15**(10), 1728–1737 (1974). https://doi.org/10.1063/1.1666533
99. S.G. Karshenboim, V.B. Smirnov, *Precision Physics of Simple Atomic Systems*. Lecture Notes in Physics, vol. 627 (Springer, Berlin, 2003). https://doi.org/10.1007/b13865
100. S.G. Karshenboim, F. Bassani, F. Pavone, M. Inguscio, T. Hänsch, *The Hydrogen Atom: Precision Physics of Simple Atomic Systems*. Lecture Notes in Physics, vol. 570 (Springer, Berlin, 2001). https://doi.org/10.1007/3-540-45395-4
101. I.B. Khriplovich, Fundamental symmetries and atomic physics. Phys. Scr. **T112**(1), 52–62 (2004). https://doi.org/10.1238/physica.topical.112a00052

102. Y.S. Kivshar, T.J. Alexander, S.K. Turitsyn, Nonlinear modes of a macroscopic quantum oscillator. Phys. Lett. A **278**(4), 225–230 (2001). https://doi.org/10.1016/S0375-9601(00)00774-X
103. R. Koekoek, P.A. Lesky, R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics (Springer, Berlin, 2010). https://doi.org/10.1007/978-3-642-05014-5
104. H.T. Koelink, On Jacobi and continuous Hahn polynomials. Proc. Am. Math. Soc. **124**(3), 887–898 (1996). https://doi.org/10.1090/S0002-9939-96-03190-5
105. F. Köhler-Langes, *The Electron Mass and Calcium Isotope Shifts: High-Precision Measurements of Bound-Electron g-Factors of Highly Charged Ions*. Springer Theses (Springer, Berlin, 2017). https://doi.org/10.1007/978-3-319-50877-1
106. C. Koutschan (2019). https://math.la.asu.edu/~suslov/curres/index.htm. See Mathematica notebook: Koutschan.nb.
107. C. Koutschan, P. Paule, S.K. Suslov, Relativistic Coulomb integrals and Zeilberger's holonomic systems approach II, in *Algebraic and Algorithmic Aspects of Differential and Integral Operators*, ed. by M. Barkatou, T. Cluzeau, G. Regensburger, M. Rosenkranz. Lecture Notes in Computer Science; AADIOS 2012, vol. 8372 (Springer, Berlin, 2014), pp. 135–145. https://doi.org/10.1007/978-3-642-54479-8_6
108. C. Koutschan, E. Suazo, S.K. Suslov, Fundamental laser modes in paraxial optics: from computer algebra and simulations to experimental observation. Appl. Phys. B **121**(3), 315–336 (2015). https://doi.org/10.1007/s00340-015-6231-9
109. C. Krattenthaler, S.I. Kryuchkov, A. Mahalov, S.K. Suslov, On the problem of electromagnetic-field quantization. Int. J. Theor. Phys. **52**(12), 4445–4460 (2013). https://doi.org/10.1007/s10773-013-1764-3
110. S.I. Kryuchkov, S.K. Suslov, J.M. Vega-Guzmán, The minimum-uncertainty squeezed states for atoms and photons in a cavity. J. Phys. B Atom. Mol. Opt. Phys. **46**(10), 104007 (2013). https://doi.org/10.1088/0953-4075/46/10/104007 (IOP=Institute Of Physics SELECT and HIGHLIGHT for 2013)
111. S.I. Kryuchkov, N.A. Lanfear, S.K. Suslov, The role of the Pauli-Lubański vector for the Dirac, Weyl, Proca, Maxwell and Fierz-Pauli equations. Phys. Scr. **91**(3), 035301 (2016). https://doi.org/10.1088/0031-8949/91/3/035301
112. S.I. Kryuchkov, E. Suazo, S.K. Suslov, Time-dependent photon statistics in variable media. Math. Methods Appl. Sci. (2018). https://doi.org/10.1002/mma.5285
113. L.D. Landau, E.M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, 3rd edn. (Butterworth-Heinemann, Pergamon, 1977). https://doi.org/10.1016/C2013-0-02793-4
114. N. Lanfear, S.K. Suslov, The time-dependent Schrödinger groups equation, Riccati equation and Airy functions. arXiv e-prints (2009), pp. 1–28
115. N. Lanfear, R.M. López, S.K. Suslov, Exact wave functions for generalized harmonic oscillators. J. Russ. Laser Res. **32**(4), 352–361 (2011). https://doi.org/10.1007/s10946-011-9223-1
116. P. Leach, S. Andriopoulos, The Ermakov equation: a commentary. Appl. Anal. Discrete Math. **2**(2), 146–157 (2008). https://doi.org/10.2298/AADM0802146L
117. D. Leibfried, R. Blatt, C. Monroe, D. Wineland, Quantum dynamics of single trapped ions. Rev. Mod. Phys. **75**(1), 281–324 (2003). https://doi.org/10.1103/RevModPhys.75.281
118. U. Leonhardt, H. Paul, Measuring the quantum state of light. Prog. Quant. Electron. **19**(2), 89–130 (1995). https://doi.org/10.1016/0079-6727(94)00007-L
119. H.R. Lewis, W.B. Riesenfeld, An exact quantum theory of the time-dependent harmonic oscillator and of a charged particle in a time-dependent electromagnetic field. J. Math. Phys. **10**(8), 1458–1473 (1969). https://doi.org/10.1063/1.1664991
120. R. López, S. Suslov, The Cauchy problem for a forced harmonic oscillator. Revista Mexicana de Física **55**, 196–215 (2009)
121. R.M. López, S.K. Suslov, J.M. Vega-Guzmán, On a hidden symmetry of quantum harmonic oscillators. J. Differ. Equ. Appl. **19**(4), 543–554 (2013). https://doi.org/10.1080/10236198.2012.658384

122. R.M. López, S.K. Suslov, J.M. Vega-Guzmán, Reconstructing the Schrödinger groups. Phys. Scr. **87**(3), 038112 (2013). https://doi.org/10.1088/0031-8949/87/03/038112
123. R.M. López, S.K. Suslov, J. Vega-Guzmán (2019). https://math.la.asu.edu/~suslov/curres/index.htm. See Mathematica notebook: HarmonicOscillatorGroup.nb
124. I.A. Malkin, V.I. Man'ko, *Dynamic Symmetry and Coherent States of Quantum Systems* (in Russian) (Nauka, Moscow, 1979)
125. N.H. March, The viral theorem for Dirac's equation. Phys. Rev. **92**(2), 481–482 (1953). https://doi.org/10.1103/PhysRev.92.481
126. H. Margenau, Relativistic magnetic moment of a charged particle. Phys. Rev. **57**(5), 383–386 (1940). https://doi.org/10.1103/PhysRev.57.383
127. M.E. Marhic, Oscillating Hermite-Gaussian wave functions of the harmonic oscillator. Lettere al Nuovo Cimento (1971–1985) **22**(9), 376–378 (1978). https://doi.org/10.1007/BF02820587
128. R.P. Martínez-y-Romero, Relativistic hydrogen atom revisited. Am. J. Phys. **68**(11), 1050–1055 (2000). https://doi.org/10.1119/1.1286314
129. R.J. McKee, μ-Atomic hyperfine structure in the K, L, and M lines of U^{238} and Th^{232}. Phys. Rev. **180**(4), 1139–1158 (1969). https://doi.org/10.1103/PhysRev.180.1139
130. W.A. McKinley, Hellmann-Feynman theorems in classical and quantum mechanics. Am. J. Phys. **39**(8), 905–910 (1971). https://doi.org/10.1119/1.1986322
131. M. Meiler, R. Cordero-Soto, S.K. Suslov, Solution of the Cauchy problem for a time-dependent Schrödinger equation. J. Math. Phys. **49**(7), 072102 (2008). https://doi.org/10.1063/1.2938698
132. E. Merzbacher, *Quantum Mechanics*, 3rd edn. (Wiley, Hoboken, 1998)
133. A. Messiah, *Quantum Mechanics: Two Volumes Bound as One*. Dover Books on Physics (Dover Publications, New York, 1999)
134. W. Miller, *Symmetry and Separation of Variables*. Encyclopedia of Mathematics and its Applications, vol. 4 (Cambridge University Press, Cambridge, 1984). https://doi.org/10.1017/CBO9781107325623
135. P.J. Mohr, G. Plunien, G. Soff, QED corrections in heavy atoms. Phys. Rep. **293**(5), 227–369 (1998). https://doi.org/10.1016/S0370-1573(97)00046-X
136. U. Niederer, The maximal kinematical invariance group of the free Schrödinger equations. Helv. Phys. Acta **45**(5), 802–810 (1972). https://doi.org/10.5169/seals-114417
137. U. Niederer, The maximal kinematical invariance group of the harmonic oscillator. Helv. Phys. Acta **46**(2), 191–200 (1973). https://doi.org/10.5169/seals-114478
138. A.F. Nikiforov, V.B. Uvarov, *Special Functions of Mathematical Physics: A Unified Introduction with Applications* (Birkhäuser, Boston, 1988). https://doi.org/10.1007/978-1-4757-1595-8_1
139. A.F. Nikiforov, S.K. Suslov, V.B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*. Springer Series in Computational Physics (Springer, Berlin, 1991). https://doi.org/10.1007/978-3-642-74748-9
140. A.F. Nikiforov, V.G. Novikov, V.B. Uvarov, *Quantum-Statistical Models of Hot Dense Matter: Methods for Computation Opacity and Equation of State*. Progress in Mathematical Physics, vol. 37 (Birkhäuser, Basel, 2005). https://doi.org/10.1007/b137687
141. S. Pasternack, On the mean value of r^s for Keplerian systems. Proc. Natl. Acad. Sci. **23**(2), 91–94 (1937). https://doi.org/10.1073/pnas.23.2.91
142. S. Pasternack, A generalization of the polynomial $F_n(x)$. Lond. Edinburgh Dublin Philos. Mag. J. Sci. **28**(187), 209–226 (1939). https://doi.org/10.1080/14786443908521175
143. P. Paule, S.K. Suslov, Relativistic Coulomb integrals and Zeilberger's holonomic systems approach. I, in *Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions, Texts & Monographs in Symbolic Computation*, ed. by C. Schneider, J. Blümlein (Springer, Vienna, 2013), pp. 225–241. https://doi.org/10.1007/978-3-7091-1616-6_9
144. L. Pitaevskii, S. Stringari, *Bose-Einstein Condensation, Illustrated*. International Series of Monographs on Physics, reprint edn. (A Clarendon Press, Oxford, 2003)

145. P. Pyykkö, E. Pajanne, M. Inokuti, Hydrogen-like relativistic corrections for electric and magnetic hyperfine integrals. Int. J. Quant. Chem. **7**(4), 785–806 (1973). https://doi.org/10.1002/qua.560070415
146. W.C. Qiang, S.H. Dong, Radial position–momentum uncertainties for the Dirac hydrogen-like atoms. J. Phys. A Math. Gen. **39**(27), 8663–8673 (2006). https://doi.org/10.1088/0305-4470/39/27/007
147. D. Rainville, *Special Functions*. AMS Chelsea Publishing Series (Chelsea Publishing Company, Hartford, 1971)
148. M.E. Rose, *Elementary Theory of Angular Momentum*. Dover Books on Physics and Chemistry (Wiley, New York, 1957); reprinted by Dover, New York (1995)
149. M.E. Rose, *Relativistic Electron Theory*, 1st edn. (Wiley, New York, 1961)
150. M.E. Rose, T.A. Welton, The virial theorem for a Dirac particle. Phys. Rev. Lett. **86**(3), 432–433 (1952). https://doi.org/10.1103/PhysRev.86.432.2
151. S.I. Rosencrans, Perturbation algebra of an elliptic operator. J. Math. Anal. Appl. **56**(2), 317–329 (1976). https://doi.org/10.1016/0022-247X(76)90045-7
152. F. Rosicky, F. Mark, The relativistic virial theorem by the elimination method and nonrelativistic approximations to this theorem. J. Phys. B Atom. Mol. Phys. **8**(16), 2581–2587 (1975). https://doi.org/10.1088/0022-3700/8/16/014
153. P. Rusev, *Analytic Functions and Classical Orthogonal Polynomials*. Bulgarian Mathematical Monographs, vol. 3 (Publishing House of the Bulgarian Academy of Sciences, Sofia, 1984). With a Russian summary
154. B. Sanborn, S.K. Suslov, L. Vinet, Dynamic invariants and the Berry phase for generalized driven harmonic oscillators. J. Russ. Laser Res. **32**(5), 486–494 (2011). https://doi.org/10.1007/s10946-011-9238-7
155. R.M. Schectman, R.H. Good, Generalizations of the virial theorem. Am. J. Phys. **25**(4), 219–225 (1957). https://doi.org/10.1119/1.1934404
156. L.I. Schiff, *Quantum Mechanics*. International Series in Pure and Applied Physics, 3rd edn. (McGraw-Hill, New York, 1968)
157. E. Schrödinger, Der stetige Übergang von der Mikro-zur Makro Mechanik (in German). Naturwissenschaften **14**(28), 664–666 (1926). https://doi.org/10.1007/BF01507634. http://www.nobelprize.org/nobel_prizes/physics/laureates/1933/schrodinger-bio.html and http://www.zbp.univie.ac.at/schrodinger/euebersicht.htm
158. E. Schrödinger, Quantisierung als Eigenwertproblem (in German). Ann. Phys. **79**(6), 489–527 (1926). https://doi.org/10.1002/andp.19263840602; see also *Collected Papers on Wave Mechanics*, Blackie & Son Ltd, London and Glascow, 1928, pp. 13–40, for English translation of Schrödinger's original paper
159. V.M. Shabaev, Recurrence formulas and some exact relations for radial integrals with Dirac and Schrödinger wave functions (in Russian). Vestnik Leningradskogo Universiteta Fizika Khimiya **4**(1), 15–19 (1984)
160. V.M. Shabaev, Generalizations of the virial relations for the Dirac equation in a central field and their applications to the Coulomb field. J. Phys. B Atom. Mol. Opt. Phys. **24**(21), 4479–4488 (1991). https://doi.org/10.1088/0953-4075/24/21/004
161. V.M. Shabaev, Hyperfine structure of hydrogen-like ions. J. Phys. B Atom. Mol. Opt. Phys. **27**(24), 5825–5832 (1994). https://doi.org/10.1088/0953-4075/27/24/006
162. V.M. Shabaev, Relativistic recoil corrections to the atomic energy levels, in *The Hydrogen Atom: Precision Physics of Simple Atomic Systems*, ed. by S.G. Karshenboim, F. Bassani, F.S. Pavone, M. Inguscio, T.W. Hänsch. Lecture Notes in Physics, vol. 570 (Springer, Berlin, 2001), pp. 714–726. https://doi.org/10.1007/3-540-45395-4_51
163. V. Shabaev, Two-time Green's function method in quantum electrodynamics of high-*Z* few-electron atoms. Phys. Rep. **356**(3), 119–228 (2002). https://doi.org/10.1016/S0370-1573(01)00024-2

164. V.M. Shabaev, Virial relations for the Dirac equation and their applications to calculations of hydrogen-like atoms, in *Precision Physics of Simple Atomic Systems*, ed. by S.G. Karshenboim, V.B. Smirnov. Lecture Notes in Physics, vol. 627 (Springer, Berlin, 2003), pp. 97–113. https://doi.org/10.1007/978-3-540-45059-7_6
165. V.M. Shabaev, Quantum electrodynamics of heavy ions and atoms: current status and prospects. Phys. Usp. **51**(11), 1175–1180 (2008). https://doi.org/10.1070/PU2008v051n11ABEH006801
166. V.M. Shabaev, A.N. Artemyev, T. Beier, G. Plunien, V.A. Yerokhin, G. Soff, Recoil correction to the ground-state energy of hydrogenlike atoms. Phys. Rev. A **57**, 4235–4239 (1998). https://doi.org/10.1103/PhysRevA.57.4235
167. V.M. Shabaev, D.A. Glazov, N.S. Oreshkina, A.V. Volotka, G. Plunien, H.J. Kluge, W. Quint, g-factor of heavy ions: a new access to the fine structure constant. Phys. Rev. Lett. **96**, 253002 (2006). https://doi.org/10.1103/PhysRevLett.96.253002
168. A. Sommerfeld, Zur Quantentheorie der Spektrallinien I–II. Ann. Phys. **51**(17), 1–94; **18**, 125–167 (1916). https://doi.org/10.1002/andp.19163561702
169. E. Suazo, S.K. Suslov, Soliton-like solutions for the nonlinear Schrödinger equation with variable quadratic Hamiltonians. J. Russ. Laser Res. **33**(1), 63–83 (2012). https://doi.org/10.1007/s10946-012-9261-3
170. E. Suazo, S.K. Suslov, J.M. Vega-Guzmán, The Riccati differential equation and a diffusion-type equation. New York J. Math. **17 A**, 225–244 (2011)
171. E. Suazo, S.K. Suslov, J.M. Vega-Guzmán, The Riccati system and a diffusion-type equation. MDPI-Mathematics **2**(2), 96–118 (2014). https://doi.org/10.3390/math2020096
172. S.K. Suslov, Matrix elements of Lorentz boosts and the orthogonality of Hahn polynomials on a contour. Sov. J. Nucl. Phys. **36**(4), 621–622 (1982)
173. S.K. Suslov, Hahn polynomials in the Coulomb problem (in Russian). Sov. J. Nucl. Phys. **40**(1), 79–82 (1984).
174. S.K. Suslov, The theory of difference analogues of special functions of hypergeometric type. Russ. Math. Surv. **44**(2), 227–278 (1989). https://doi.org/10.1070/RM1989v044n02ABEH002045
175. S.K. Suslov, *An Introduction to Basic Fourier Series*. Developments in Mathematics, vol. 9 (Springer, Boston, 2003). https://doi.org/10.1007/978-1-4757-3731-8
176. S.K. Suslov, Expectation values in relativistic Coulomb problems. J. Phys. B Atom. Mol. Opt. Phys. **42**(18), 185003 (2009). https://doi.org/10.1088/0953-4075/42/18/185003
177. S.K. Suslov, Dynamical invariants for variable quadratic Hamiltonians. Phys. Scr. **81**(5), 055006 (2010). https://doi.org/10.1088/0031-8949/81/05/055006
178. S.K. Suslov, Mathematical structure of relativistic Coulomb integrals. Phys. Rev. A **81**, 032110 (2010). https://doi.org/10.1103/PhysRevA.81.032110
179. S.K. Suslov, Relativistic Kramers–Pasternack recurrence relations. J. Phys. B Atom. Mol. Opt. Phys. **43**(7), 074006 (2010). https://doi.org/10.1088/0953-4075/43/7/074006
180. S.K. Suslov, On integrability of nonautonomous nonlinear Schrödinger equations. Proc. Am. Math. Soc. **140**(9), 3067–3082 (2012). https://doi.org/10.1090/S0002-9939-2011-11176-6
181. S.K. Suslov, An analogue of the Berry phase for simple harmonic oscillators. Phys. Scr. **87**(3), 038118 (2013). https://doi.org/10.1088/0031-8949/87/03/038118
182. S.K. Suslov (2019). https://math.la.asu.edu/~suslov/curres/index.htm. See Mathematica notebooks: BerrySummary.nb, Fourier.nb and Heisenberg.nb
183. S.K. Suslov, B. Trey, The Hahn polynomials in the nonrelativistic and relativistic Coulomb problems. J. Math. Phys. **49**(1), 012104 (2008). https://doi.org/10.1063/1.2830804
184. G. Szegő, *Orthogonal Polynomials*. American Mathematical Society: Colloquium Publication, vol. 23, 4th edn. (American Mathematical Society, Providence, 1975)
185. P.L. Tchebychef, Sur l'interpolation par la méthode des moindres carrés. Mémoires de l'Académie Impériale des sciences de St.-Pétersbourg, VII^e serie **1**(15), 1–24 (1859). Also in Oeuvres I, pp. 473–498

186. P.L. Tchebychef, Sur l'interpolation des valeurs équidistantes, in *Oeuvres*, P.L. Chebyshev, A.A. Markov, N. Sonin, vol. I (Chelsea Publishing Company, Hartford, 1962), pp. 542–560. Reprint of the 1864 edition
187. P.L. Tchebychef, Sur l'interpolation des valeurs équidistantes, in *Oeuvres*, ed. by P.L. Chebyshev, A.A. Markov, N. Sonin, vol. II (Chelsea Publishing Company, New York, 1962), pp. 219–242. Reprint of the 1875 edition
188. J. Touchard, Nombres exponentiels et nombres de Bernoulli. Can. J. Math. **8**, 305–320 (1956). https://doi.org/10.4153/cjm-1956-034-1
189. D.A. Varshalovich, A.N. Moskalev, V.K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988). https://doi.org/10.1142/0270
190. N.J. Vilenkin, *Special Functions and the Theory of Group Representations*. Translations of Mathematical Monographs, vol. 22 (American Mathematical Society, Providence, 1968). Translated from the Russian by V. N. Singh
191. L. Vinet, A. Zhedanov, Representations of the Schrödinger group and matrix orthogonal polynomials. J. Phys. A Math. Theor. **44**(35), 355201 (2011). https://doi.org/10.1088/1751-8113/44/35/355201
192. E. Vrscay, H. Hamidian, Rayleigh-Schrödinger perturbation theory at large order for radial relativistic Hamiltonians using hypervirial Hellmann–Feynman theorems. Phys. Lett. A **130**(3), 141–146 (1988). https://doi.org/10.1016/0375-9601(88)90417-3
193. G.N. Watson, *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library, vol. 2 (Cambridge University Press, Cambridge 1995)
194. S. Weinberg, *The Quantum Theory of Fields: Volumes 1–3* (Cambridge University Press, Cambridge, 1998)
195. E.T. Whittaker, G.N. Watson, *A Course of Modern Analysis: An Introduction to the General Theory of Infinite Processes and of Analytic Functions; with an Account of the Principal Transcendental Functions*. Cambridge Mathematical Library (Cambridge University Press, Cambridge, 1950) [Reprint of the 4th (1927) edition]
196. E.P. Wigner, *The Application of Group Theory to the Special Functions of Mathematical Physics*. Lectures During the Spring Term of 1955 (Princeton University Press, Princeton, 1955)
197. E.P. Wigner, *Group Theory: And Its Application to the Quantum Mechanics of Atomic Spectra*. Pure and Applied Physics, vol. 5 (Academic, New York, 1959). https://doi.org/10.1016/B978-0-12-750550-3.50003-3. Trans. from the German
198. J.A. Wilson, Hypergeometric series recurrence relations and some new orthogonal functions. Ph.D. thesis, The University of Wisconsin, Madison, 1978
199. J.A. Wilson, Orthogonal functions from Gram determinants. SIAM J. Math. Anal. **22**(4), 1147–1155 (1991). https://doi.org/10.1137/0522074
200. K.B. Wolf, On time-dependent quadratic quantum Hamiltonians. SIAM J. Appl. Math. **40**(3), 419–431 (2005). https://doi.org/10.1137/0140035
201. M.K.F. Wong, H.Y. Yeh, Exact solution of the Dirac-Coulomb equation and its application to bound-state problems. I external fields. Phys. Rev. A **27**(5), 2300–2304 (1983). https://doi.org/10.1103/PhysRevA.27.2300
202. M. Wyman, L. Moser, On some polynomials of Touchard. Can. J. Math. **8**, 321–322 (1956). https://doi.org/10.4153/CJM-1956-035-9
203. K.H. Yeon, K.K. Lee, C.I. Um, T.F. George, L.N. Pandey, Exact quantum theory of a time-dependent bound quadratic Hamiltonian system. Phys. Rev. A **48**(4), 2716–2720 (1993). https://doi.org/10.1103/PhysRevA.48.2716
204. V.A. Yerokhin, Z. Harman, Two-loop QED corrections with closed fermion loops for the bound-electron *g* factor. Phys. Rev. A **88**(4), 042502 (2013). https://doi.org/10.1103/PhysRevA.88.042502

205. J.Q. You, F. Nori, Atomic physics and quantum optics using superconducting circuits. Nature **474**, 589–597 (2011). https://doi.org/10.1038/nature10122
206. H.P. Yuen, Two-photon coherent states of the radiation field. Phys. Rev. A **13**(6), 2226–2243 (1976). https://doi.org/10.1103/PhysRevA.13.2226
207. A.V. Zhukov, Exact quantum theory of a time-dependent system with quadratic Hamiltonian. Phys. Lett. A **256**(5), 325–328 (1999). https://doi.org/10.1016/S0375-9601(99)00247-9

Orthogonal and Multiple Orthogonal Polynomials, Random Matrices, and Painlevé Equations

Walter Van Assche

Abstract Orthogonal polynomials and multiple orthogonal polynomials are interesting special functions because there is a beautiful theory for them, with many examples and useful applications in mathematical physics, numerical analysis, statistics and probability and many other disciplines. In these notes we give an introduction to the use of orthogonal polynomials in random matrix theory, we explain the notion of multiple orthogonal polynomials, and we show the link with certain non-linear difference and differential equations known as Painlevé equations.

Keywords Orthogonal polynomials · Random matrices · Multiple orthogonal polynomials · Painlevé equations

Mathematics Subject Classification (2000) Primary 33C45, 42C05, 60B20, 33E17; Secondary 15B52, 34M55, 41A21

1 Introduction

For these lecture notes I assume the reader is familiar with the basic theory of orthogonal polynomials, in particular the classical orthogonal polynomials (Jacobi, Laguerre, Hermite) should be known. In this introduction we will fix the notation and terminology. Let μ be a positive measure on the real line for which all the moments m_n, $n \in \mathbb{N} = \{0, 1, 2, 3, \dots\}$ exist, where

$$m_n = \int_{\mathbb{R}} x^n \, d\mu(x).$$

W. Van Assche (✉)
Department of Mathematics, KU Leuven, Leuven, Belgium
e-mail: walter.vanassche@kuleuven.be

M. Foupouagnigni, W. Koepf (eds.), *Orthogonal Polynomials*,
Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-030-36744-2_22

The orthonormal polynomials $(p_n)_{n\in\mathbb{N}}$ are such that $p_n(x) = \gamma_n x^n + \cdots$, with $\gamma_n > 0$, satisfying the orthogonality condition

$$\int_{\mathbb{R}} p_n(x)p_m(x)\,d\mu(x) = \delta_{m,n}, \qquad m,n \in \mathbb{N}.$$

It is well known that the zeros of p_n are real and simple, and we denote them by

$$x_{1,n} < x_{2,n} < \cdots < x_{n,n}.$$

Orthonormal polynomials on the real line always satisfy a three-term recurrence relation

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_n p_n(x) + a_n p_{n-1}(x), \qquad n \geq 1, \tag{1.1}$$

with initial condition $p_0 = 1/\sqrt{m_0}$ and $p_{-1} = 0$, with recurrence coefficients $a_{n+1} > 0$ and $b_n \in \mathbb{R}$ for $n \geq 0$. Often we will also use monic orthogonal polynomials, which we denote by capital letters:

$$P_n(x) = \frac{1}{\gamma_n} p_n(x) = x^n + \cdots.$$

Their recurrence relation is of the form

$$P_{n+1}(x) = (x - b_n)P_n(x) - a_n^2 P_{n-1}(x), \tag{1.2}$$

with initial conditions $P_0 = 1$ and $P_{-1} = 0$. The classical families of orthogonal polynomials are

- The *Jacobi* polynomials $P_n^{(\alpha,\beta)}$, for which

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta\,dx = 0, \qquad m \neq n,$$

  with parameters $\alpha, \beta > -1$.
- The *Laguerre polynomials* $L_n^{(\alpha)}$ for which

$$\int_0^\infty L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)x^\alpha e^{-x}\,dx = 0, \qquad m \neq n,$$

  with parameter $\alpha > -1$.
- The *Hermite polynomials* $H_n(x)$ for which

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}\,dx = 0, \qquad m \neq n.$$

Usually these polynomials are neither normalized nor monic but another normalization is used (for historical reasons) and one has to be a bit careful with some of the general formulas for orthonormal or monic orthogonal polynomials.

The matrix

$$H_n = \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_{n-1} \\ m_1 & m_2 & m_3 & \cdots & m_n \\ m_2 & m_3 & m_4 & \cdots & m_{n+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-2} \end{pmatrix} = \left(m_{i+j-2}\right)_{i,j=1}^{n}$$

is the *Hankel matrix* with the moments of the orthogonality measure μ. The *Hankel determinant* is

$$D_n = \det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_{n-1} \\ m_1 & m_2 & m_3 & \cdots & m_n \\ m_2 & m_3 & m_4 & \cdots & m_{n+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-2} \end{pmatrix} = \det\left(m_{i+j-2}\right)_{i,j=1}^{n}. \tag{1.3}$$

If the support of μ contains infinitely many points, then $D_n > 0$ for all $n \in \mathbb{N}$.

The monic orthogonal polynomials $P_n(x)$ are given by

$$P_n(x) = \frac{1}{D_n} \det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & m_3 & \cdots & m_{n+1} \\ m_2 & m_3 & m_4 & \cdots & m_{n+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-1} \\ 1 & x & x^2 & \cdots & x^n \end{pmatrix}, \tag{1.4}$$

and

$$\frac{1}{\gamma_n^2} = \int_{\mathbb{R}} P_n^2(x)\, d\mu(x) = \frac{D_{n+1}}{D_n}. \tag{1.5}$$

The *Christoffel-Darboux kernel* is defined as

$$K_n(x, y) = \sum_{k=0}^{n-1} \gamma_k^2 P_k(x) P_k(y) = \sum_{k=0}^{n-1} p_k(x) p_k(y).$$

This Christoffel-Darboux kernel is a reproducing kernel: for every polynomial q_{n-1} of degree $\leq n-1$ one has

$$\int K_n(x,y)q_{n-1}(y)\,d\mu(y) = q_{n-1}(x).$$

If f is a function in $L^2(\mu)$, then

$$\int K_n(x,y)f(y)\,d\mu(y) = f_{n-1}(x)$$

gives a polynomial of degree $\leq n-1$ which is the least squares approximant of f in the space of polynomials of degree $\leq n-1$. The Christoffel-Darboux kernel is a sum of n terms containing all the polynomials $p_0, p_1, \ldots, p_{n-1}$, but there is a nice formula that expresses the kernel in just two terms containing the polynomials p_{n-1} and p_n only:

Property 1.1 The Christoffel-Darboux formula is

$$\sum_{k=0}^{n-1} \gamma_k^2 P_k(x)P_k(y) = \gamma_{n-1}^2 \frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{x-y},$$

and its confluent version is

$$\sum_{k=0}^{n-1} \gamma_k^2 P_k^2(x) = \gamma_{n-1}^2 \Big(P_n'(x)P_{n-1}(x) - P_{n-1}'(x)P_n(x) \Big).$$

The version for orthonormal polynomials is

Property 1.2 The Christoffel-Darboux formula is

$$\sum_{k=0}^{n-1} p_k(x)p_k(y) = a_n \frac{p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)}{x-y},$$

and its confluent version is

$$\sum_{k=0}^{n-1} p_k^2(x) = a_n \Big(p_n'(x)p_{n-1}(x) - p_{n-1}'(x)p_n(x) \Big).$$

2 Orthogonal Polynomials and Random Matrices

The link between orthogonal polynomials and random matrices is via the Christoffel-Darboux kernel and Heine's formula for orthogonal polynomials, see Property 2.1. Useful references for random matrices are Mehta's book [31], the book by Anderson et al. [1], and Deift's monograph [11]. First of all, let $x_1, x_2, \ldots, x_n$ be real or complex numbers, then we define the *Vandermonde determinant* as

$$\Delta_n(x_1, \ldots, x_n) = \det \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}. \tag{2.1}$$

This Vandermonde determinant can be evaluated explicitly:

$$\Delta_n = \prod_{i>j} (x_i - x_j).$$

From this it is clear that $\Delta_n \neq 0$ when all the x_i are distinct, and if $x_1 < x_2 < \cdots < x_n$ then $\Delta_n > 0$. Heine's formula expresses the Hankel determinant with the moments of a measure μ as an n-fold integral:

Property 2.1 (Heine) The Hankel determinants D_n in (1.3) can be written as

$$D_n = \frac{1}{n!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Delta_n^2(x_1, \ldots, x_n)\, d\mu(x_1) \cdots d\mu(x_n), \tag{2.2}$$

where Δ_n is the Vandermonde determinant (2.1). Furthermore, the monic orthogonal polynomial $P_n(x)$ is also given by an n-fold integral

$$P_n(x) = \frac{1}{n! D_n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n} (x - x_i)\; \Delta_n^2(x_1, \ldots, x_n)\, d\mu(x_1) \cdots d\mu(x_n). \tag{2.3}$$

Proof If we write all the moments in the first row of (1.3) as an integral and use linearity of the determinant (for one row), then

$$D_n = \int_{-\infty}^{\infty} \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ m_1 & m_2 & m_3 & \cdots & m_n \\ m_2 & m_3 & m_4 & \cdots & m_{n+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-2} \end{pmatrix} d\mu(x_1).$$

Repeating this for every row gives

$$D_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ x_2 & x_2^2 & x_2^3 & \cdots & x_2^n \\ x_3^2 & x_3^3 & x_3^4 & \cdots & x_3^{n+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_n^{n-1} & x_n^n & x_n^{n+1} & \cdots & x_n^{2n-2} \end{pmatrix} d\mu(x_1) \cdots d\mu(x_n).$$

In each row we can take out the common factors to find

$$D_n = \int_{\mathbb{R}^n} \prod_{j=1}^n x_j^{j-1} \, \Delta_n(x_1, \ldots, x_n) \, d\mu(x_1) \cdots d\mu(x_n).$$

Now write the integral over $\mathbb{R}^n$ as a sum of integrals over all simplices $x_{i_1} < x_{i_2} < \cdots < x_{i_n}$, where $\sigma = (i_1, i_2, \ldots, i_n)$ is a permutation of $(1, 2, \ldots, n)$. Then

$$D_n = \sum_{\sigma \in S_n} \int_{x_{\sigma(1)} < \cdots < x_{\sigma(n)}} \prod_{j=1}^n x_j^{j-1} \, \Delta_n(x_1, x_2, \ldots, x_n) \, d\mu(x_1) \cdots d\mu(x_n).$$

With the change of variables $x_{\sigma(j)} = y_j$ one has $x_j = y_{\tau(j)}$, with $\tau = \sigma^{-1}$ and

$$D_n = \int_{y_1 < \cdots < y_n} \sum_{\tau \in S_n} \prod_{j=1}^n y_{\tau(j)}^{j-1} \, \Delta_n(y_{\tau(1)}, \ldots, y_{\tau(n)}) \, d\mu(y_1) \cdots d\mu(y_n).$$

Observe that $\Delta_n(y_{\tau(1)}, \ldots, y_{\tau(n)}) = \mathrm{sign}(\tau) \Delta_n(y_1, \ldots, y_n)$, so that

$$D_n = \int_{y_1 < \cdots < y_n} \left(\sum_{\tau \in S_n} \mathrm{sign}(\tau) \prod_{j=1}^n y_{\tau(j)}^{j-1} \right) \Delta_n(y_1, \ldots, y_n) \, d\mu(y_1) \cdots d\mu(y_n).$$

Now use

$$\sum_{\tau \in S_n} \mathrm{sign}(\tau) \prod_{j=1}^n y_{\tau(j)}^{j-1} = \Delta_n(y_1, \ldots, y_n)$$

to find

$$D_n = \int_{y_1 < \cdots < y_n} \Delta_n^2(y_1, \ldots, y_n) \, d\mu(y_1) \cdots d\mu(y_n).$$

This is an integral over one simplex $y_1 < y_2 < \cdots < y_n$ in $\mathbb{R}^n$. This integral is the same for every simplex, and since there are $n!$ simplices (because there are $n!$ permutations of $(1, 2, \ldots, n)$), we find the required formula (2.2).

The proof for formula (2.3) is similar, using the determinant expression (1.4) for the monic orthogonal polynomial. □

It is remarkable that Szegő writes in his book [40]:

> [These] Formulas ... are not suitable in general for derivation of properties of the polynomials in question. To this end we shall generally prefer the orthogonality property itself, or other representations derived by means of the orthogonality property.

Heine's formulas have now become crucial in the theory of random matrices.

2.1 Point Processes

A n-point process is a stochastic process where a set of n points $\{X_1, \ldots, X_n\}$ is selected, and the joint distribution of the random variables $(X_1, X_2, \ldots, X_n)$ is given. Since we are dealing with a set of n random numbers, the order of the random variables is irrelevant and hence we use a probability distribution which is invariant under permutations. Our interest is in the n-point process where the joint probability distribution has a density (with respect to the product measure $d\mu(x_1) \ldots d\mu(x_n)$) given by

$$P(x_1, x_2, \ldots, x_n) = \frac{1}{n! D_n} \Delta_n^2(x_1, \ldots, x_n), \tag{2.4}$$

where we mean that

$$\text{Prob}(X_1 \le y_1, \ldots, X_n \le y_n) = \int_{-\infty}^{y_1} \ldots \int_{-\infty}^{y_n} P(x_1, \ldots, x_n)\, d\mu(x_1) \cdots d\mu(x_n).$$

Observe that by Heine's formula (2.2) this is indeed a probability distribution since it is positive and integrates over $\mathbb{R}^n$ to one. The points in this n-point process are not independent and the factor $\Delta_n^2(x_1, \ldots, x_n)$ describes the dependence of the points. Two points are unlikely to be close together because then $\Delta_n^2(x_1, \ldots, x_n) = \prod_{j>i}(x_j - x_i)^2$ is small and by the maximum likelihood principle the points will prefer to choose a position that maximizes $\Delta_n^2(x_1, \ldots, x_n)$. This n-point process therefore has points that repel each other.

An important property of this n-point process is that it is a *determinantal point process*. To see this, we will express the probability density in terms of the Christoffel-Darboux kernel. We need a few important properties of that kernel.

Property 2.2 The Christoffel-Darboux kernel satisfies

$$\int_{-\infty}^{\infty} K_n(x,y)K_n(y,z)\,d\mu(y) = K_n(x,z),$$

and

$$\int_{-\infty}^{\infty} K_n(x,x)\,d\mu(x) = n.$$

Proof The first property follows from the reproducing property of the Christoffel-Darboux kernel. For the second property we have

$$\int_{-\infty}^{\infty} K_n(x,x)\,d\mu(x) = \sum_{k=0}^{n-1}\int_{-\infty}^{\infty} p_k^2(x)\,d\mu(x) = n.$$

□

Property 2.3 The density (2.4) can be written as

$$P(x_1,x_2,\ldots,x_n) = \frac{1}{n!}\det\bigl(K_n(x_i,x_j)\bigr)_{i,j=1}^{n},$$

where K_n is the Christoffel-Darboux kernel.

Proof If we add rows in the Vandermonde determinant (2.1), then

$$\Delta_n(x_1,\ldots,x_n) = \det\begin{pmatrix} P_0(x_1) & P_0(x_2) & P_0(x_3) & \cdots & P_0(x_n) \\ P_1(x_1) & P_1(x_2) & P_1(x_3) & \cdots & P_1(x_n) \\ P_2(x_1) & P_2(x_2) & P_2(x_3) & \cdots & P_2(x_n) \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ P_{n-1}(x_1) & P_{n-1}(x_2) & P_{n-1}(x_3) & \cdots & P_{n-1}(x_n) \end{pmatrix},$$

for any sequence $(P_0,P_1,P_2,\ldots,P_{n-1})$ of monic polynomials. If we take the monic orthogonal polynomials, then

$$\begin{aligned}&\left(\prod_{j=0}^{n-1}\gamma_j^2\right)\Delta_n^2(x_1,\ldots,x_n)\\ &= \det\begin{pmatrix} P_0(x_1) & P_1(x_1) & \cdots & P_{n-1}(x_1) \\ P_0(x_2) & P_1(x_2) & \cdots & P_{n-1}(x_2) \\ P_0(x_3) & P_1(x_3) & \cdots & P_{n-1}(x_3) \\ \vdots & \vdots & \ldots & \vdots \\ P_0(x_n) & P_1(x_n) & \cdots & P_{n-1}(x_n) \end{pmatrix} \Gamma_n \begin{pmatrix} P_0(x_1) & P_0(x_2) & \cdots & P_0(x_n) \\ P_1(x_1) & P_1(x_2) & \cdots & P_1(x_n) \\ P_2(x_1) & P_2(x_2) & \cdots & P_2(x_n) \\ \vdots & \vdots & \ldots & \vdots \\ P_{n-1}(x_1) & P_{n-1}(x_2) & \cdots & P_{n-1}(x_n) \end{pmatrix},\end{aligned}$$

where $\Gamma_n = \mathrm{diag}(\gamma_0^2, \gamma_1^2, \ldots, \gamma_{n-1}^2)$. Then use (1.5) to find that $\prod_{j=0}^{n-1} \gamma_j^2 = 1/D_n$, so that

$$\Delta_n^2(x_1, \ldots, x_n) = D_n \det\Big(\sum_{k=0}^{n-1} \gamma_k^2 P_k(x_i) P_k(x_j)\Big)_{i,j=1}^n,$$

which combined with (2.4) gives the required result. □

For this reason we call the n-point process with density (2.4) the *Christoffel-Darboux point process*.

2.2 Determinantal Point Process

The fact that the density $P(x_1, \ldots, x_n)$ can be written as a determinant of a kernel function $K(x, y)$ that satisfies Property 2.2 is important and allows to compute correlation functions for k points $k \leq n$ of the point process, in particular the probability density of one point (for $k = 1$).

Definition 2.4 For $k \leq n$ the kth correlation function is

$$\rho_k(x_1, \ldots, x_k) = \det\Big(K_n(x_i, x_j)\Big)_{i,j=1}^k.$$

The interpretation of these kth correlation functions is the following: if $A_i \cap A_j = \emptyset$ $(i \neq j)$, and $N(A)$ is the number of points in A, then

$$\int_{A_1} \int_{A_2} \cdots \int_{A_k} \rho_k(x_1, \ldots, x_k)\, d\mu(x_1) \cdots d\mu(x_k) = \mathbb{E}\left(\prod_{i=1}^k N(A_i)\right).$$

The kth correlation function can also be seen as the density of the marginal distribution of k points in the n-point process, up to a normalization factor:

Property 2.5 The kth correlation function is obtained from $P(x_1, \ldots, x_n)$ by

$$\rho_k(x_1, x_2, \ldots, x_k) = \frac{n!}{(n-k)!} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-k} P(x_1, \ldots, x_n)\, d\mu(x_{k+1}) \cdots d\mu(x_n).$$

Proof For $k = n - 1$ we have, by expanding the determinant along the last row,

$$\int_{-\infty}^{\infty} P(x_1, \dots, x_n)\, d\mu(x_n)$$
$$= \frac{1}{n!} \sum_{k=1}^{n-1} \int_{-\infty}^{\infty} (-1)^{n+k} K_n(x_n, x_k) \det\Big(K_n(x_i, x_j)\Big)_{1\le i\ne n,\, j\ne k\le n} d\mu(x_n)$$
$$+ \frac{1}{n!} \int_{-\infty}^{\infty} K_n(x_n, x_n) \det\Big(K_n(x_i, x_j)\Big)_{i,j=1}^{n-1} d\mu(x_n).$$

By Property 2.2 the last term is $1/(n-1)!\rho_{n-1}(x_1, \dots, x_{n-1})$. Expanding the remaining determinant along the last column gives

$$\frac{1}{n!} \sum_{k=1}^{n-1} \sum_{\ell=1}^{n-1} (-1)^{n+k} (-1)^{n-1+\ell} \int_{-\infty}^{\infty} K_n(x_n, x_k) K_n(x_\ell, x_n)$$
$$\times \det\Big(K_n(x_i, x_j)\Big)_{1\le i\ne \ell,\, j\ne k\le n-1} d\mu(x_n).$$

The determinant does not contain x_n, so the remaining integration can be done using Property 2.2 and gives

$$\frac{1}{n!} \sum_{k=1}^{n-1} \sum_{\ell=1}^{n-1} (-1)^{k+\ell-1} K_n(x_\ell, x_k) \det\Big(K_n(x_i, x_j)\Big)_{1\le i\ne \ell,\, j\ne k\le n-1}.$$

The sum over ℓ gives the $(n-1) \times (n-1)$ determinant (recall that column k which contains $K_n(x_i, x_k)$ is missing since $j \ne k$)

$$(-1)^n \det \begin{pmatrix} K_n(x_1, x_1) & K_n(x_1, x_2) & \cdots & K_n(x_1, x_{n-1}) & K_n(x_1, x_k) \\ K_n(x_2, x_1) & K_n(x_2, x_2) & \cdots & K_n(x_2, x_{n-1}) & K_n(x_2, x_k) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ K_n(x_{n-1}, x_1) & K_n(x_{n-1}, x_2) & \cdots & K_n(x_{n-1}, x_{n-1}) & K_n(x_{n-1}, x_k) \end{pmatrix},$$

and to get the last column in the kth position, we need to interchange columns $n - 1 - k$ times, which gives

$$\int_{-\infty}^{\infty} P(x_1, \dots, x_n)\, d\mu(x_n)$$
$$= \frac{-1}{n!} \sum_{k=1}^{n-1} \rho_{n-1}(x_1, \dots, x_{n-1}) + \frac{1}{(n-1)!} \rho_{n-1}(x_1, \dots, x_{n-1}),$$

and hence

$$\rho_{n-1}(x_1,\ldots,x_{n-1}) = n!\int_{-\infty}^{\infty} P(x_1,\ldots,x_n)\,d\mu(x_n).$$

To prove the case for all $k = n - m$ one uses induction on m, for which we just proved the case $m = 1$. □

Definition 2.6 A point process on $\mathbb{R}$ with correlation functions ρ_k is a *determinantal point process* if there exists a kernel $K(x, y)$ such that for every k and every $x_1,\ldots,x_k \in \mathbb{R}$

$$\rho_k(x_1,x_2,\ldots,x_k) = \det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{k}.$$

The following theorem shows that Property 2.2 is indeed crucial.

Theorem 2.7 *Suppose $K : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is a kernel such that*

- $\int_{-\infty}^{\infty} K(x,x)\,dx = n \in \mathbb{N}$,
- *For every $x_1,\ldots,x_n \in \mathbb{R}$, one has $\det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{k} \geq 0$.*
- $K(x,y) = \int_{-\infty}^{\infty} K(x,s)K(s,y)\,ds$.

Then

$$P(x_1,\ldots,x_n) = \frac{1}{n!}\det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{n}$$

is a probability density on $\mathbb{R}^n$ which is invariant under permutations of coordinates. The associated n-point process is determinantal.

The most important example (at least in the context of this section) is when $d\mu(x) = w(x)\,dx$, and then one can take

$$K(x,y) = K_n(x,y)\sqrt{w(x)}\sqrt{w(y)}.$$

2.3 Random Matrices

To see the relation with random matrices, we claim that the eigenvalues of certain random matrices of order n form a determinantal point process with the Christoffel-Darboux kernel for a particular family of orthogonal polynomials. The *Gaussian unitary ensemble* (GUE) consists of Hermitian random matrices $\mathbf{M}$ of order n with random entries

$$\mathbf{M}_{k,\ell} = X_{k,\ell} + iY_{k,\ell}, \quad \mathbf{M}_{\ell,k} = X_{k,\ell} - iY_{k,\ell}, \qquad k < \ell,$$

$$\mathbf{M}_{k,k} = X_{k,k}, \qquad 1 \leq k \leq n,$$

where all $X_{k,\ell}$, $Y_{k,\ell}$, $X_{k,k}$ are independent normal random variables with mean zero and variance $\frac{1}{4n}$ (if $k < \ell$) or $\frac{1}{2n}$ (if $k = \ell$). The multivariate density is

$$\frac{1}{Z_n} \prod_{k<\ell} e^{-2n(x_{k,\ell}^2+y_{k,\ell}^2)} \prod_{k=1}^{n} e^{-nx_{k,k}^2} \prod_{k<\ell} dx_{k,\ell} dy_{k,\ell} \prod_{k=1}^{n} dx_{k,k},$$

where Z_n is normalizing constant. But this is also equal to

$$\frac{1}{Z_n} \exp(-n\mathrm{Tr}\, M^2)\, dM$$

where $M_{k,\ell} = (x_{k,\ell} + iy_{k,\ell})$ for $k < \ell$, $M_{k,k} = x_{k,k}$, and $M = M^*$.

We are mostly interested in the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the random matrix $\mathbf{M}$. To find the density of the eigenvalues, we use the change of variables: $M \mapsto (\Lambda, U)$, where U is a unitary matrix for which

$$M = U\Lambda U^*,$$

and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, and then integrate over the unitary part U, which leaves only the eigenvalues. This change of variables is done using the Weyl integration formula (see, e.g., [1, §4.1.3]):

Theorem 2.8 (Weyl Integration Formula) *For the change of variables $M = U\Lambda U^*$ one has*

$$dM = c_n \prod_{i<j} (\lambda_i - \lambda_j)^2\, d\lambda_1 \cdots d\lambda_n\, dU,$$

where c_n is a constant and dU is the Haar measure on the unitary group.

We will use a simplified version of this result, for which one does not need the Haar measure on the unitary group. This works when the expression $f(M)$ that we want to integrate only depends on the eigenvalues of M. Let $\mathcal{H}_n$ be the Hermitian matrices of order n.

Definition 2.9 A function $f : \mathcal{H}_n \to \mathbb{C}$ is a *class function* if

$$f(UMU^*) = f(M)$$

for all unitary matrices U.

Theorem 2.10 (Weyl Integration Formula for Class Functions) *For an integrable class function f we have*

$$\int f(M)\, dM = c_n \int_{\mathbb{R}^n} f(\lambda_1, \ldots, \lambda_n) \prod_{i<j} (\lambda_i - \lambda_j)^2\, d\lambda_1 \cdots d\lambda_n,$$

with

$$c_n = \frac{\pi^{n(n-1)/2}}{\prod_{j=1}^{n} j!}.$$

The characteristic polynomial of a matrix M only depends on the eigenvalues, hence $\det(xI - M)$ is a class function. For random matrices in GUE one finds for the average characteristic function

$$\mathbb{E}\det(xI - \mathbf{M}) = \frac{1}{D_n}\int_{\mathbb{R}^n}\prod_{i=1}^{n}(x - x_i)\ \Delta_n^2(x_1,\ldots,x_n)e^{-n(x_1^2+\cdots+x_n^2)}\,dx_1\cdots dx_n \tag{2.5}$$

and by (2.3) this is the monic *Hermite polynomial* $H_n(\sqrt{n}x)$. More generally, the eigenvalues of a random matrix in GUE form a determinantal point process with the Christoffel-Darboux kernel of (scaled) Hermite polynomials. The average number of eigenvalues of $\mathbf{M}$ in $[a, b]$ is in terms of the correlation function $\rho_1(x)$:

$$\mathbb{E}\big(N([a,b])\big) = \int_a^b K_n(x,x)\,e^{-nx^2}\,dx.$$

2.4 Random Matrix Ensembles

Here we give a few more random matrix ensembles for which the eigenvalues form a determinantal point process with the Christoffel-Darboux kernel of classical orthogonal polynomials.

- We already defined GUE (Gaussian Unitary Ensemble): this contains random matrices in $\mathcal{H}_n$ with density

  $$\frac{1}{Z_n}\exp(-n\mathrm{Tr}\ M^2)\,dM.$$

  The average characteristic polynomial is

  $$\mathbb{E}\det(xI - \mathbf{M}) = \text{(scaled) Hermite polynomial.}$$

  This suggests that on the average the eigenvalues behave like the zeros of (scaled) Hermite polynomials. This is indeed true, but for this one needs the correlation function ρ_1 and the result that

  $$\lim_{n\to\infty}\frac{1}{n}\int_a^b f(x)K_n(x,x)e^{-nx^2}\,dx = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} f(x_{j,n}/\sqrt{n}),$$

  where $x_{1,n},\ldots,x_{n,n}$ are the zeros of the Hermite polynomial H_n.

- The Wishart ensemble. Let $\mathbf{M}$ be a $n \times m$ matrix $(m \geq n)$ with independent complex Gaussian entries $X_{k,\ell} + iY_{k,\ell}$. Then $\mathbf{MM}^*$ has the Wishart distribution with density

$$\frac{1}{C_n} |\det W|^{m-n} \exp(-Tr W).$$

The average characteristic polynomial is

$$\mathbb{E} \det(xI - \mathbf{MM}^*) = \text{ Laguerre polynomial with } \alpha = m - n.$$

Observe that $\mathbf{MM}^*$ is a positive definite matrix so that all the eigenvalues are positive. On the average they behave like the zeros of Laguerre polynomials.

- Truncated unitary matrices. Let U be a random unitary matrix of order $(m+k) \times (m+k)$ and let $\mathbf{V}$ be the $m \times n$ upper left corner $(m \geq n)$. Then $\mathbf{V}^*\mathbf{V}$ is an $n \times n$ matrix and

$$\mathbb{E} \det(xI - \mathbf{V}^*\mathbf{V}) = \text{Jacobi polynomial on } [0, 1], \quad \alpha = m - n, \beta = k - n.$$

Unitary matrices have their eigenvalues on the unit circle, and a truncated unitary matrix has its singular values (the eigenvalues of $\mathbf{V}^*\mathbf{V}$) in [0, 1]. These eigenvalues behave on the average like the zeros of Jacobi polynomials.

Exercise Let $\mathbf{M}_n$ be the Hermitian random matrix with entries

$$(\mathbf{M}_n)_{k,\ell} = \begin{cases} X_{k,\ell} + iY_{k,\ell}, & k < \ell, \\ X_{\ell,k} - iY_{\ell,k}, & k > \ell, \\ X_{k,k}, & k = \ell, \end{cases}$$

where $X_{k,\ell}, Y_{k,\ell}$ $(k < \ell)$ and $X_{k,k}$ $(1 \leq k \leq n)$ are independent random variables with means $\mathbb{E}(X_{k,\ell}) = \mathbb{E}(Y_{k,\ell}) = \mathbb{E}(X_{k,k}) = 0$ and variances $\mathbb{E}(X_{k,\ell}^2) = \mathbb{E}(Y_{k,\ell}^2) = \mathbb{E}(X_{k,k}^2) = \sigma^2 > 0$. Show that $P_n(x) = \mathbb{E} \det(xI_n - \mathbf{M}_n)$ satisfies the three-term recurrence relation

$$P_n(x) = xP_{n-1}(x) - 2(n-1)\sigma^2 P_{n-2}(x),$$

with $P_0(x) = 1$ and $P_1(x) = x$. Identify this $P_n(x)$ as $\sigma^n H_n(x/2\sigma)$, where H_n is the Hermite polynomial of degree n. This shows that the Hermite polynomial is the average characteristic polynomial of a large class of Hermitian random matrices, not only GUE.

So far we found that on the average the eigenvalues of random matrices from these ensembles behave like zeros of orthogonal polynomials. To get more information about individual eigenvalues, for example the largest eigenvalue or the smallest eigenvalue, one needs a more detailed analysis of the point process. In particular one needs to investigate the asymptotic behavior of the Christoffel-Darboux kernels. In particular, to understand the spacing between the eigenvalues in the neighborhood of x^* in the bulk of the spectrum, one needs results for

$$\lim_{n\to\infty} \frac{1}{n} K_n\big(x^* + \frac{u}{n}, x^* + \frac{v}{n}\big),$$

or, when x^* is at the end of the spectrum,

$$\lim_{n\to\infty} \frac{1}{n^\gamma} K_n\big(x^* + \frac{u}{n^\gamma}, x^* + \frac{v}{n^\gamma}\big),$$

where γ depends on the nature of the endpoint (hard or soft edge). This will give kernels of well-known point processes.

An important quantity of interest is the probability $p_A(m)$ that there are exactly m eigenvalues in the set $A \subset \mathbb{R}$. If there are m eigenvalues in A, then the number of ordered k-tuples in A is $\binom{m}{k}$ and thus

$$\sum_{m=k}^{\infty} \binom{m}{k} p_A(m) = \frac{1}{k!} \int_{A^k} \rho_k(x_1, \ldots, x_k)\, d\mu(x_1) \cdots d\mu(x_k), \qquad k \geq 1,$$

because this is the expected number of ordered k-tuples in A. For $k = 0$ one has

$$\sum_{m=0}^{\infty} p_A(m) = 1,$$

therefore

$$1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^k} \rho_k(x_1, \ldots, x_k)\, d\mu(x_1) \cdots d\mu(x_k) = \sum_{k=0}^{\infty} \sum_{m=k}^{\infty} (-1)^k \binom{m}{k} p_A(m).$$

Changing the order of summation (we assume that this is allowed) and using

$$\sum_{k=0}^{m} (-1)^k \binom{m}{k} = \delta_{m,0},$$

we find that

$$p_A(0) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^k} \rho_k(x_1, \ldots, x_k)\, d\mu(x_1) \cdots d\mu(x_k).$$

This is the so-called *gap probability*: the probability to find no eigenvalues in A. For a determinantal point process, such as the eigenvalues of various random matrices, this gap probability is in fact the Fredholm determinant $\det(I - K_A)$ of the operator $K_A : L^2(A) \to L^2(A)$ defined by

$$K_A f(x) = \int_A K_n(x, y) f(y)\, d\mu(y), \qquad x \in A.$$

The asymptotic behavior as the size n of the random matrices increases to infinity, then gives the Fredholm determinant $\det(I - K_A)$ of the operator K_A that uses the kernel $K(x, y)$ which is the limit of the Christoffel-Darboux kernel $K_n(x, y)$ as described above. The lesson to be learned from this is that the asymptotic behavior of orthogonal polynomials and their Christoffel-Darboux kernel gives important insight in the behavior of eigenvalues of random matrices.

3 Multiple Orthogonal Polynomials

In this section we will explain the notion of multiple orthogonal polynomials. Useful references are Ismail's book [20, Ch. 23], Nikishin and Sorokin's book [33, Ch. 4] and the papers [2, 29, 48]. Instead of orthogonality conditions with respect to one measure on the real line, the orthogonality will be with respect to r measures, where $r \geq 1$. For $r = 1$ one has the usual orthogonal polynomials, but for $r \geq 2$ one gets two types of multiple orthogonal polynomials.

Let $r \in \mathbb{N}$ and let $\mu_1, \ldots, \mu_r$ be positive measures on the real line, for which all the moments exist. We use *multi-indices* $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$ and denote their *length* by $|\vec{n}| = n_1 + n_2 + \cdots + n_r$.

Definition 3.1 (Type I) Type I multiple orthogonal polynomials for $\vec{n}$ consist of the vector $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ of r polynomials, with $\deg A_{\vec{n},j} \leq n_j - 1$, for which

$$\int x^k \sum_{j=1}^{r} A_{\vec{n},j}(x)\, d\mu_j(x) = 0, \qquad 0 \leq k \leq |\vec{n}| - 2,$$

with normalization

$$\int x^{|\vec{n}|-1} \sum_{j=1}^{r} A_{\vec{n},j}(x)\, d\mu_j(x) = 1.$$

Definition 3.2 (Type II) The type II multiple orthogonal polynomial for $\vec{n}$ is the **monic** polynomial $P_{\vec{n}}$ of degree $|\vec{n}|$ for which

$$\int x^k P_{\vec{n}}(x)\, d\mu_j(x) = 0, \qquad 0 \leq k \leq n_j - 1,$$

for $1 \leq j \leq r$.

The conditions for type I and type II multiple orthogonal polynomials give a system of $|\vec{n}|$ linear equations for the $|\vec{n}|$ unknown coefficients of the polynomials. This system may not have a solution, or when a solution exists it may not be unique. A multi-index $\vec{n}$ is said to be *normal* if the type I vector $(A_{\vec{n},1}, \dots A_{\vec{n},r})$ exists and is unique, and this is equivalent with the existence and uniqueness of the monic type II multiple orthogonal polynomial $P_{\vec{n}}$, because the matrix of the linear system for type II is the transpose of the matrix for the type I linear system. Hence $\vec{n}$ is a normal multi-index if and only if

$$\det \begin{pmatrix} M_{n_1}^{(1)} \\ M_{n_2}^{(2)} \\ \vdots \\ M_{n_r}^{(r)} \end{pmatrix} \neq 0,$$

where

$$M_{n_j}^{(j)} = \begin{pmatrix} m_0^{(j)} & m_1^{(j)} & \cdots & m_{|\vec{n}|-1}^{(j)} \\ m_1^{(j)} & m_2^{(j)} & \cdots & m_{|\vec{n}|}^{(j)} \\ \vdots & \vdots & \cdots & \vdots \\ m_{n_j-1}^{(j)} & m_{n_j}^{(j)} & \cdots & m_{|\vec{n}|+n_j-2}^{(j)} \end{pmatrix}$$

are rectangular Hankel matrices containing the moments

$$m_k^{(j)} = \int x^k\, d\mu_j(x).$$

3.1 *Special Systems*

Interesting systems of measures $(\mu_1, \dots, \mu_r)$ are those for which all the multi-indices are normal. We call such systems *perfect*. Here we will describe two such systems.

Definition 3.3 (Angelesco System) The measures $(\mu_1, \ldots, \mu_r)$ are an *Angelesco system* if the supports of the measures are subsets of disjoint intervals Δ_j, i.e., $\mathrm{supp}(\mu_j) \subset \Delta_j$ and $\Delta_i \cap \Delta_j = \emptyset$ whenever $i \neq j$.

Usually one allows that the intervals are touching, i.e., $\overset{\circ}{\Delta}_i \cap \overset{\circ}{\Delta}_j = \emptyset$ whenever $i \neq j$.

Theorem 3.4 (Angelesco, Nikishin) *The type II multiple orthogonal polynomial $P_{\vec{n}}$ for an Angelesco system has exactly n_j distinct zeros on $\overset{\circ}{\Delta}_j$ for $1 \leq j \leq r$.*

This means that the type II multiple orthogonal polynomial $P_{\vec{n}}$ can be factored as $P_{\vec{n}}(x) = \prod_{j=1}^{r} p_{\vec{n},j}(x)$, where $p_{\vec{n},j}$ has all its zeros on Δ_j. In fact, $p_{\vec{n},j}$ is an ordinary orthogonal polynomial of degree n_j on the interval Δ_j for the measure $\prod_{i \neq j} p_{\vec{n},i}(x)\, d\mu_j(x)$:

$$\int_{\Delta_j} x^k p_{\vec{n},j}(x) \prod_{i \neq j} p_{\vec{n},i}\, d\mu(x) = 0, \qquad 0 \leq k \leq n_j - 1.$$

Observe that for $i \neq j$ the polynomial $p_{\vec{n},i}(x)$ has constant sign on Δ_j.

Corollary 3.5 *Every multi-index $\vec{n}$ is normal (an Angelesco system is perfect).*

Exercise Show that every $A_{\vec{n},j}$ has $n_j - 1$ zeros on $\overset{\circ}{\Delta}_j$.

For another system of measures, which are all supported on the same interval $[a, b]$, we need to recall the notion of a Chebyshev system.

Definition 3.6 The functions $\varphi_1, \ldots, \varphi_n$ are a **Chebyshev system** on $[a, b]$ if every linear combination $\sum_{i=1}^{n} a_i \varphi_i$ with $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$ has at most $n - 1$ zeros on $[a, b]$.

We can then define an *Algebraic Chebyshev system*:

Definition 3.7 (AT-System) The measures $(\mu_1, \ldots, \mu_r)$ are an *AT-system* on the interval $[a, b]$ if the measures are all absolutely continuous with respect to a positive measure μ on $[a, b]$, i.e., $d\mu_j(x) = w_j(x)\, d\mu(x)$ $(1 \leq j \leq r)$, and for every $\vec{n}$ the functions

$$w_1(x), x w_1(x), \ldots, x^{n_1-1} w_1(x),\ w_2(x), x w_2(x), \ldots, x^{n_2-1} w_2(x),$$
$$\ldots, w_r(x), x w_r(x), \ldots, x^{n_r-1} w_r(x)$$

are a Chebyshev system on $[a, b]$.

For an AT-system we have some control of the zeros of the type I and type II multiple orthogonal polynomials.

Theorem 3.8 *For an AT-system the function*

$$Q_{\vec{n}}(x) = \sum_{j=1}^{r} A_{\vec{n},j}(x) w_j(x)$$

has exactly $|\vec{n}| - 1$ *sign changes on* (a, b)*. Furthermore, the type II multiple orthogonal polynomial* $P_{\vec{n}}$ *has exactly* $|\vec{n}|$ *distinct zeros on* (a, b)*.*

Corollary 3.9 *Every multi-index in an AT-system is normal (an AT-system is perfect).*

A very special system of measures was introduced by Nikishin in 1980.

Definition 3.10 (Nikishin System for $r = 2$) A *Nikishin system* of order $r = 2$ consists of two measures (μ_1, μ_2), both supported on an interval Δ_2, and such that

$$\frac{d\mu_2(x)}{d\mu_1(x)} = \int_{\Delta_1} \frac{d\sigma(t)}{x - t},$$

where σ is a positive measure on an interval Δ_1 and $\Delta_1 \cap \Delta_2 = \emptyset$.

Nikishin showed that indices with $n_1 \geq n_2$ are perfect. Driver and Stahl [12] proved the more general statement.

Theorem 3.11 (Nikishin, Driver-Stahl) *A Nikishin system of order two is perfect.*

In order to define a Nikishin system of order $r > 2$ we need some notation. We write $\langle \sigma_1, \sigma_2 \rangle$ for the measure which is absolutely continuous with respect to σ_1 and for which the Radon-Nikodym derivative is the Stieltjes transform of σ_2:

$$d\langle \sigma_1, \sigma_2 \rangle(x) = \left(\int \frac{d\sigma_2(t)}{x - t} \right) d\sigma_1(x).$$

Nikishin systems of order r can then be defined by induction.

Definition 3.12 (Nikishin System for General r) A *Nikishin system* of order r on an interval Δ_r is a system of r measures $(\mu_1, \mu_2, \ldots, \mu_r)$ supported on Δ_r such that $\mu_j = \langle \mu_1, \sigma_j \rangle$, $2 \leq j \leq r$, where $(\sigma_2, \ldots, \sigma_r)$ is a Nikishin system of order $r - 1$ on an interval Δ_{r-1} and $\Delta_r \cap \Delta_{r-1} = \emptyset$.

Fidalgo Prieto and López Lagomasino proved [13]

Theorem 3.13 *Every Nikishin system is perfect.*

In most cases the measures $(\mu_1, \ldots, \mu_r)$ are absolutely continuous with respect to one fixed measure μ:

$$d\mu_j(x) = w_j(x)\, d\mu(x), \qquad 1 \le j \le r.$$

We then define the *type I function*

$$Q_{\vec{n}}(x) = \sum_{j=1}^{r} A_{\vec{n},j}(x) w_j(x).$$

The type I functions and the type II polynomials then are very complementary: they form a biorthogonal system for many multi-indices.

Property 3.14 (Biorthogonality)

$$\int P_{\vec{n}}(x) Q_{\vec{m}}(x)\, d\mu(x) = \begin{cases} 0, & \text{if } \vec{m} \le \vec{n}, \\ 0, & \text{if } |\vec{n}| \le |\vec{m}| - 2, \\ 1, & \text{if } |\vec{n}| = |\vec{m}| - 1. \end{cases}$$

3.2 Nearest Neighbor Recurrence Relations

The usual orthogonal polynomials (the case $r = 1$) on the real line always satisfy a three-term recurrence relation that expresses $xp_n(x)$ in terms of the polynomials with neighboring degrees p_{n+1}, p_n, p_{n-1}. A similar result is true for multiple orthogonal polynomials, but there are more neighbors for a multi-index. Indeed, the multi-index $\vec{n}$ has r neighbors from above by adding 1 to one of the components of $\vec{n}$. We denote these neighbors from above by $\vec{n} + \vec{e}_k$ for $1 \le k \le r$, where $\vec{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ with 1 in position k. There are also r neighbors from below, namely $\vec{n} - \vec{e}_j$, for $1 \le j \le r$. The nearest neighbor recurrence relations for type II multiple orthogonal polynomials are [45]

$$\begin{aligned} xP_{\vec{n}}(x) &= P_{\vec{n}+\vec{e}_1}(x) + b_{\vec{n},1} P_{\vec{n}}(x) + \sum_{j=1}^{r} a_{\vec{n},j} P_{\vec{n}-\vec{e}_j}(x), \\ &\vdots \\ xP_{\vec{n}}(x) &= P_{\vec{n}+\vec{e}_r}(x) + b_{\vec{n},r} P_{\vec{n}}(x) + \sum_{j=1}^{r} a_{\vec{n},j} P_{\vec{n}-\vec{e}_j}(x). \end{aligned}$$

Observe that one always uses the same linear combination of the neighbors from below. The nearest neighbor recurrence relations for type I multiple orthogonal polynomials are

$$\begin{aligned}
xQ_{\vec{n}}(x) &= Q_{\vec{n}-\vec{e}_1}(x) + b_{\vec{n}-\vec{e}_1,1}Q_{\vec{n}}(x) + \sum_{j=1}^{r} a_{\vec{n},j}Q_{\vec{n}+\vec{e}_j}(x),\\
&\vdots\\
xQ_{\vec{n}}(x) &= Q_{\vec{n}-\vec{e}_r}(x) + b_{\vec{n}-\vec{e}_r,r}Q_{\vec{n}}(x) + \sum_{j=1}^{r} a_{\vec{n},j}Q_{\vec{n}+\vec{e}_j}(x).
\end{aligned}$$

These are using the same recurrence coefficients $a_{\vec{n},j}$, but there is a shift for the recurrence coefficients $b_{\vec{n},k}$. For $r \geq 2$ the recurrence coefficients $\{a_{\vec{n},j}, 1 \leq j \leq r\}$ and $\{b_{\vec{n},k}, 1 \leq k \leq r\}$ are connected:

Theorem 3.15 (Van Assche [45]) *The recurrence coefficients* $(a_{\vec{n},1}, \ldots, a_{\vec{n},r})$ *and* $(b_{\vec{n},1}, \ldots, b_{\vec{n},r})$ *satisfy the partial difference equations*

$$\begin{aligned}
b_{\vec{n}+\vec{e}_i,j} - b_{\vec{n},j} &= b_{\vec{n}+\vec{e}_j,i} - b_{\vec{n},i},\\
\sum_{k=1}^{r} a_{\vec{n}+\vec{e}_j,k} - \sum_{k=1}^{r} a_{\vec{n}+\vec{e}_i,k} &= \det\begin{pmatrix} b_{\vec{n}+\vec{e}_j,i} & b_{\vec{n},i}\\ b_{\vec{n}+\vec{e}_i,j} & b_{\vec{n},j}\end{pmatrix},\\
\frac{a_{\vec{n},i}}{a_{\vec{n}+\vec{e}_j,i}} &= \frac{b_{\vec{n}-\vec{e}_i,j} - b_{\vec{n}-\vec{e}_j,i}}{b_{\vec{n},j} - b_{\vec{n},i}},
\end{aligned}$$

for all $1 \leq i \neq j \leq r$.

By combining the equations of the nearest neighbor recurrence relations, one can also find a recurrence relation of order $r+1$ for the multiple orthogonal polynomials along a path from $\vec{0}$ to $\vec{n}$ in $\mathbb{N}^r$. Let $(\vec{n}_k)_{k\geq 0}$ be a path in $\mathbb{N}^r$ starting from $\vec{n}_0 = \vec{0}$, such that $\vec{n}_{k+1} - \vec{n}_k = \vec{e}_i$ for some $1 \leq i \leq r$. Then

$$xP_{\vec{n}_k}(x) = P_{\vec{n}_{k+1}}(x) + \sum_{j=0}^{r} \beta_{\vec{n}_k,j}P_{\vec{n}_{k-j}}(x).$$

These $\beta_{\vec{n}_k,j}$ coefficients can be expressed in terms of the recurrence coefficients in the nearest neighbor recurrence relations, but the explicit expression is rather complicated for general r. An important case is the *stepline*:

$$\vec{n}_k = (\overbrace{i+1,\ldots,i+1}^{j}, \underbrace{i,\ldots i}_{r-j}), \qquad k = ri + j,\ 0 \leq j \leq r-1.$$

This recurrence relation of order $r+1$ can be expressed in terms of a Hessenberg matrix with r diagonals below the main diagonal:

$$x \begin{pmatrix} P_{\vec{n}_0}(x) \\ P_{\vec{n}_1}(x) \\ P_{\vec{n}_2}(x) \\ \vdots \\ P_{\vec{n}_k}(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} \beta_{\vec{n}_0,0} & 1 & 0 & 0 & 0 & 0 & \cdots \\ \beta_{\vec{n}_1,1} & \beta_{\vec{n}_1,0} & 1 & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & \ddots & 1 & 0 & 0 & \cdots \\ \beta_{\vec{n}_r,r} & \beta_{\vec{n}_r,r-1} & \cdots & \beta_{\vec{n}_r,0} & 1 & 0 & \cdots \\ 0 & \beta_{\vec{n}_{r+1},r} & \beta_{\vec{n}_{r+1},r-1} & \cdots & \beta_{\vec{n}_{r+1},0} & 1 & \cdots \\ 0 & 0 & \beta_{\vec{n}_{r+2},r} & \beta_{\vec{n}_{r+2},r-1} & \cdots & \beta_{\vec{n}_{r+2},0} & \cdots \\ 0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} P_{\vec{n}_0}(x) \\ P_{\vec{n}_1}(x) \\ P_{\vec{n}_2}(x) \\ \vdots \\ P_{\vec{n}_k}(x) \\ \vdots \end{pmatrix}.$$

3.3 Christoffel-Darboux Formula

The Christoffel-Darboux kernel, which is the important reproducing kernel for orthogonal polynomials, has a counterpart in the theory of multiple orthogonal polynomials. It uses both the type I and type II multiple orthogonal polynomials, and is a sum over a path from $\vec{0}$ to $\vec{n}$ as described before. The Christoffel-Darboux kernel is defined as

$$K_{\vec{n}}(x,y) = \sum_{k=0}^{N-1} P_{\vec{n}_k}(x) Q_{\vec{n}_{k+1}}(y)$$

where $\vec{n}_0 = \vec{0}$, $\vec{n}_N = \vec{n}$ and the path in $\mathbb{N}^r$ is such that $\vec{n}_{k+1} - \vec{n}_k = \vec{e}_i$ for some i satisfying $1 \le i \le r$, i.e., in every step the multi-index is increased by 1 in one component. This definition seems to depend on the choice of the path from $\vec{0}$ to $\vec{n}$, but surprisingly this kernel is independent of that chosen path. This is a consequence of the relations between the recurrence coefficients given by Theorem 3.15 and is best explained by the following analogue of the Christoffel-Darboux formula for orthogonal polynomials:

Theorem 3.16 (Daems and Kuijlaars) *Let $(\vec{n}_k)_{0\le k\le N}$ be a path in $\mathbb{N}^r$ starting from $\vec{n}_0 = \vec{0}$ and ending in $\vec{n}_N = \vec{n}$ (where $N = |\vec{n}|$), such that $\vec{n}_{k+1} - \vec{n}_k = \vec{e}_i$ for some $1 \le i \le r$. Then*

$$(x-y)\sum_{k=0}^{N-1} P_{\vec{n}_k}(x) Q_{\vec{n}_{k+1}}(y) = P_{\vec{n}}(x) Q_{\vec{n}}(y) - \sum_{j=1}^{r} a_{\vec{n},j} P_{\vec{n}-\vec{e}_j}(x) Q_{\vec{n}+\vec{e}_j}(y).$$

Proof This was first proved in [9] and a proof based on the nearest neighbor recurrence relations can be found in [45]. □

The sum depends only on the endpoint $\vec{n}$ of the path in $\mathbb{N}^r$ and not on the path from $\vec{0}$ to this point. In many cases this Christoffel-Darboux kernel can be used to generate a determinantal process by using Theorem 2.7 and the biorthogonality in Property 3.14. The only thing which is not obvious is the positivity $K_{\vec{n}}(x,x) \geq 0$, which needs to be checked separately. See [24] for more details about such determinantal processes.

3.4 Hermite-Padé Approximation

Multiple orthogonal polynomials have their roots in Hermite-Padé approximation, which was introduced by Hermite and investigated in detail by Padé (for $r = 1$). Hermite-Padé approximation is a method to approximate r functions simultaneously by rational functions. Multiple orthogonal polynomials appear when one uses Hermite-Padé approximation near infinity. Let $(f_1, \ldots, f_r)$ be r Markov functions, i.e.,

$$f_j(z) = \int \frac{d\mu_j(x)}{z-x} = \sum_{k=0}^{\infty} \frac{m_k^{(j)}}{z^{k+1}}.$$

Definition 3.17 (Type I Hermite-Padé Approximation) Type I Hermite-Padé approximation is to find r polynomials $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$, with $\deg A_{\vec{n},j} \leq n_j - 1$, and a polynomial $B_{\vec{n}}$ such that

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \mathcal{O}\left(\frac{1}{z^{|\vec{n}|}}\right), \qquad z \to \infty. \tag{3.1}$$

The solution is that $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ is the type I multiple orthogonal polynomial vector, and

$$B_{\vec{n}}(z) = \int \sum_{j=1}^{r} \frac{A_{\vec{n},j}(z) - A_{\vec{n},j}(x)}{z-x} \, d\mu_j(x).$$

The error in this approximation problem can also be expressed in terms of the type I multiple orthogonal polynomials. One has

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \int \sum_{j=1}^{r} \frac{A_{\vec{n},j}(x)}{z-x} \, d\mu_j(x),$$

and the orthogonality properties of the type I multiple orthogonal polynomials indeed show that (3.1) holds.

Definition 3.18 (Type II Hermite-Padé Approximation) Type II Hermite-Padé approximation is to find a polynomial $P_{\vec{n}}$ of degree $\leq |\vec{n}|$ and polynomials

$Q_{\vec{n},1}, \ldots, Q_{\vec{n},r}$ such that

$$P_{\vec{n}}(z) f_j(z) - Q_{\vec{n},j}(z) = \mathcal{O}\left(\frac{1}{z^{n_j+1}}\right), \qquad z \to \infty, \tag{3.2}$$

for $1 \leq j \leq r$.

The solution for this approximation problem is to take the type II multiple orthogonal polynomial $P_{\vec{n}}$ and

$$Q_{\vec{n},j}(z) = \int \frac{P_{\vec{n}}(z) - P_{\vec{n}}(x)}{z - x} \, d\mu_j(x).$$

Observe that this approximation problem is to find rational approximants to each f_j with a *common denominator*, and this common denominator turns out to be the type II multiple orthogonal polynomial. The error can again be expressed in terms of the multiple orthogonal polynomial:

$$P_{\vec{n}}(z) f_j(z) - Q_{\vec{n},j}(z) = \int \frac{P_{\vec{n}}(x)}{z - x} \, d\mu_j(x),$$

which can be verified by using the orthogonality conditions for the type II multiple orthogonal polynomial.

Hermite-Padé approximants are used frequently in number theory to find good rational approximants for real numbers and to prove irrationality and transcendence of some important real numbers. Hermite used these approximants (but at 0 rather than ∞) to prove that e is a transcendental number.

3.5 Multiple Hermite Polynomials

As an example we will describe multiple Hermite polynomials in some detail and explain some applications where they are used. The type II multiple Hermite polynomials $H_{\vec{n}}$ satisfy

$$\int_{-\infty}^{\infty} H_{\vec{n}}(x) x^k e^{-x^2 + c_j x} \, dx = 0, \qquad 0 \leq k \leq n_j - 1$$

for $1 \leq j \leq r$, with $c_i \neq c_j$ whenever $i \neq j$. This condition on the parameters $c_1, \ldots, c_r$ guarantees that every multi-index $\vec{n}$ is normal, since the measures with weight function $e^{-x^2 + c_j x}$ $(1 \leq j \leq r)$ form an AT-system. These multiple orthogonal polynomials can be obtained by using the *Rodrigues formula*

$$e^{-x^2} H_{\vec{n}}(x) = \frac{(-1)^{|\vec{n}|}}{2^{|\vec{n}|}} \left(\prod_{j=1}^{r} e^{-c_j x} \frac{d^{n_j}}{dx^{n_j}} e^{c_j x} \right) e^{-x^2}.$$

Exercise Show that the differential operators

$$e^{-c_j x} \frac{d^{n_j}}{dx^{n_j}} e^{c_j x}, \qquad 1 \leq j \leq r$$

are commuting. Use this (and integration by parts) to show that this indeed gives the type II multiple Hermite polynomial.

By using this Rodrigues formula (and the Leibniz rule for the nth derivative of a product), one finds the *explicit expression*

$$H_{\vec{n}}(x) = \frac{(-1)^{|\vec{n}|}}{2^{|\vec{n}|}} \sum_{k_1=0}^{n_1} \cdots \sum_{k_r=0}^{n_r} \binom{n_1}{k_1} \cdots \binom{n_r}{k_r} c_1^{n_1-k_1} \cdots c_r^{n_r-k_r} (-1)^{|\vec{k}|} H_{|\vec{k}|}(x),$$

where H_n are the usual Hermite polynomials. The nearest neighbor recurrence relations for multiple Hermite polynomials are quite simple:

$$x H_{\vec{n}}(x) = H_{\vec{n}+\vec{e}_k}(x) + \frac{c_k}{2} H_{\vec{n}}(x) + \frac{1}{2} \sum_{j=1}^{r} n_j H_{\vec{n}-\vec{e}_j}(x), \qquad 1 \leq k \leq r.$$

They also have some useful differential properties: there are r *raising operators*

$$\left(e^{-x^2+c_j x} H_{\vec{n}-\vec{e}_j}(x) \right)' = -2 e^{-x^2+c_j x} H_{\vec{n}}(x), \qquad 1 \leq j \leq r,$$

and one *lowering operator*

$$H'_{\vec{n}}(x) = \sum_{j=1}^{r} n_j H_{\vec{n}-\vec{e}_j}(x).$$

By combining these raising operators and the lowering operator one finds a *differential equation* of order $r+1$:

$$\left(\prod_{j=1}^{r} D_j \right) D H_{\vec{n}}(x) = -2 \left(\sum_{j=1}^{r} n_j \prod_{i \neq j} D_i \right) H_{\vec{n}}(x),$$

where

$$D = \frac{d}{dx}, \qquad D_j = e^{x^2 - c_j x} D e^{-x^2 + c_j x}.$$

One can also find some integral representations (see [5])

$$H_{\vec{n}}(x) = \frac{1}{\sqrt{\pi} i} \int_{-i\infty}^{i\infty} e^{(s-x)^2} \prod_{j=1}^{r} \left(s - \frac{c_j}{2}\right)^{n_j} ds.$$

For the type I multiple Hermite polynomials one has

$$e^{-x^2 + c_k x} A_{\vec{n},k}(x) = \frac{1}{\sqrt{\pi} 2\pi i} \oint_{\Gamma_k} e^{-(t-x)^2} \prod_{j=1}^{r} \left(t - \frac{c_j}{2}\right)^{-n_j} dt,$$

where Γ_k is a closed contour encircling $c_k/2$ once and none of the other $c_j/2$, and

$$Q_{\vec{n}}(x) = \sum_{k=1}^{r} e^{-x^2 + c_k x} A_{\vec{n},k}(x) = \frac{1}{\sqrt{\pi} 2\pi i} \oint_{\Gamma} e^{-(t-x)^2} \prod_{j=1}^{r} \left(t - \frac{c_j}{2}\right)^{-n_j} dt,$$

where Γ is a closed contour encircling all $c_j/2$.

3.5.1 Random Matrices

These multiple Hermite polynomials are useful for investigating *random matrices with external source* [4]. Let $\mathbf{M}$ be a random $N \times N$ Hermitian matrix and consider the *ensemble* with probability distribution

$$\frac{1}{Z_N} \exp\Big(-\mathrm{Tr}(M^2 - AM)\Big)\, dM, \qquad dM = \prod_{i=1}^{N} dM_{i,i} \prod_{1 \le i < j \le N} dM_{i,j},$$

where A is a fixed Hermitian matrix (the *external source*). The average characteristic polynomial is a multiple Hermite polynomial:

Property 3.19 Suppose A has eigenvalues $c_1, \ldots, c_r$ with multiplicities $n_1, \ldots, n_r$, then

$$\mathbb{E}\Big(\det(\mathbf{M} - zI_N)\Big) = (-1)^{|\vec{n}|} H_{\vec{n}}(z).$$

Furthermore, the eigenvalues form a determinantal process with the Christoffel-Darboux kernel for multiple Hermite polynomials:

Property 3.20 The density of the eigenvalues is given by

$$P_N(\lambda_1, \ldots, \lambda_N) = \frac{1}{N!} \det\Big(K_N(\lambda_i, \lambda_j)\Big)_{i,j=1}^N,$$

where the kernel is given by

$$K_N(x, y) = e^{-(x^2+y^2)/2} \sum_{k=0}^{N-1} H_{\vec{n}_k}(x) Q_{\vec{n}_{k+1}}(y),$$

with $(\vec{n}_k)_{0\leq k\leq N}$ a path from $\vec{0}$ to $\vec{n}$ in $\mathbb{N}^r$ and

$$Q_{\vec{n}}(y) = \sum_{j=1}^r A_{\vec{n},j}(y) e^{c_j y}.$$

This means that we can also find the correlation functions:

Property 3.21 The m-point correlation function

$$\rho_m(\lambda_1, \ldots, \lambda_m) = \frac{N!}{(N-m)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P_N(\lambda_1, \ldots, \lambda_N)\, d\lambda_{m+1} \ldots d\lambda_N$$

is given by

$$\rho_m(\lambda_1, \ldots, \lambda_m) = \det\Big(K_N(\lambda_i, \lambda_j)\Big)_{i,j=1}^m,$$

where the kernel is given by

$$K_N(x, y) = e^{-(x^2+y^2)/2} \sum_{k=0}^{N-1} H_{\vec{n}_k}(x) Q_{\vec{n}_{k+1}}(y).$$

3.5.2 Non-intersecting Brownian Motions

Another interesting problem where multiple Hermite polynomials are appearing is to find what happens with n independent Brownian motions (in fact, n Brownian bridges) with the constraint that they are not allowed to intersect, see [10].

The density of the probability that the n non-intersecting paths, leaving $(t = 0)$ at $a_1, \ldots, a_n$ and arriving $(t = 1)$ at $b_1, \ldots, b_n$, are at $x_1, \ldots, x_n$ at time $t \in (0, 1)$ is (Karlin and McGregor [22])

$$p_{n,t}(x_1, \ldots, x_n) = \frac{1}{Z_n} \det\Big(P(t, a_j, x_k)\Big)_{j,k=1}^n \det\Big(P(1-t, b_j, x_k)\Big)_{j,k=1}^n,$$

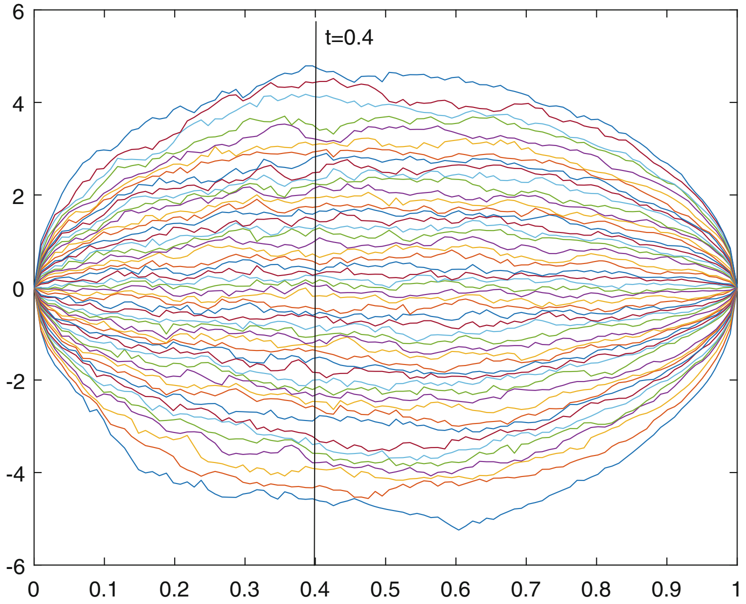


Fig. 1 Non-intersecting Brownian motions

where

$$P(t,a,x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(x-a)^2}.$$

When $a_1,\ldots,a_n \to 0$ and $b_1,\ldots,b_n \to 0$ (see Fig. 1) then

$$p_{n,t}(x_1,\ldots,x_n) = \frac{1}{n!}\det\Big(K_n(x_j,x_k)\Big)_{j,k=1}^n,$$

where the kernel is given by

$$K_n(x,y) = e^{-\frac{x^2}{4t}-\frac{y^2}{4(1-t)}} \sum_{k=0}^{n-1} H_k(\frac{x}{\sqrt{2t}}) H_k(\frac{y}{\sqrt{2(1-t)}}).$$

This kernel is related to the Christoffel-Darboux kernel for the usual Hermite polynomials.

When $a_1,\ldots,a_n \to 0$ and $b_1,\ldots,b_{n/2} \to -b, b_{n/2+1},\ldots,b_n \to b$ (see Fig. 2) then

$$p_{n,t}(x_1,\ldots,x_n) = \frac{1}{n!}\det\Big(K_n(x_j,x_k)\Big)_{j,k=1}^n,$$

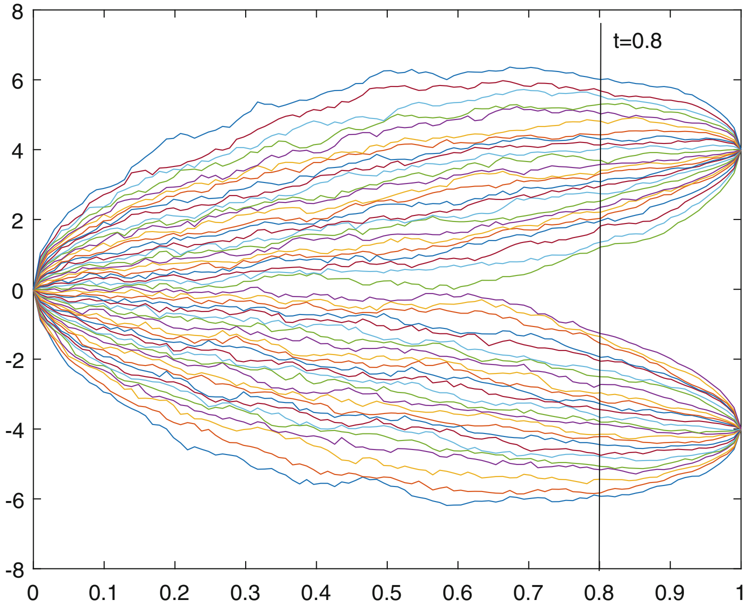


Fig. 2 Non-intersecting Brownian motions (two arriving points)

with

$$K_n(x,y) = e^{-\frac{x^2}{4t}-\frac{y^2}{4(1-t)}} \sum_{k=0}^{n-1} H_{\vec{n}_k}\left(\frac{x}{\sqrt{2t}}\right) Q_{\vec{n}_{k+1}}\left(\frac{y}{\sqrt{2(1-t)}}\right),$$

with multiple orthogonal polynomials for the weights

$$e^{-x^2-2bx}, \quad e^{-x^2+2bx}.$$

This kernel is related to the Christoffel-Darboux kernel for multiple Hermite polynomials. An interesting phenomenon appears: for small values of t the points at level t accumulate on one interval, but for larger values of t in $(0, 1)$ the points accumulate on two disjoint intervals. There is a phase transition at a critical point $t_c \in (0, 1)$. A detailed asymptotic analysis of the kernel near this point will require a special function satisfying a third order differential equation (the Pearsey equation) which is a limiting case of the third order differential equation of multiple Hermite polynomials. The limiting kernel is known as the Pearsey kernel.

3.6 *Multiple Laguerre Polynomials*

The Laguerre weight is

$$w(x) = x^\alpha e^{-x}, \qquad x \in [0,\infty),\ \alpha > -1.$$

There are two easy ways to obtain multiple Laguerre polynomials:

1. Changing the parameter α to $\alpha_1, \ldots, \alpha_r$. This gives *multiple Laguerre polynomials of the first kind.*
2. Changing the exponential decay at infinity from e^{-x} to $e^{-c_j x}$ with parameters $c_1, \ldots, c_r$. This gives *multiple Laguerre polynomials of the second kind.*

3.6.1 Multiple Laguerre Polynomials of the First Kind

Type II multiple Laguerre of the first kind $L_{\vec{n}}^{\vec{\alpha}}(x)$ satisfy

$$\int_0^\infty x^k L_{\vec{n}}^{\vec{\alpha}}(x) x^{\alpha_j} e^{-x}\, dx = 0, \qquad 0 \le k \le n_j - 1,$$

for $1 \le j \le r$. In order that all multi-indices are normal we need to have parameters $\alpha_j > -1$ and $\alpha_i - \alpha_j \notin \mathbb{Z}$ whenever $i \neq j$, in which case the r measures form an AT-system. The multiple orthogonal polynomials can be found from the *Rodrigues formula*

$$(-1)^{|\vec{n}|} e^{-x} L_{\vec{n}}^{\vec{\alpha}}(x) = \prod_{j=1}^r \left(x^{-\alpha_j} \frac{d^{n_j}}{dx^{n_j}} x^{n_j+\alpha_j} \right) e^{-x}.$$

An explicit formula is

$$L_{\vec{n}}^{\vec{\alpha}}(x) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_r=0}^{n_r} (-1)^{|\vec{k}|} \frac{n_1!}{(n_1-k_1)!} \cdots \frac{n_r!}{(n_r-k_r)!}$$
$$\times \binom{n_r+\alpha_r}{k_r} \binom{n_r+n_{r-1}+\alpha_{r-1}-k_r}{k_{r-1}} \cdots \binom{|\vec{n}|-|\vec{k}|+k_1+\alpha_1}{k_1} x^{|\vec{n}|-|\vec{k}|}.$$

Another explicit expression with hypergeometric functions is

$$(-1)^{|\vec{n}|} e^{-x} L_{\vec{n}}^{\vec{\alpha}}(x) = \prod_{j=1}^r (\alpha_j+1)_{n_j}\, {}_rF_r \left(\left. \begin{matrix} n_1+\alpha_1+1, \ldots, n_r+\alpha_r+1 \\ \alpha_1+1, \ldots, \alpha_r+1 \end{matrix} \right| -x \right).$$

The nearest neighbor recurrence relations are

$$xL_{\vec{n}}(x) = L_{\vec{n}+\vec{e}_k}(x) + b_{\vec{n},k}L_{\vec{n}}(x) + \sum_{j=1}^{r} a_{\vec{n},j}L_{\vec{n}-\vec{e}_j}(x)$$

with

$$a_{\vec{n},j} = n_j(n_j+\alpha_j)\prod_{i=1,i\neq j}^{r} \frac{n_j+\alpha_j-\alpha_i}{n_j-n_i+\alpha_j-\alpha_i},$$

and

$$b_{\vec{n},k} = |\vec{n}| + n_k + \alpha_k + 1.$$

These multiple Laguerre polynomials also have some differential properties. There are r *raising operators*

$$\frac{d}{dx}\left(x^{\alpha_j+1}e^{-x}L_{\vec{n}-\vec{e}_j}^{\vec{\alpha}+\vec{e}_j}(x)\right) = -x^{\alpha_j}e^{-x}L_{\vec{n}}^{\vec{\alpha}}(x), \qquad 1 \le j \le r.$$

and there is one *lowering operator*

$$\frac{d}{dx}L_{\vec{n}}^{\vec{\alpha}}(x) = \sum_{j=1}^{r} \frac{\prod_{i=1}^{r}(n_i+\alpha_i-\alpha_j)}{\prod_{i=1,i\neq j}^{r}(\alpha_i-\alpha_j)} L_{\vec{n}-\vec{e}_j}^{\vec{\alpha}+\vec{e}_j}(x).$$

Combining them gives the *differential equation*

$$\left(\prod_{j=1}^{r} D_j\right) DL_{\vec{n}}^{\vec{\alpha}}(x) = -\sum_{j=1}^{r} \frac{\prod_{i=1}^{r}(n_i+\alpha_i-\alpha_j)}{\prod_{i=1,i\neq j}^{r}(\alpha_i-\alpha_j)} \left(\prod_{i\neq j} D_i\right) L_{\vec{n}}^{\vec{\alpha}}(x).$$

$$D = \frac{d}{dx}, \qquad D_j = x^{-\alpha_j}e^x D x^{\alpha_j+1}e^{-x}.$$

3.6.2 Multiple Laguerre Polynomials of the Second Kind

Type II multiple Laguerre polynomials of the second kind $L_{\vec{n}}^{\alpha,\vec{c}}(x)$ satisfy

$$\int_0^\infty x^k L_{\vec{n}}^{\alpha,\vec{c}}(x) x^\alpha e^{-c_j x}\, dx = 0, \qquad 0 \le k \le n_j - 1,$$

for $1 \le j \le r$. The parameters need to satisfy $\alpha > -1$ and $c_j > 0$ with $c_i \ne c_j$ whenever $i \ne j$. The *Rodrigues formula* is

$$(-1)^{|\vec{n}|} \prod_{j=1}^r c_j^{n_j}\, x^\alpha L_{\vec{n}}^{\alpha,\vec{c}}(x) = \prod_{j=1}^r \left(e^{c_j x} \frac{d^{n_j}}{dx^{n_j}} e^{-c_j x} \right) x^{|\vec{n}|+\alpha},$$

which allows to find the explicit expression

$$L_{\vec{n}}^{\alpha,\vec{c}}(x) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_r=0}^{n_r} \binom{n_1}{k_1} \cdots \binom{n_r}{k_r} \binom{|\vec{n}|+\alpha}{|\vec{k}|} (-1)^{|\vec{k}|} \frac{|\vec{k}|!}{c_1^{k_1} \cdots c_r^{k_r}} x^{|\vec{n}|-|\vec{k}|}.$$

The nearest neighbor recurrence relations are

$$x L_{\vec{n}}(x) = L_{\vec{n}+\vec{e}_k}(x) + b_{\vec{n},k} L_{\vec{n}}(x) + \sum_{j=1}^r a_{\vec{n},j} L_{\vec{n}-\vec{e}_j}(x),$$

with

$$a_{\vec{n},j} = \frac{n_j(|\vec{n}|+\alpha)}{c_j^2}, \quad b_{\vec{n},k} = \frac{|\vec{n}|+\alpha+1}{c_k} + \sum_{j=1}^r \frac{n_j}{c_j}.$$

The differential properties include r *raising operators*

$$\frac{d}{dx}\left(x^{\alpha+1} e^{-c_j x} L_{\vec{n}-\vec{e}_j}^{\alpha+1,\vec{c}}(x) \right) = -c_j x^\alpha e^{-c_j x} L_{\vec{n}}^{\alpha,\vec{c}}(x), \qquad 1 \le j \le r.$$

and one *lowering operator*

$$\frac{d}{dx} L_{\vec{n}}^{\alpha,\vec{c}}(x) = \sum_{j=1}^r n_j L_{\vec{n}-\vec{e}_j}^{\alpha+1,\vec{c}}(x).$$

They give the *differential equation*

$$\left(\prod_{j=1}^r D_j \right) x^{\alpha+1} D L_{\vec{n}}^{\alpha,\vec{c}}(x) = -\sum_{j=1}^r c_j n_j \left(\prod_{i \ne j} D_i \right) x^\alpha L_{\vec{n}}^{\alpha,\vec{c}}(x),$$

where

$$D = \frac{d}{dx}, \quad D_j = e^{c_j x} D e^{-c_j x}.$$

3.6.3 Random Matrices: Wishart Ensemble

Wishart (1928) introduced the *Wishart distribution* for $N \times N$ positive definite Hermitian matrices

$$\mathbf{M} = \mathbf{X}\mathbf{X}^*, \qquad \mathbf{X} \in \mathbb{C}^{N\times(N+p)},$$

where all the columns of $\mathbf{X}$ are independent and have a multivariate Gauss distribution with covariance matrix Σ. The density for the Wishart distribution is

$$\frac{1}{Z_N} e^{-\mathrm{Tr}(\Sigma^{-1}M)} (\det M)^p \, dM.$$

If $\Sigma = I_N$ then Laguerre polynomials (with $\alpha = p$) play an important role. If Σ^{-1} has eigenvalues $c_1, \ldots, c_r$ with multiplicities $n_1, \ldots, n_r$, then we need multiple Laguerre polynomials of the second kind. The average characteristic polynomial is

$$\mathbb{E}\Big(\det(\mathbf{M} - zI_N)\Big) = (-1)^{|\vec{n}|} L_{\vec{n}}^{p,\vec{c}}(z).$$

3.7 *Jacobi-Piñeiro Polynomials*

There are several ways to find multiple Jacobi polynomials. Here we only mention one way which uses the same differential operators as the multiple Laguerre polynomials of the first kind. The Jacobi-Piñeiro polynomials $P_{\vec{n}}^{(\vec{\alpha},\beta)}$ satisfy

$$\int_0^1 P_{\vec{n}}^{(\vec{\alpha},\beta)}(x) x^k x^{\alpha_j} (1-x)^\beta \, dx = 0, \qquad 0 \le k \le n_j - 1,$$

for $1 \le j \le r$. Hence we are using Jacobi weights $x^\alpha(1-x)^\beta$ on the interval $[0, 1]$, with $\alpha, \beta > -1$ but with r different parameters $\alpha_1, \ldots, \alpha_r$. In order to have a perfect system we require $\alpha_i - \alpha_j \notin \mathbb{Z}$ whenever $i \neq j$. They can be obtained using the *Rodrigues formula*

$$(-1)^{|\vec{n}|} \left(\prod_{j=1}^r (|\vec{n}| + \alpha_j + \beta)_{n_j} \right) (1-x)^\beta P_{\vec{n}}^{(\vec{\alpha},\beta)}(x) = \prod_{j=1}^r \left(x^{-\alpha_j} \frac{d^{n_j}}{dx^{n_j}} x^{n_j+\alpha_j} \right) (1-x)^{\beta+|\vec{n}|}.$$

An expression in terms of generalized hypergeometric functions is

$$(-1)^{|\vec{n}|}\left(\prod_{j=1}^{r}(|\vec{n}|+\alpha_j+\beta)_{n_j}\right)(1-x)^{\beta}P_{\vec{n}}^{(\vec{\alpha},\beta)}(x)$$
$$=\prod_{j=1}^{r}(\alpha_j+1)_{n_j}\;{}_{r+1}F_r\left(\left.\begin{matrix}-|\vec{n}|-\beta,\alpha_1+n_1+1,\ldots,\alpha_r+n_r+1\\ \alpha_1+1,\ldots,\alpha_r+1\end{matrix}\right|x\right).$$

This hypergeometric function does not terminate when β is not an integer. Another useful expression is

$$(-1)^{|\vec{n}|}P_{\vec{n}}^{(\vec{\alpha},\beta)}(x)$$
$$=\frac{n_1!\cdots n_r!}{\prod_{j=1}^{r}(|\vec{n}|+\alpha_j+\beta)_{n_j}}\sum_{k_1=0}^{n_1}\cdots\sum_{k_r=0}^{n_r}(-1)^{|\vec{k}|}\prod_{j=1}^{r}\binom{n_j+\alpha_j+\sum_{i=1}^{j-1}k_i}{n_j-k_j}$$
$$\times\binom{|\vec{n}|+\beta}{|\vec{k}|}\frac{|\vec{k}|!x^{|\vec{k}|}(1-x)^{|\vec{n}|-|\vec{k}|}}{k_1!\cdots k_r!}.$$

Again there are r raising differential operators and one lowering operator and the recurrence coefficients are known explicitly. These polynomials are useful for rational approximation of polylogarithms, and in particular for the zeta function $\zeta(k)$ at integers. The polylogarithms are defined by

$$\mathrm{Li}_k(z)=\sum_{n=1}^{\infty}\frac{z^n}{n^k},\qquad |z|<1,$$

and one has

$$\mathrm{Li}_{k+1}(1/z)=\frac{(-1)^k}{k!}\int_0^1\frac{\log^k(x)}{z-x}\,dx.$$

Simultaneous rational approximation to $\mathrm{Li}_1(1/z),\ldots,\mathrm{Li}_r(1/z)$ can be done using Hermite-Padé approximation with a limiting case of Jacobi-Piñeiro polynomials where $\beta=0$ and $\alpha_1=\alpha_2=\cdots=\alpha_r=0$, which is possible when $n_1\geq n_2\geq\cdots\geq n_r$. This is particularly interesting if we let $z\to 1$, since $\mathrm{Li}_k(1)=\zeta(k)$. Apéry's construction of good rational approximants for $\zeta(3)$ (proving that $\zeta(3)$ is irrational) essentially makes use of these multiple orthogonal polynomials, see, e.g. [43].

4 Orthogonal Polynomials and Painlevé Equations

In this section we describe how orthogonal polynomials are related to non-linear difference and differential equations, in particular to discrete Painlevé equations and the six Painlevé differential equations. For a recent discussion on this relation between orthogonal polynomials and Painlevé equations we refer to the monograph [46]. Other useful references are [7, 8, 44].

Painlevé equations (discrete and continuous) appear at various places in the theory of orthogonal polynomials, in particular

- The recurrence coefficients of some semiclassical orthogonal polynomials satisfy discrete Painlevé equations.
- The recurrence coefficients of orthogonal polynomials with a Toda-type evolution satisfy Painlevé differential equations for which special solutions depending on special functions (Airy, Bessel, (confluent) hypergeometric, parabolic cylinder functions) are relevant.
- Rational solutions of Painlevé equations can be expressed in terms of Wronskians of orthogonal polynomials.
- The local asymptotics for orthogonal polynomials at critical points is often using special transcendental solutions of Painlevé equations.

In this section we will only deal with the first two of these.

What are Painlevé (differential) equations? They are second order nonlinear differential equations

$$y'' = R(y', y, x), \qquad R \text{ rational},$$

that have the *Painlevé property*: **The general solution is free from movable branch points**. The only singularities which may depend on the initial conditions are poles. Painlevé and his collaborators found 50 families (up to Möbius transformations), all of which could be reduced to known equations and six new equations (new at least at the beginning of the twentieth century). The six Painlevé equations are

$$\mathrm{P_I} \quad y'' = 6y^2 + x,$$

$$\mathrm{P_{II}} \quad y'' = 2y^3 + xy + \alpha, \tag{4.1}$$

$$\mathrm{P_{III}} \quad y'' = \frac{(y')^2}{y} - \frac{y'}{x} + \frac{\alpha y^2 + \beta}{x} + \gamma y^3 + \frac{\delta}{y}, \tag{4.2}$$

$$\mathrm{P_{IV}} \quad y'' = \frac{(y')^2}{2y} + \frac{3}{2}y^3 + 4xy^2 + 2(x^2 - \alpha)y + \frac{\beta}{y}, \tag{4.3}$$

$$\mathrm{P_V} \quad y'' = \left(\frac{1}{2y} + \frac{1}{y-1}\right)(y')^2 - \frac{y'}{x} + \frac{(y-1)^2}{x^2}\left(\alpha y + \frac{\beta}{y}\right) + \frac{\gamma y}{x} + \frac{\delta y(y+1)}{y-1}, \tag{4.4}$$

$$\mathrm{P_{VI}} \quad y'' = \frac{1}{2}\left(\frac{1}{y} + \frac{1}{y-1} + \frac{1}{y-x}\right)(y')^2 - \left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{y-x}\right)y' + \frac{y(y-1)(y-x)}{x^2(x-1)^2}\left(\alpha + \frac{\beta x}{y^2} + \frac{\gamma(x-1)}{(y-1)^2} + \frac{\delta x(x-1)}{(y-x)^2}\right),$$

Discrete Painlevé equations are somewhat more difficult to describe. Roughly speaking they are second order nonlinear recurrence equations for which the continuous limit is a Painlevé equation. They have the *singularity confinement* property, but this property is not sufficient to characterize discrete Painlevé equations. A quote by Kruskal [23] is:

> Anything simpler becomes trivially integrable, anything more complicated becomes hopelessly non-integrable.

A more correct description is that they are nonlinear recurrence relations with 'nice' symmetry and geometry. A full classification of discrete (and continuous) Painlevé equations has been found by Sakai [36]. This is based on *rational surfaces* associated with affine root systems. It describes the space of initial values which parametrizes all the solutions (Okamoto [34]). A fine tuning of this classification was given recently by Kajiwara, Noumi and Yamada [21]: they also include the *symmetry*, i.e., the group of Bäcklund transformations, which are transformations that map a solution of a Painlevé equation to another solution with different parameters. A partial list of discrete Painlevé equations is:

$$\text{d-P}_{\mathrm{I}} \quad x_{n+1} + x_n + x_{n-1} = \frac{z_n + a(-1)^n}{x_n} + b, \tag{4.5}$$

$$\text{d-P}_{\mathrm{II}} \quad x_{n+1} + x_{n-1} = \frac{x_n z_n + a}{1 - x_n^2}, \tag{4.6}$$

$$\text{d-P}_{\mathrm{IV}} \quad (x_{n+1} + x_n)(x_n + x_{n-1}) = \frac{(x_n^2 - a^2)(x_n^2 - b^2)}{(x_n + z_n)^2 - c^2},$$

$$\text{d-P}_{\mathrm{V}} \quad \frac{(x_{n+1} + x_n - z_{n+1} - z_n)(x_n + x_{n-1} - z_n - z_{n-1})}{(x_{n+1} + x_n)(x_n + x_{n-1})} = \frac{[(x_n - z_n)^2 - a^2][(x_n - z_n)^2 - b^2]}{(x_n - c^2)(x_n - d^2)},$$

where $z_n = \alpha n + \beta$ and a, b, c, d are constants.

$$\text{q-P}_{\mathrm{III}} \quad x_{n+1}x_{n-1} = \frac{(x_n - aq_n)(x_n - bq_n)}{(1 - cx_n)(1 - x_n/c)},$$

$$\text{q-P}_{\mathrm{V}} \quad (x_{n+1}x_n - 1)(x_n x_{n-1} - 1) = \frac{(x_n - a)(x_n - 1/a)(x_n - b)(x_n - 1/b)}{(1 - cx_n q_n)(1 - x_n q_n/c)},$$

$$\text{q-P}_{\text{VI}} \quad \frac{(x_n x_{n+1} - q_n q_{n+1})(x_n x_{n-1} - q_n q_{n-1})}{(x_n x_{n+1} - 1)(x_n x_{n-1} - 1)} = \frac{(x_n - a q_n)(x_n - q_n/a)(x_n - b q_n)(x_n - q_n/b)}{(x_n - c)(x_n - 1/c)(x_n - d)(x_n - 1/d)},$$

where $q_n = q_0 q^n$ and a, b, c, d are constants.

$$\alpha\text{-d-P}_{\text{IV}} \quad (x_n + y_n)(x_{n+1} + y_n) = \frac{(y_n - a)(y_n - b)(y_n - c)(y_n - d)}{(y_n + \gamma - z_n)(y_n - \gamma - z_n)}$$

$$(x_n + y_n)(x_n + y_{n-1}) = \frac{(x_n + a)(x_n + b)(x_n + c)(x_n + d)}{(x_n + \delta - z_{n+1/2})(x_n - \delta - z_{n+1/2})}.$$

The latter corresponds to d-P($E_6^{(1)}/A_2^{(1)}$) where $E_6^{(1)}$ is the surface type and $A_2^{(1)}$ is the symmetry type. Sakai's classification (surface type) corresponds to the following diagram:

$$\begin{array}{ccccccccccccccccc}
E_8^e & & & & & & & & & & & & & & & & A_1^q \\
\downarrow & & & & & & & & & & & & & & & \nearrow & \\
E_8^q & \to & E_7^q & \to & E_6^q & \to & D_5^q & \to & A_4^q & \to & (A_2+A_1)^q & & \to & (A_1+A_1)^q & & \to & A_1^q \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & | \quad \downarrow & & & | \quad \downarrow & & & \\
E_8^d & \to & E_7^d & \to & E_6^d & \to & D_4^c & \to & A_3^c & \to & | \quad (2A_1)^c & & \to & | \quad A_1^c & & & \\
& & & & & & & & & \searrow & \downarrow & & \searrow & \downarrow & & & \\
& & & & & & & & & & A_2^c & & \to & A_1^c & & &
\end{array}$$

4.1 Compatibility and Lax Pairs

There is a general philosophy behind the reason why Painlevé equations appear for the recurrence coefficients of orthogonal polynomials. Orthogonal polynomials $P_n(x)$ are really functions of two variables: a discrete variable n and a continuous variable x. The three term recurrence relation (1.2) gives a difference equation in the variable n, and if the measure is absolutely continuous with a weight function w that satisfies a Pearson equation

$$\frac{d}{dx}[\sigma(x)w(x)] = \tau(x)w(x), \tag{4.7}$$

where σ and τ are polynomials, then the orthogonal polynomials also satisfy differential relations in the variable x. If $\deg \sigma \leq 2$ and $\deg \tau = 1$ then we are dealing with classical orthogonal polynomials which satisfy the second order differential equation

$$\sigma(x)y''(x) + \tau(x)y'(x) + \lambda_n y(x) = 0,$$

where $\lambda_n = -n(n-1)\sigma''/2 - n\tau'$. In the semiclassical case we still have the Pearson equation (4.7) but we allow $\deg\sigma > 2$ or $\deg\tau \neq 1$. In that case there is a *structure relation*

$$\sigma(x)\frac{d}{dx}P_n(x) = \sum_{k=n-t}^{n+s-1} A_{n,k}P_k(x), \tag{4.8}$$

where $s = \deg\sigma$ and $t = \max\{\deg\tau, \deg\sigma - 1\}$. The structure relation (4.8) and the three-term recurrence relation (1.2) have to be *compatible*: if we differentiate the terms in the recurrence relation (1.2) and replace all the $P_k'(x)$ using the structure relation (4.8), then we get a linear combination of a finite number of orthogonal polynomials that is equal to 0. Since (orthogonal) polynomials are linearly independent in the linear space of polynomials, the coefficients in this linear combination have to be zero, and this gives relations between the recurrence coefficient a_n^2, b_n and the coefficients $A_{n,k}$ in the structure relation. Eliminating these $A_{n,k}$ gives recurrence relations for the a_n^2, b_n, which turn out to be non-linear. If they are of second order, then we can identify them as discrete Painlevé equations. In this way the three-term recurrence relation and the structure relation can be considered as a *Lax pair* for the obtained discrete Painlevé equation.

In order to get to the Painlevé differential equation, we need to introduce an extra continuous parameter t. For this we will use an exponential modification of the measure μ and investigate orthogonal polynomials for the measure $d\mu_t(x) = e^{xt}\,d\mu(x)$, whenever all the moments of this modified measure exist. We will denote the monic orthogonal polynomials by $P_n(x;t)$ and in this way the orthogonal polynomial is now a function of three variables n, x, t. The behavior for the parameter t is given by:

Theorem 4.1 *The monic orthogonal polynomials $P_n(x;t)$ for the measure $d\mu_t(x) = e^{xt}\,d\mu(x)$ satisfy*

$$\frac{d}{dt}P_n(x;t) = C_n(t)P_{n-1}(x;t), \tag{4.9}$$

where $C_n(t)$ depends only on t and n.

Proof First of all, since $P_n(x;t)$ is a monic polynomial, the derivative $\frac{d}{dt}P_n(x;t)$ is a polynomial of degree $\leq n-1$. We will show that it is orthogonal to x^k for $0 \leq k \leq n-2$ for the measure $e^{xt}\,d\mu(x)$, so that it is proportional to $P_{n-1}(x;t)$, which proves (4.9). We start from the orthogonality relations

$$\int P_n(x;t)x^k e^{xt}\,d\mu(x) = 0, \qquad 0 \leq k \leq n-1,$$

and take derivatives with respect to t to find

$$\int \left(\frac{d}{dt}P_n(x;t)\right)x^k e^{xt}\,d\mu(x) + \int P_n(x;t)x^{k+1}e^{xt}\,d\mu(x) = 0, \qquad 0 \leq k \leq n-1.$$

The second integral vanishes for $0 \le k \le n-2$ by orthogonality, hence

$$\int \left(\frac{d}{dt} P_n(x;t)\right) x^k e^{xt}\, d\mu(x) = 0, \qquad 0 \le k \le n-2,$$

which is what we needed to prove. □

This relation is not new, see e.g. [39, §4], but has not been sufficiently appreciated in the literature. If we now check the compatibility between (4.9) and the three-term recurrence relation (1.2), then we find differential-difference equations for the recurrence coefficients a_n^2, b_n.

Theorem 4.2 (Toda Equations) *The recurrence coefficients $a_n^2(t)$ and $b_n(t)$ for the orthogonal polynomials $P_n(x;t)$ satisfy*

$$\frac{d}{dt} a_n^2(t) = a_n^2(b_n - b_{n-1}), \qquad n \ge 1, \tag{4.10}$$

$$\frac{d}{dt} b_n(t) = a_{n+1}^2 - a_n^2, \qquad n \ge 0, \tag{4.11}$$

with $a_0^2 = 0$.

Proof If we take derivatives with respect to t in the three-term recurrence relations (1.2), then

$$x\frac{d}{dt}P_n(x;t) = \frac{d}{dt}P_n(x;t) + b_n'(t)P_n(x;t) + b_n\frac{d}{dt}P_n(x) \\ + (a_n^2)'(t)P_{n-1}(x;t) + a_n^2\frac{d}{dt}P_{n-1}(x;t).$$

Use (4.9) to find

$$xC_nP_{n-1}(x;t) = C_{n+1}P_n(x;t) + b_n'P_n(x;t) + b_nC_nP_{n-1}(x;t) \\ + (a_n^2)'P_{n-1}(x;t) + a_n^2C_{n-1}P_{n-2}(x;t).$$

If we compare this with (1.2) (with n shifted to $n-1$), then we find

$$C_{n+1} + b_n' = C_n, \tag{4.12}$$

$$C_n(b_{n-1} - b_n) = (a_n^2)' \tag{4.13}$$

$$a_n^2 C_{n-1} = a_{n-1}^2 C_n. \tag{4.14}$$

From (4.14) we find that a_n^2/C_n does not depend on n, so that $a_n^2/C_n = a_1^2/C_1$ and from (4.12) we find that $C_1 = -b_0'(t)$. A simple exercise shows that $b_0'(t) = a_1^2(t)$

so that $C_n(t) = -a_n^2(t)$. If we use this in (4.13), then we find (4.10). If we use it in (4.12), then we find (4.11). $\square$

The system (4.10)–(4.11) is closely related to a chain of interacting particles with exponential interaction with their neighbors, introduced by Toda [41] in 1967. If $x_n(t)$ is the position of particle n, then the Toda system of equations is

$$x_n''(t) = \exp(x_{n-1} - x_n) - \exp(x_n - x_{n+1}).$$

The relation with orthogonal polynomials was made by Flaschka [15, 16] and Manakov [28], who suggested the change of variables

$$a_n(t) = \exp(-[x_n - x_{n-1}]/2), \quad b_n = -x_n'(t),$$

which gives the system (4.10)–(4.11).

If we are dealing with symmetric orthogonal polynomials, i.e., when the measure is symmetric and all the odd moments are zero, then the three-term recurrence relation simplifies to

$$x P_n(x) = P_{n+1}(x) + a_n^2 P_{n-1}(x), \qquad n \geq 0. \tag{4.15}$$

A symmetric modification of the measure is given by $d\mu_t(x) = e^{tx^2}\, d\mu(x)$ and the relation becomes

$$\frac{d}{dt} P_n(x;t) = C_n(t) P_{n-2}(x;t). \tag{4.16}$$

The compatibility between (4.15) and (4.16) then gives:

Theorem 4.3 (Langmuir Lattice) *Let μ be a symmetric positive measure on $\mathbb{R}$ for which all the moments exist and let μ_t be the measure for which $d\mu_t(x) = e^{tx^2}\, d\mu(x)$, where $t \in \mathbb{R}$ is such that all the moments of μ_t exist. Then the recurrence coefficients of the orthogonal polynomials for μ_t satisfy the differential-difference equations*

$$\frac{d}{dt} a_n^2 = a_n^2 (a_{n+1}^2 - a_{n-1}^2), \qquad n \geq 1. \tag{4.17}$$

Proof If we differentiate (4.15) with respect to t and then use (4.16), then we find

$$x C_n P_{n-2}(x;t) = C_{n+1} P_{n-1}(x;t) + (a_n^2)' P_{n-1}(x;t) + a_n^2 C_{n-1} P_{n-3}(x;t).$$

Comparing with (4.15) (with n replaced by $n-2$) gives

$$(a_n^2)' = C_n - C_{n+1}, \tag{4.18}$$

$$a_n^2 C_{n-1} = a_{n-2}^2 C_n. \tag{4.19}$$

From (4.19) it follows that $a_n^2 a_{n-1}^2/C_n$ is constant and therefore equal to $a_2^2 a_1^2/C_2$. Now $C_2(t) = -(a_1^2)'$ and one can easily compute a_1^2, a_2^2 and $(a_1^2)'$ in terms of the moments m_0, m_2, m_4 to find that $a_2^2 a_1^2/C_2 = -1$, so that $a_n^2 a_{n-1}^2 = -C_n$. If one uses this in (4.18), then one finds (4.17). □

This differential-difference equation is known as the Langmuir lattice or the Kac-van Moerbeke lattice. We will now illustrate this with a number of explicit examples.

4.2 Discrete Painlevé I

Let us consider orthogonal polynomials for the weight function $w(x) = e^{-x^4+tx^2}$ on $(-\infty, \infty)$. The symmetry $w(-x) = w(x)$ of this weight function implies that the recurrence coefficients b_n in (1.1) or (1.2) vanish and the three-term recurrence relation is (4.15). The orthogonal polynomials also have a nice differential property: the *structure relation* is

$$P_n'(x) = A_n P_{n-1}(x) + C_n P_{n-3}(x), \tag{4.20}$$

for certain sequences $(A_n)_n$ and $(C_n)_n$. Indeed, we can express P_n' in terms of the orthogonal polynomials as

$$P_n'(x) = \sum_{k=0}^{n-1} c_{n,k} P_k(x),$$

where

$$c_{n,k} \int_{-\infty}^{\infty} P_k^2(x) e^{-x^4+tx^2}\, dx = \int_{-\infty}^{\infty} P_n'(x) P_k(x) e^{-x^4+tx^2}\, dx.$$

Using integration by parts gives

$$\begin{aligned} c_{n,k}/\gamma_k^2 &= -\int_{-\infty}^{\infty} P_n(x) \big(P_k(x) e^{-x^4+tx^2}\big)'\, dx \\ &= -\int_{-\infty}^{\infty} P_n(x) P_k'(x) e^{-x^4+tx^2}\, dx \\ &\quad + \int_{-\infty}^{\infty} P_n(x) P_k(x)(4x^3 - 2tx) e^{-x^4+tx^2}\, dx, \end{aligned}$$

and the last two integrals are zero for $0 \le k < n-3$ by orthogonality, so that only $c_{n,n-1}$, $c_{n,n-2}$ and $c_{n,n-3}$ are left. The symmetry of w implies that $P_{2n}(x)$ is an even polynomial and $P_{2n+1}(x)$ is an odd polynomial for every n, hence $c_{n,n-2} = 0$. Taking $A_n = c_{n,n-1}$ and $C_n = c_{n,n-3}$ then gives the structure relation.

We now have a recurrence relation (4.15) which describes the behavior of $P_n(x)$ in the (discrete) variable n, and a structure relation (4.20) which describes the behavior of $P_n(x)$ in the (continuous) variable x. Both relations have to be compatible: if we differentiate (4.15) and then use (4.20) to replace all the derivatives, then comparing coefficients of the polynomials p_k gives the *compatibility relations*

$$4a_n^2\left(a_{n+1}^2 + a_n^2 + a_{n-1}^2 - \frac{t}{2}\right) = n. \tag{4.21}$$

This simple non-linear recurrence relation is known as discrete Painlevé I (d-$\mathrm{P_I}$) and is a special case of (4.5) we gave earlier. This particular equation was already in work of Shohat [37] in 1939, who extended earlier work of Laguerre [25] from 1885. Later it was obtained again by Freud [18] in 1976, who was unaware of the work of Shohat. The special positive solution needed to get the recurrence coefficients was analyzed by Nevai [32] and Lew and Quarles [26]. An asymptotic expansion was found by Máté-Nevai-Zaslavsky [30]. Only later (in 1991) it was recognized as a discrete Painlevé equation by Fokas, Its and Kitaev [17] who coined the name d-$\mathrm{P_I}$. Magnus [27] used the extra parameter t and showed that, as a function of t, the recurrence coefficient $a_n(t)$ satisfies the differential equation Painlevé IV, as we will see later.

The discrete Painlevé equation (4.21) easily allows to find the asymptotic behavior as $n \to \infty$:

Theorem 4.4 (Freud) *The recurrence coefficients for the weight* $w(x) = e^{-x^4+tx^2}$ *on* $(-\infty, \infty)$ *satisfy*

$$\lim_{n\to\infty} \frac{a_n}{n^{1/4}} = \frac{1}{\sqrt[4]{12}}.$$

Observe that (4.21) is a second order recurrence relation, so one needs two initial conditions a_0 and a_1 to generate all the recurrence coefficients. It turns out that the recurrence coefficients are a special solution with $a_0 = 0$ for which all a_n are positive for $n \geq 1$. This means that there is only one special initial value a_1 that gives a positive solution. Put $x_n = a_n^2$, then (for $t = 0$)

$$x_n(x_{n+1} + x_n + x_{n-1}) = an, \qquad a = 1/4. \tag{4.22}$$

Theorem 4.5 (Lew and Quarles, Nevai) *There is a unique solution of* (4.22) *for which* $x_0 = 0$ *and* $x_n > 0$ *for all* $n \geq 1$.

Hence one should not use this recurrence relation (4.22) to generate the recurrence coefficients starting from $x_0 = 0$ and x_1, because a small error in x_1 will produce a sequence for which not all the terms are positive. A small perturbation in the initial condition x_1 has a very important effect on the solution as $n \to \infty$. This is not unusual for non-linear recurrence relations. Instead it is better to generate the positive solution by using a fixed point algorithm, because the positive solution turns out to be the fixed point of a contraction in an appropriate normed space of infinite sequences. See, e.g., [46, §2.3].

4.3 Langmuir Lattice and Painlevé IV

We will modify the measure μ by multiplying it with the symmetric function e^{tx^2}, where t is a real parameter. This gives the Langmuir lattice (4.17). We can combine this with the discrete Painlevé equation (4.21) to find a differential equation for $a_n^2(t)$ as a function of the variable t. Put $a_n^2 = x_n$, then

$$n = 4x_n(x_{n+1} + x_n + x_{n-1} - t/2), \tag{4.23}$$

$$x_n' = x_n(x_{n+1} - x_{n-1}), \tag{4.24}$$

where the $'$ denotes the derivative with respect to t. Differentiate (4.24) to find

$$x_n'' = x_n'(x_{n+1} - x_{n-1}) + x_n(x_{n+1}' - x_{n-1}').$$

Replace x_{n+1}' and x_{n-1}' by (4.24), then

$$x_n'' = x_n'(x_{n+1} - x_{n-1}) + x_n\Big(x_{n+1}(x_{n+2} - x_n) - x_{n-1}(x_n - x_{n-2})\Big).$$

Eliminate x_{n+1} and x_{n-1} using (4.23)–(4.24) to find

$$x_n'' = \frac{(x_n')^2}{2x_n} + \frac{3x_n^3}{2} - tx_n^2 + x_n\left(\frac{n}{2} + \frac{t^2}{8}\right) - \frac{n^2}{32x_n}.$$

This is Painlevé IV if we use the transformation $2x_n(t) = y(-t/2)$. This means that Painlevé IV has a solution which can be described completely in terms of the moments of $w(x) = e^{-x^4+tx^2}$, since $a_n^2 = \gamma_{n-1}^2/\gamma_n^2$ and by (1.5) $\gamma_n^2 = D_n/D_{n+1}$, where D_n is the Hankel determinant (1.3) containing the moments. Notice that all the odd moments m_{2n+1} are zero, and for the even moments one has

$$m_{2n} = \int_{\mathbb{R}} x^{2n} e^{-x^4+tx^2}\,dx = \frac{d^n}{dt^n} m_0.$$

Hence the special solution $a_n^2(t)$ of Painlevé IV is in terms of $m_0(t)$ only, and this is a special function:

$$m_0(t) = \int_{-\infty}^{\infty} e^{-x^4+tx^2}\,dx = 2^{-1/4}\sqrt{\pi}e^{t^2/8}D_{-1/2}(-\sqrt{t/2}),$$

where $D_{-1/2}$ is a parabolic cylinder function.

4.4 Singularity Confinement

In this section we will explain the notion of singularity confinement for the discrete Painlevé I equation

$$4x_n(x_{n+1}+x_n+x_{n-1})=n.$$

From this equation one finds

$$x_{n+1}=\frac{n}{4x_n}-x_n-x_{n-1}.$$

If $x_n=0$ then x_{n+1} becomes infinite. This need not be a problem, but problems arise later when we have to add or subtract infinities. So we need to be careful and suppose that $x_n=\epsilon$ is small. Then

$$x_{n+1}=\frac{n}{4\epsilon}-\epsilon-x_{n-1},$$

and

$$x_{n+2}=-\frac{n}{4\epsilon}+x_{n-1}+\epsilon+\mathcal{O}(\epsilon^2),$$

and

$$x_{n+3}=-\epsilon+\mathcal{O}(\epsilon^2),$$

and one more

$$x_{n+4}=x_{n-1}+\frac{2-8x_{n-1}^2}{n}\epsilon+\mathcal{O}(\epsilon^2),$$

and for $\epsilon\to 0$ we see that x_{n+4} is finite again and recovers the value x_{n-1} we had before we started to get singularities. The singularities are confined to x_{n+1} and x_{n+2} and one can continue the recurrence relation from x_{n+4}. This has some meaning in terms of the orthogonal polynomials for the weight e^{-x^4}, but we have to consider this weight on the set $\mathbb{R}\cup i\mathbb{R}$ and look for orthogonal polynomials $(R_n)_n$ for which

$$\alpha\int_{-\infty}^{\infty}R_n(x)R_m(x)e^{-x^4}\,dx+\beta\int_{-i\infty}^{+i\infty}R_n(x)R_m(x)e^{-x^4}\,|dx|=0,\quad n\neq m,$$

with $\alpha,\beta>0$. They satisfy the recurrence relation

$$xR_n(x)=R_{n+1}(x)+c_nR_{n-1}(x)$$

and the recurrence coefficients $(c_n)_n$ still satisfy (4.22) but with initial condition $c_0 = 0$ and $c_1 = \frac{(\alpha-\beta)m_2}{(\alpha+\beta)m_0}$. If $\alpha = \beta$ then $c_1 = 0$ generates a singularity for d-P$_\mathrm{I}$ and gives $c_2 = \infty$, hence R_3 does not exist if we define it using (1.4). The singularity, however, is confined to a finite number of terms. We have

Property 4.6 For $\alpha = \beta$ one has $D_{4n-1} = D_{4n-2} = 0$ for the Hankel determinants, so that R_{4n-1} and R_{4n-2} as defined by (1.4) do not exist for $n \geq 1$. Furthermore

$$R_{4n}(x) = r_n(x^4), \qquad R_{4n+1}(x) = x s_n(x^4).$$

The polynomials r_n and s_n can be identified as Laguerre polynomials with parameter $\alpha = -3/4$ and $\alpha = 1/4$ respectively. The problem with R_{4n-1} and R_{4n-2} is not so much that they do not exist, but rather that they are not unique.

Exercise Show that for every $a \in \mathbb{R}$ the polynomials $(x^2 + ax)s_n(x^4)$ are monic polynomials of degree $4n + 2$ that are orthogonal to x^k for $0 \leq k \leq 4n + 1$, so that the monic orthogonal polynomial R_{4n+2} is not unique. In a similar way $(x^3 + ax^2 + bx)s_n(x^4)$ are monic polynomials of degree $4n + 3$ that are orthogonal to x^k for $0 \leq k \leq 4n + 2$ for every $a, b \in \mathbb{R}$ so that the monic orthogonal polynomial R_{4n+3} is not unique.

4.5 Generalized Charlier Polynomials

Our next example is a family of discrete orthogonal polynomials $P_n(x)$, which satisfy

$$\sum_{k=0}^{\infty} P_n(k) P_m(k) \frac{c^k}{(\beta)_k k!} = 0, \qquad n \neq m.$$

Without the factor $(\beta)_k$ the polynomials are the Charlier polynomials, but with the factor $(\beta)_k$ we have a semiclassical family of discrete orthogonal polynomials. The case $\beta = 1$ was investigated in [47] and the general case in [38], see also [46, §3.2]. The structure relation for discrete orthogonal polynomials is now in terms of a difference operator instead of a differential operator. For these generalized Charlier polynomials it is

$$\Delta P_n(x) = A_n P_{n-1}(x) + B_n P_{n-2}(x), \tag{4.25}$$

where Δ is the forward difference operator acting on a function f by

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

and $(A_n)_n$ and $(B_n)_n$ are certain sequences. If one works out the compatibility of (1.2) and (4.25), then one finds

$$b_n + b_{n-1} - n + \beta = \frac{cn}{a_n^2},$$
$$(a_{n+1}^2 - c)(a_n^2 - c) = c(b_n - n)(b_n - n + \beta - 1).$$

This corresponds to a limiting case of discrete Painlevé with surface/symmetry $D_4^{(1)}$ in Sakai's classification.

If we put $c = c_0 e^t$, then the weights with parameter c are a Toda modification of the weights with parameter c_0,

$$\frac{c^k}{(\beta)_k k!} = e^{tk} \frac{c_0^k}{(\beta)_k k!},$$

and hence the recurrence coefficients satisfy the Toda equations given in Theorem 4.2. Put $x_n(t) = a_n^2$ and $y_n(t) = b_n$, then

$$(x_n - c)(x_{n+1} - c) = c(y_n - n)(y_n - n + \beta - 1),$$
$$y_n + y_{n-1} - n + \beta = \frac{cn}{x_n},$$

and if $x_n' = dx_n/dc$, $y_n' = dy_n/dc$, the Toda lattice equations are

$$cx_n' = x_n(y_n - y_{n-1}),$$
$$cy_n' = x_{n+1} - x_n.$$

Eliminate y_{n-1} and x_{n+1} (this requires quite a few computations) and put $x_n = \frac{c}{1-y}$, then $y(c)$ satisfies (after even more computations)

$$y'' = \frac{1}{2}\left(\frac{1}{2y} + \frac{1}{y-1}\right)(y')^2 - \frac{y'}{c} + \frac{(1-y)^2}{c^2}\left(\frac{n^2 y}{2} - \frac{(\beta-1)^2}{2y}\right) - \frac{2y}{c}.$$

This is a Painlevé V differential equation as in (4.4) with $\delta = 0$. Such an equation can always be transformed to Painlevé III.

4.6 Discrete Painlevé II

We will now give an example of a family of orthogonal polynomials on the unit circle, for which the recurrence coefficients satisfy a discrete Painlevé equation.

Orthogonal polynomials on the unit circle (OPUC) are defined by the orthogonality relations

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi_n(z)\overline{\varphi_m(z)}v(\theta)\,d\theta = \delta_{m,n}, \qquad z = e^{i\theta}, \quad \varphi_n(z) = \kappa_n z^n + \cdots$$

where $\kappa_n > 0$. We denote the monic polynomials by $\Phi_n = \varphi_n/\kappa_n$. They satisfy a nice recurrence relation

$$z\Phi_n(z) = \Phi_{n+1}(z) + \overline{\alpha_n}\Phi_n^*(z), \tag{4.26}$$

where $\Phi_n^*(z) = z^n\overline{\Phi}_n(1/z)$ is the reversed polynomial. The recurrence coefficients $\alpha_n = -\overline{\Phi_{n+1}(0)}$ are nowadays known as *Verblunsky coefficients*, but earlier they were also known as Schur parameters or reflection coefficients. Let $v(\theta) = e^{t\cos\theta}$ for $\theta \in [-\pi, \pi]$. The trigonometric moments for this weight function are modified Bessel functions

$$\frac{1}{2\pi}\int_0^{2\pi} e^{in\theta}v(\theta)\,d\theta = I_n(t),$$

which is why Ismail [20, Example 8.4.3] calls them *modified Bessel polynomials*. The symmetry $v(-\theta) = v(\theta)$ implies that $\alpha_n(t)$ are real-valued. If we write

$$v(\theta) = \hat{v}(z), \quad z = e^{i\theta},$$

then

$$\hat{v}(z) = \exp\left(t\frac{z+\frac{1}{z}}{2}\right),$$

and this function satisfies the Pearson equation

$$\hat{v}'(z) = \frac{t}{2}\left(1 - \frac{1}{z^2}\right)\hat{v}(z).$$

As a consequence the orthogonal polynomials satisfy a structure relation:

Property 4.7 The monic orthogonal polynomials for $v(\theta) = e^{t\cos\theta}$ satisfy

$$\Phi_n'(z) = n\Phi_{n-1}(z) + B_n\Phi_{n-2}(z), \tag{4.27}$$

for some sequence $(B_n)_n$. In fact, one has

$$B_n = \frac{t}{2}\frac{\kappa_{n-2}^2}{\kappa_n^2}.$$

We now have two equations: the recurrence relation (4.26) and the structure relation (4.27), and we can check their compatibility. They will be compatible if the recurrence coefficients satisfy the following non-linear relation:

Theorem 4.8 (Periwal and Shevitz [35]) *The Verblunsky coefficients for the weight $v(\theta) = e^{t\cos\theta}$ satisfy*

$$-\frac{t}{2}(\alpha_{n+1} + \alpha_{n-1}) = \frac{(n+1)\alpha_n}{1-\alpha_n^2},$$

with initial values

$$\alpha_{-1} = -1, \quad \alpha_0 = \frac{I_1(t)}{I_0(t)}.$$

Let $x_n = \alpha_{n-1}$, then

$$x_{n+1} + x_{n-1} = \frac{\alpha n x_n}{1-x_n^2}, \qquad \alpha = -\frac{2}{t}, \tag{4.28}$$

and this is a particular case of discrete Painlevé II (d-$\mathrm{P_{II}}$) given in (4.6). We need a solution with $x_0 = -1$ and $|x_n| < 1$ for $n \geq 1$, because for Verblunsky coefficients one always has $|\alpha_n| < 1$. Such a solution is unique.

Theorem 4.9 *Suppose $\alpha > 0$. Then there is a unique solution of (4.28) for which $x_0 = -1$ and $-1 < x_n < 1$. The solution corresponds to $x_1 = I_1(-2/\alpha)/I_0(-2/\alpha)$ and is negative for every $n \geq 0$.*

A proof of this result can be found in [46, §3.3] for $\alpha > 1$; a proof for $0 < \alpha \leq 1$ has not been published and we invite the reader to come up with such a proof. This special solution converges to zero (fast).

4.7 The Ablowitz-Ladik Lattice and Painlevé III

The lattice equations corresponding to orthogonal polynomials on the unit circle are the *Ablowitz-Ladik lattice* equations (or the Schur flow).

Theorem 4.10 *Let ν be a positive measure on the unit circle which is symmetric (the Verblunsky coefficients are real). Let ν_t be the modified measure $d\nu_t(\theta) = e^{t\cos\theta}\, d\nu(\theta)$, with $t \in \mathbb{R}$. The Verblunsky coefficients $(\alpha_n(t))_n$ for the measure ν_t then satisfy*

$$2\alpha_n' = (1-\alpha_n^2)(\alpha_{n+1} - \alpha_{n-1}), \qquad n \geq 0.$$

We can now combine the discrete Painlevé II equation

$$\alpha_{n+1} + \alpha_{n-1} = \frac{-2n\alpha_n}{t(1-\alpha_n^2)}$$

with the Ablowitz-Ladik equation

$$\alpha_{n+1} - \alpha_{n-1} = \frac{2\alpha_n'}{1-\alpha_n^2}.$$

Eliminate α_{n+1} and α_{n-1} to find

$$\alpha_n'' = -\frac{\alpha_n}{1-\alpha_n^2}(\alpha_n')^2 - \frac{\alpha_n'}{t} - \alpha_n(1-\alpha_n^2) + \frac{(n+1)^2}{t^2}\frac{\alpha_n}{1-\alpha_n^2}.$$

Exercise If one puts $\alpha_n = \frac{1+y}{1-y}$, then show that y satisfies the Painlevé V differential equation (4.4) with $\gamma = 0$.

Painlevé V with $\gamma = 0$ can always be transformed to Painlevé III. A direct approach was given by Hisakado [19] and Tracy and Widom [42]. They showed that the ratio $w_n(t) = \alpha_n(t)/\alpha_{n-1}(t)$ satisfies Painlevé III.

4.8 *Some More Examples*

Several more examples have been worked out in the literature the past few years. Here is a short sample.

4.8.1 Generalized Meixner Polynomials

These are discrete orthogonal polynomials

$$\sum_{k=0}^{\infty} P_n(k)P_m(k)\frac{(\gamma)_k a^k}{(\beta)_k k!} = 0, \qquad n \neq m,$$

which were considered in [8, 14, 38]. Put $a_n^2 = na - (\gamma-1)u_n$, and $b_n = n + \gamma - \beta + a - \frac{\gamma-1}{a}v_n$, then

$$(u_n+v_n)(u_{n+1}+v_n) = \frac{\gamma-1}{a^2}v_n(v_n-a)\left(v_n - a\frac{\gamma-\beta}{\gamma-1}\right),$$

$$(u_n+v_n)(u_n+v_{n-1}) = \frac{u_n}{u_n - \frac{an}{\gamma-1}}(u_n+a)\left(u_n + a\frac{\gamma-\beta}{\gamma-1}\right).$$

The initial values are

$$a_0^2 = 0, \quad b_0 = \frac{\gamma a}{\beta} \frac{M(\gamma+1, \beta+1, a)}{M(\gamma, \beta, a)},$$

where $M(a, b, z)$ is Kummer's confluent hypergeometric function. This is asymmetric discrete Painlevé IV or d-P($E_6^{(1)}/A_2^{(1)}$). If we put

$$v_n(a) = \frac{a\Big(ay' - (1+\beta-2\gamma)y^2 + (n+1-a+\beta-2\gamma)y - n\Big)}{2(\gamma-1)(y-1)y},$$

then

$$y'' = \left(\frac{1}{2y} + \frac{1}{y-1}\right)(y')^2 - \frac{y'}{a} + \frac{(y-1)^2}{a^2}\big(Ay + \frac{B}{y}\big) + \frac{Cy}{a} + \frac{Dy(y+1)}{y-1}$$

with

$$A = \frac{(\beta-1)^2}{2}, \quad B = -\frac{n^2}{2}, \quad C = n - \beta + 2\gamma, \quad D = -\frac{1}{2},$$

which is Painlevé V given in (4.4).

4.8.2 Modified Laguerre Polynomials

Chen and Its [6] (see also [46, §4.4]) looked at orthogonal polynomials for the weight function $w(x) = x^\alpha e^{-x} e^{-t/x}$ on $[0, \infty)$. This is a modification of the Laguerre weight with an exponential function that has an essential singularity at 0. Put $b_n = 2n + \alpha + 1 + c_n$, $a_n^2 = n(n+\alpha) + y_n + \sum_{j=0}^{n-1} c_j$, and $c_n = 1/x_n$, then

$$\begin{aligned} x_n + x_{n-1} &= \frac{nt - (2n+\alpha)y_n}{y_n(y_n - t)}, \\ y_n + y_{n+1} &= t - \frac{2n+\alpha+1}{x_n} - \frac{1}{x_n^2}. \end{aligned}$$

This corresponds to the discrete Painlevé equation d-P($(2A_1)^{(1)}/D_6^{(1)}$). The exponential modification is not of Toda type but belongs to a similar class of modifications (the Toda hierarchy). With some effort one can find the differential equation

$$c_n'' = \frac{(c_n')^2}{c_n} - \frac{c_n'}{t} + (2n+\alpha+1)\frac{c_n^2}{t^2} + \frac{c_n^3}{t^2} + \frac{\alpha}{t} - \frac{1}{c_n}$$

which is Painlevé III given in (4.2).

4.8.3 Modified Jacobi Polynomials

Basor et al. [3] (see also [46, §5.2]) considered the weight $w(x) = (1-x)^\alpha(1+x)^\beta e^{-tx}$. This is a Toda modification of the weight function for Jacobi polynomials. In this case one has

$$
\begin{aligned}
tb_n &= 2n+1+\alpha+\beta-t-2R_n,\\
t(t+R_n)a_n^2 &= n(n+\beta)-(2n+\alpha+\beta)r_n-\frac{tr_n(r_n+\alpha)}{R_n},
\end{aligned}
$$

where r_n and R_n satisfy the recurrence relations

$$
\begin{aligned}
2t(r_{n+1}+r_n) &= 4R_n^2-2R_n(2n+1+\alpha+\beta-t)-2\alpha t,\\
n(n+\beta)-(2n+\alpha+\beta)r_n &= r_n(r_n+\alpha)\left(\frac{t^2}{R_nR_{n-1}}+\frac{t}{R_n}+\frac{t}{R_{n-1}}\right),
\end{aligned}
$$

and for $y = 1+t/R_n$ one has the differential equation

$$
\begin{aligned}
y'' = \frac{3y-1}{2y(y-1)}(y')^2-\frac{y'}{t}+2(2n+1+\alpha+\beta)\frac{y}{t}-\frac{2y(y+1)}{y-1}\\
+\frac{(y-1)^2}{t^2}\left(\frac{\alpha^2y}{2}-\frac{\beta^2}{2y}\right),
\end{aligned}
$$

which is Painlevé V given in (4.4).

4.8.4 q-Orthogonal Polynomials

There are also examples of families of q-orthogonal polynomials for which one can find q-discrete Painlevé equations for the recurrence coefficients. In this case the structure relation uses the q-difference operator D_q for which

$$
D_qf(x) = \frac{f(x)-f(qx)}{x(1-q)}.
$$

If we consider the weight

$$
w(x) = \frac{x^\alpha}{(-x^2;q^2)_\infty(-q^2/x^2;q^2)_\infty}, \qquad x \in [0,\infty)
$$

then the recurrence coefficients (after some transformation) satisfy q-discrete Painlevé III

$$x_{n-1}x_{n+1} = \frac{(x_n + q^{-\alpha})^2}{(q^{n+\alpha}x_n + 1)^2}.$$

For the weight

$$w(x) = \frac{x^\alpha(-p/x^2; q^2)_\infty}{(-x^2; q^2)_\infty(-q^2/x^2; q^2)_\infty}, \qquad x \in [0, \infty)$$

one finds q-discrete Painlevé V

$$(z_n z_{n-1} - 1)(z_n z_{n+1} - 1) = \frac{(z_n + \sqrt{q^{2-\alpha}/p})^2(z_n\sqrt{pq^{\alpha-2}})^2}{(q^{n+\alpha/2-1}\sqrt{p}z_n + 1)^2}.$$

and for

$$w(x) = x^\alpha(q^2x^2; q^2)_\infty, \qquad x \in \{q^k, k = 0, 1, 2, 3, \ldots\}$$

one again finds q-discrete Painlevé V. Observe that sometimes the weights are on $[0, \infty)$ but they can also be on the discrete set $\{q^n, n \in \mathbb{N}\}$. See [46, §5.4] for more details.

4.9 Wronskians and Special Function Solutions

There is a good explanation why these Toda modifications of orthogonal polynomials often give rise to Painlevé differential equations. In fact the solutions that we need for the recurrence coefficients are special solutions of the Painlevé equations in terms of special functions, such as the Airy functions, the Bessel functions, parabolic cylinder functions, the confluent hypergeometric function and the hypergeometric function. Such special function solutions are often in terms of Wronskians of one of these special functions. We can easily explain where these Wronskians are coming from, by using the theory of orthogonal polynomials. Indeed, we return to our Hankel determinants D_n given in (1.3). They contain the moments m_n, which for a Toda modification are

$$m_n(t) = \int_{\mathbb{R}} x^n e^{xt}\, d\mu(x) = \frac{d^n}{dt^n}\int_{\mathbb{R}} e^{xt}\, \mu(x) = \frac{d^n}{dx^n}m_0(t).$$

Hence all the moments are obtained from the moment $m_0(t)$ by differentiation, and the Hankel determinant (1.3) becomes

$$D_n = \det \begin{pmatrix} m_0 & m_0' & m_0'' & \cdots & m_0^{(n-1)} \\ m_0' & m_0'' & m_0''' & \cdots & m_0^{(n)} \\ m_0'' & m_0''' & m_0^{(4)} & \cdots & m_0^{(n+1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_0^{(n-1)} & m_0^{(n)} & m_0^{(n+1)} & \cdots & m_0^{(2n-2)} \end{pmatrix}$$

which is the Wronskian of the functions $m_0, m_0', m_0'', \ldots, m_0^{(n-1)}$,

$$D_n = \mathrm{Wr}(m_0, m_0', m_0'', \ldots, m_0^{(n-1)}).$$

The recurrence coefficient a_n^2 can be expressed in terms of these Hankel determinants as

$$a_n^2(t) = \frac{\gamma_{n-1}^2}{\gamma_n^2} = \frac{D_{n+1}(t)D_{n-1}(t)}{D_n^2(t)},$$

where we used (1.5). The recurrence coefficients b_n can also be found in terms of determinants. If we write $P_n(x) = x^n + \delta_n x^{n-1} + \cdots$ and compare the coefficients of x^n in the recurrence relation (1.2), then $b_n = \delta_n - \delta_{n+1}$. The coefficient δ_n can be obtained from (1.4) from which we see that $\delta_n = -D_n^*/D_n$, where D_n^* is obtained from D_n by replacing the last column $(m_{n-1}, m_n, \ldots, m_{(2n-2)})^T$ by moments of one order higher $(m_n, m_{n+1}, \ldots, m_{2n-1})^T$. If we take a derivative of the Wronskian, then

$$\frac{d}{dt} D_n = \mathrm{Wr}(m_0, m_0', m_0'', \ldots, m_0^{(n-2)}, m_0^{(n)}) = D_n^*,$$

so that

$$b_n(t) = \frac{D_{n+1}'(t)}{D_{n+1}(t)} - \frac{D_n'(t)}{D_n(t)}.$$

This gives explicit expressions of the recurrence coefficients $a_n^2(t)$ and $b_n(t)$ in terms of Wronskians generated from one seed function $m_0(t)$.

Acknowledgements Many thanks to Mama Foupouagnigni and Wolfram Koepf for organizing the workshop *Introduction to Orthogonal Polynomials and Applications* in Douala, Cameroon, and for encouraging me to write this survey. Also thanks to Arno Kuijlaars with whom I am sharing a course on *Orthogonal Polynomials and Random Matrices* at KU Leuven, which was very useful for the material in Sect. 2. The author was supported by FWO research project G.0864.16N and by EOS project 30889451.

References

1. G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, *Cambridge Studies in Advanced Mathematics*, vol. 118 (Cambridge University Press, Cambridge, 2010)
2. A.I. Aptekarev, Multiple orthogonal polynomials. J. Comput. Appl. Math. **99**(1–2), 423–447 (1998)
3. E. Basor, Y. Chen, T. Ehrhardt, Painlevé V and time dependent Jacobi polynomials. J. Phys. A Math. Theor. **43**(1), 015204 (25 pp.) (2010)
4. P.M. Bleher, A.B.J. Kuijlaars, Random matrices with external source and multiple orthogonal polynomials. Int. Math. Res. Not. **2004**(3), 109–129 (2004)
5. P.M. Bleher, A.B.J. Kuijlaars, Integral representations for multiple Hermite and multiple Laguerre polynomials. Ann. Inst. Fourier, Grenoble **55**(6), 2001–2014 (2005)
6. Y. Chen, A. Its, Painlevé III and a singular linear statistics in Hermitian random matrix ensembles. J. Approx. Theory **162**(2), 270–297 (2010)
7. P.A. Clarkson, Painlevé equations—nonlinear special functions, in Lecture Notes in Mathematics, vol. 1883 (Springer, Berlin, 2006), pp. 331–411
8. P.A. Clarkson, Recurrence coefficients for discrete orthogonal polynomials and the Painlevé equations. J. Phys. A Math. Theor. **46**(18), 185205 (18 pp.) (2013)
9. E. Daems, A.B.J. Kuijlaars, A Christoffel-Darboux formula for multiple orthogonal polynomials. J. Approx. Theory **130**(2), 190–202 (2004)
10. E. Daems, A.B.J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions. J. Approx. Theory **146**(1), 91–114 (2007)
11. P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*. Courant Lecture Notes in Mathematics, vol. 3 (American Mathematical Society and Courant Institute of Mathematical Sciences, New York, 2000)
12. K. Driver, H. Stahl, Simultaneous rational approximants to Nikishin-systems. I, II. Acta Sci. Math. (Szeged) **60**(1–2), 245–263 (1995); **61**(1–4), 261–284 (1995)
13. U. Fidalgo Prieto, G. López Lagomasino, Nikishin systems are perfect. Constr. Approx. **34**, 297–356 (2011)
14. G. Filipuk, W. Van Assche, Recurrence coefficients of a new generalization of the Meixner polynomials. SIGMA **7**, 068 (2011)
15. H. Flaschka, The Toda lattice. I. Existence of integrals. Phys. Rev. B (3) **9**, 1924–1925 (1974)
16. H. Flaschka, On the Toda lattice. II. Inverse-scattering solution. Progr. Theoret. Phys. **51**, 703–716 (1974)
17. A.S. Fokas, A.R. Its, A.V. Kitaev, Discrete Painlevé equations and their appearance in quantum gravity. Commun. Math. Phys. **142**(2), 313–344 (1991)
18. G. Freud, On the coefficients in the recursion formulae of orthogonal polynomials. Proc. Royal Irish Acad. Sect. A **76**(1), 1–6 (1976)
19. M. Hisakado, Unitary matrix models and Painlevé III. Mod. Phys. Lett. **A11**, 3001–3010 (1996)
20. M.E.H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, Encyclopedia of Mathematics and its Applications, vol. 98 (Cambridge University Press, Cambridge, 2005)
21. K. Kajiwara, M. Noumi, Y. Yamada, Geometric aspects of Painlevé equations. J. Phys. A Math. Theor. **50**(7), 073001 (164 pp.) (2017)
22. S. Karlin, J. McGregor, Coincidence probabilities. Pacific J. Math. **9**, 1141–1164 (1959)
23. M.D. Kruskal, B. Grammaticos, T. Tamizhmani, Three lessons on the Painlevé property and the Painlevé equations. Springer Lecture Notes in Physics **644**, 1–15 (2004)
24. A.B.J. Kuijlaars, Multiple orthogonal polynomial ensembles, in *Recent Trends in Orthogonal Polynomials and Approximation Theory*, Contemporary Mathematics, vol. 507 (American Mathematical Society, Providence, 2010), pp. 155–176
25. E. Laguerre, Sur la réduction en fractions continues d'une fraction qui satisfait à une équation différentielle linéaire du premier ordre dont les coefficients sont rationnels. J. Math. Pures Appl. (4) **1**, 135–166 (1995)

26. J.S. Lew, D.A. Quarles, Nonnegative solutions of a nonlinear recurrence. J. Approx. Theory **38**, 357–379 (1983)
27. A.P. Magnus, Painlevé type differential equations for the recurrence coefficients of semi-classical orthogonal polynomials. J. Comput. Appl. Math. **57**, 215–237 (1995)
28. S.V. Manakov, Complete integrability and stochastization of discrete dynamical systems. Sov. Phys. – JETP **40**(2), 269–274 (1974); Translated from Zh. Eksper. Teoret. Fiz. **67**(2), 543–555 (1974) (in Russian)
29. A. Martínez-Finkelshtein, W. Van Assche, What is ... a multiple orthogonal polynomial? Not. Amer. Math. Soc. **63**(9), 1029–1031 (2016)
30. A. Máté, P. Nevai, T. Zaslavsky, Asymptotic expansions of ratios of coefficients of orthogonal polynomials with exponential weights. Trans. Am. Math. Soc. **287**(2), 495–505 (1985)
31. M.L. Mehta, *Random Matrices*, 2nd edn. (Academic, San Diego, 1991)
32. P. Nevai, Orthogonal polynomials associated with $\exp(-x^4)$, in *Second Edmonton Conference on Approximation Theory*, Canadian Mathematical Society Conference Proceedings, vol. 3 (1983), pp. 263–285
33. E.M. Nikishin, V.N. Sorokin, *Rational Approximations and Orthogonality*, Translations of Mathematical Monographs, vol. 92 (American Mathematical Society, Providence, 1991)
34. K. Okamoto, Sur les feuilletages associés aux équations du second ordre à points critiques fixés de P. Painlevé. Japan. J. Math. **5**, 1–79 (1979)
35. V. Periwal, D. Shevitz, Unitary-matrix models as exactly solvable string theories. Phys. Rev. Letters **64**, 1326–1329 (1990)
36. H. Sakai, Rational surfaces associated with affine root systems and geometry of the Painlevé equations. Comm. Math. Phys. **220**(1), 165–229 (2001)
37. J. Shohat, A differential equation for orthogonal polynomials. Duke Math. J. **5**(2), 401–417 (1939)
38. C. Smet, W. Van Assche, Orthogonal polynomials on a bi-lattice. Constr. Approx. **36**(2), 215–242 (2012)
39. V. Spiridonov, A. Zhedanov, Self-similarity, spectral transformations and orthogonal and biorthogonal polynomials, in *Self-Similar Systems*, Dubna, ed. by V.B. Priezzhev, V.P. Spiridonov (1999), pp. 349–361
40. G. Szegő, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, Providence (1939; 4th edn. 1975)
41. M. Toda, Vibration of a chain with a non-linear interaction. J. Phys. Soc. Jpn. **22**, 431–436 (1967)
42. C.A. Tracy, H. Widom, Random unitary matrices, permutations and Painlevé. Commun. Math. Phys. **207**(3), 665–685 (1999)
43. W. Van Assche, Multiple orthogonal polynomials, irrationality and transcendence, in *Continued Fractions: from Analytic Number Theory to Constructive Approximation* (Columbia, MO, 1998), Contemporary Mathematics, vol. 236 (American Mathematical Society, Providence, 1999), pp. 325–342
44. W. Van Assche, Discrete Painlevé equations for recurrence coefficients of orthogonal polynomials, in *Difference Equations, Special Functions and Orthogonal Polynomials* (World Scientific, Singapore, 2007), pp. 687–725
45. W. Van Assche, Nearest neighbor recurrence relations for multiple orthogonal polynomials. J. Approx. Theory **163**, 1427–1448 (2011)
46. W. Van Assche, *Orthogonal Polynomials and Painlevé Equations*, Lecture Series of the Australian Mathematical Society, vol. 27 (Cambridge University Press, Cambridge, 2018)
47. W. Van Assche, M. Foupouagnigni, Analysis of non-linear recurrence relations for the recurrence coefficients of generalized Charlier polynomials. J. Nonlinear Math. Phys. **10**(suppl. 2), 231–237 (2003)
48. W. Van Assche, J.S. Geronimo, A.B.J. Kuijlaars, Riemann-Hilbert problems for multiple orthogonal polynomials, in *Special Functions 2000: Current Perspective and Future Directions*, ed. by J. Bustoz, M.E.H. Ismail, S.K. Suslov, NATO Science Series II. Mathematics, Physics and Chemistry, vol. 30 (Kluwer, Dordrecht, 2001), pp. 23–59