



Stable Spectral Methods for Time-Dependent Problems and the Preservation of Structure

Arieh Iserles¹

Received: 3 April 2023 / Revised: 5 October 2023 / Accepted: 5 October 2023

© The Author(s) 2024

Abstract

This paper is concerned with orthonormal systems in real intervals, given with zero Dirichlet boundary conditions. More specifically, our interest is in systems with a skew-symmetric differentiation matrix (this excludes orthonormal polynomials). We consider a simple construction of such systems and pursue its ramifications. In general, given any $C^1(a, b)$ weight function such that $w(a) = w(b) = 0$, we can generate an orthonormal system with a skew-symmetric differentiation matrix. Except for the case $a = -\infty$, $b = +\infty$, only few powers of that matrix are bounded and we establish a connection between properties of the weight function and boundedness. In particular, we examine in detail two weight functions: the Laguerre weight function $x^\alpha e^{-x}$ for $x > 0$ and $\alpha > 0$ and the ultraspherical weight function $(1 - x^2)^\alpha$, $x \in (-1, 1)$, $\alpha > 0$, and establish their properties. Both weights share a most welcome feature of *separability*, which allows for fast computation. The quality of approximation is highly sensitive to the choice of α , and we discuss how to choose optimally this parameter, depending on the number of zero boundary conditions.

Keywords Orthogonal polynomials · Differentiation matrices · Special functions · Spectral methods

Mathematics Subject Classification Primary 41A30 · Secondary 41A10 · 65M70

Communicated by Endre Süli.

✉ Arieh Iserles
ai@maths.cam.ac.uk

¹ Department of Applied Mathematics and Theoretical Physics, University of Cambridge Centre for Mathematical Sciences, Wilberforce Rd, Cambridge CB3 0WA, UK

Contents

- 1 Introduction and Motivation
 - 1.1 Spectral Methods
 - 1.2 The Differentiation Matrix
 - 1.3 Few Examples
 - 1.4 An Orthonormal Basis
 - 1.5 Plan of this Paper
- 2 W-Functions
 - 2.1 The Definition and Some of Its Consequences
 - 2.2 The Boundedness of \mathcal{D}^s
- 3 Separable Systems
 - 3.1 The Laguerre Family
 - 3.2 The Ultraspherical Family
 - 3.3 Counterexamples: generalised Hermite and Konoplev weights
 - 3.3.1 Generalised Hermite weights
 - 3.3.2 Konoplev weights
 - 3.3.3 A Limiting Behaviour of the $t_{m,n}^s$
- 4 Computational Aspects
 - 4.1 A Product of \mathcal{D} by a Vector
 - 4.2 The Speed of Convergence
 - 4.2.1 Ultraspherical W-Functions
 - 4.2.2 Laguerre W-Functions
 - 4.2.3 The Speed of Convergence Redux
 - 4.3 Outstanding Computational and Theoretical Challenges
- A Separability Coefficients for Laguerre Weights
- B Symmetric Separability Coefficients for Ultraspherical Weights
- References

1 Introduction and Motivation

1.1 Spectral Methods

This work is motivated by spectral methods for time-dependent partial differential equations (PDEs) of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f(x, u), \quad t \geq 0, \quad x \in \Omega, \tag{1.1}$$

where \mathcal{L} is a well-posed linear operator, defining a strongly continuous semigroup, and $\Omega \subseteq \mathbb{R}^d$, given with an initial condition for $u(x, 0)$ and appropriate boundary conditions on $\partial\Omega$. Standard examples are $\mathcal{L} = \Delta$ with $f \equiv 0$ (the *diffusion* equation) or f a cubic polynomial in u with real zeros (the *Fitzhugh–Nagumo* equation) and $\mathcal{L} = i\Delta$ with either $f = -iV(x)$ (the *linear Schrödinger* equation) or $f = -i\lambda|u|^2u$ (the *nonlinear Schrödinger* equation with standard cubic nonlinearity).

In this paper, we are concerned by spectral methods applied in tandem with a splitting approach. As an example, we may commence by approximating locally the solution of (1.1) using the Strang splitting,

$$u^{n+1} = e^{\frac{1}{2}\Delta t \mathcal{L}} e^{\Delta t f} e^{\frac{1}{2}\Delta t \mathcal{L}} u^n, \quad n \geq 0, \tag{1.2}$$

where $u^n(x)$ is an approximation to $u(x, n\Delta t)$. Here $e^{t\mathcal{L}}v$ is a shorthand for a numerical solution at t of $\partial u/\partial t = \mathcal{L}u$, $u(0) = v$, while $e^{tf}v$ denotes a numerical solution of the ordinary differential equation (ODE) $du/dt = f(x, u)$, $u(0) = v$. The splitting (1.2), which incurs a local error of $\mathcal{O}((\Delta t)^3)$, is but one example of *operatorial splittings* [1, 2, 20] and is intended here to illustrate a general point, namely that the solution of ‘complicated’ PDEs can be reduced to the solution of ‘simple’ PDEs and ODEs. Once done correctly, this procedure is consistent with eventual quality of the solution, concerning accuracy and stability alike.

Another benefit of (1.2) and of similar splittings is that it is consistent with conservation of the L_2 energy. Many *dispersive* equations, e.g. Schrödinger (linear or nonlinear), Gross–Pitaevskii, Dirac, Klein–Gordon and Korteweg–De Vries, conserve the L_2 norm of the solution. This often represents a highly significant physical feature, and it is vital to respect it under discretisation. (Note that conservation of L_2 norm automatically implies numerical stability.) Because of the special form of (1.2) (and of similar splittings), the overall numerical scheme preserves L_2 energy if both the discretisations of $e^{t\mathcal{L}}$ and e^{tf} do so. Insofar as e^{tf} is concerned, we can use the very extensive and robust existing theory [8], e.g. use a symplectic method (which automatically also preserves the L_2 norm). It is more challenging to ensure that $\|e^{t\mathcal{L}}v\|_2 = \|v\|_2$ for every v , in other words that the discretisation of $e^{t\mathcal{L}}$ is *unitary*.

1.2 The Differentiation Matrix

In this paper, we are concerned with spectral methods for time-dependent problems [3, 9, 27]. In a nutshell, we commence from a set $\Phi = \{\varphi_n\}_{n \in \mathbb{Z}_+}$, where each φ_n is defined in Ω and endowed with appropriate regularity. We assume that Φ is orthonormal in the standard L_2 inner product,

$$\int_{\Omega} \varphi_m(x)\varphi_n(x) \, dx = \delta_{m,n}, \quad m, n \in \mathbb{Z}_+,$$

and complete in $L_2(\Omega)$, and expand a solution in the basis Φ ,

$$u(x, t) = \sum_{n=0}^{\infty} u_n(t)\varphi_n(x),$$

where the coefficients $u_n(0)$ s are determined by expanding the initial condition u_0 , while the $u_n(t)$ s, $t > 0$ are typically evolved by Galerkin conditions, which for (1.1) read

$$u'_m(t) = \sum_{n=0}^{\infty} u_n(t)\langle \mathcal{L}\varphi_n + f(\cdot, \varphi_n), \varphi_m \rangle, \quad m \in \mathbb{Z}_+.$$

In a practical method, we truncate the expansion and the range of m , thereby obtaining a finite-dimensional linear system of ODEs.

Substantive advantage of spectral methods is that expansions in orthonormal bases typically converge very rapidly indeed: for example, orthogonal polynomials converge

in a finite interval to analytic (in the interval and its neighbourhood) functions at an exponential rate. Therefore, the number of degrees of freedom, compared to the more usual finite difference or finite element methods, is substantially smaller. While this is not the entire truth—finite differences and finite elements produce sparse linear algebraic systems while spectral elements yield dense matrices which sometimes can be also ill conditioned and, moreover, expanding a function in an orthonormal basis can be potentially costly—spectral methods are often the approach of choice in numerical computations. In the specific context of time-dependent problems, however, naive spectral methods are unstable [9]. This motivates us to consider the major concept of a *differentiation matrix*.

In the sequel, we restrict our narrative to the univariate case, $\Omega \subseteq \mathbb{R}$, for a number of reasons. Firstly, surprisingly, even the univariate case (as we hope to persuade the reader) is dramatically incomplete. Secondly, it lays the foundations to a multivariate case, whether by tensorial extension to parallelepipeds or by more advanced means which we intend to explore in a future paper.

The set Φ being a basis of $L_2(\Omega) \cap C^1(\Omega)$, any function therein can be expressed as a linear combination of the φ_n s, and this is particularly true with regard to the derivatives φ'_m . This yields a linear map represented by the infinite-dimensional matrix \mathcal{D} such that

$$\mathcal{D}_{m,n} = \int_{\Omega} \varphi'_m(x) \varphi_n(x) \, dx, \quad m, n \in \mathbb{Z}_+.$$

It is very simple to prove by integration by parts that the differentiation operator $\mathbf{D} = \partial/\partial x$ is skew Hermitian in the following three configurations of boundary conditions:

1. The *torus* $\Omega = \mathbb{T}$ (i.e. periodic boundary conditions);
2. The *Cauchy problem*: $\Omega = \mathbb{R}$; and
3. Zero Dirichlet boundary conditions on the boundary of $\Omega \subset \mathbb{R}$.

In that case $\|e^{t\mathbf{D}}\| = 1$ (unless otherwise stated, we assume in this paper the Euclidean norm.) and, \mathbf{D}^2 being Hermitian and negative semidefinite, $\|e^{t\mathbf{D}^2}\| \leq 1$. More generally, $\mathbf{D}^{2\ell+1}$ is skew Hermitian and $(-1)^{\ell-1} \mathbf{D}^{2\ell}$ negative semidefinite for all $\ell \in \mathbb{Z}_+$. Consequently, once $\mathcal{L} = \sum_{\ell=1}^M a_{\ell} \mathbf{D}^{\ell}$, where $(-1)^{\ell-1} a_{2\ell} \geq 0$, it is trivial to prove that $\langle \mathcal{L}u, u \rangle \leq 0$ for every u in the underlying Hilbert space.

This feature is retained by a spectral method, provided that \mathcal{D} is skew Hermitian, as is its $(N+1) \times (N+1)$ section \mathcal{D}_N . In the context of the PDE (1.1), it thus follows that, letting $\mathcal{L}_N = \sum_{\ell=1}^M a_{\ell} \mathcal{D}_N^{\ell}$, we have $w^* \mathcal{L}_N w \leq 0$ for all $w \in \mathbb{C}^{M+1}$. A consequence is that the L_2 energy is conserved for $\mathcal{L} = \mathbf{D}$ (the Schrödinger case) and it dissipates for $\mathcal{L} = \mathbf{D}^2$ (the diffusion equation case). In both cases, numerical stability comes in the wash.

1.3 Few Examples

This is the right moment to expand further on skew symmetry and differentiation matrices in a practical setting by means of few examples, restricting ourselves for simplicity to a single space dimension. Firstly, the reaction–diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u),$$

where f is a low-degree polynomial. In a finite-dimensional setting, the equation is replaced by $\mathbf{u}' = \mathcal{D}^* \mathcal{D} \mathbf{u} + f(\mathbf{u})$ and, provided \mathcal{D} is skew Hermitian, $\mathcal{D}^* \mathcal{D}$ is negative semidefinite. Once we use, e.g. the Strang splitting (1.2), the diffusion component \mathcal{D}^2 is assured to be dissipative. Another example is the convection–diffusion equation for incompressible flow (assuming, for simplicity, constant vector field),

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + c \frac{\partial u}{\partial x},$$

where $c \in \mathbb{R}$. Its semidiscretisation is $\mathbf{u}' = (\mathcal{D}^* \mathcal{D} + c \mathcal{D}) \mathbf{u}$ and, since for skew-Hermitian matrix \mathcal{D}

$$\frac{1}{2} [(\mathcal{D}^* \mathcal{D} + c \mathcal{D}) + (\mathcal{D}^* \mathcal{D} + c \mathcal{D})^*] = \mathcal{D}^* \mathcal{D},$$

a negative-semidefinite matrix, stability is assured.

Yet another example (and the underlying motivation to the work that has led to this paper) is dispersive equations of the form

$$i \frac{\partial u}{\partial t} = -\Delta u + f(x, u)$$

where $f(x, u) = V(x)u$ for linear Schrödinger equation, $f(x, u) = \lambda|u|^2 u$ for standard nonlinear Schrödinger equation and $f(x, u) = [V(x) + \lambda|u|^2]u$ for the Gross–Pitaevskii equation. All these equations conserve the L_2 energy $\|u\|$ (often known as ‘mass’ in this context), and this is a fundamental physical invariant. Once the equation is semi-discretised in the form $i \mathbf{u}' = -\mathcal{D}^* \mathcal{D} \mathbf{u} + \mathbf{F}(\mathbf{u})$ and solved, e.g. with the Strang splitting (1.2), the mass is conserved because $i \mathcal{D}^* \mathcal{D}$ is skew Hermitian.

Our last example is the KdV equation

$$\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u \frac{\partial u}{\partial x} = 0.$$

Strang splitting (1.2) means that we solve separately $\partial u / \partial t = -\mathcal{D}^3 u$ and scalar Burgers equations $\partial u_m / \partial t = -u_m (\mathcal{D} u)_m$. In each case, once \mathcal{D} is skew Hermitian, the solution is stable and L_2 energy conserved.

An important take-away lesson of our four examples is that we typically work (for real \mathcal{D}) with a specific power (or powers) of the differentiation matrix, predicated by the space derivatives present in the differential equation. In particular, we require not just \mathcal{D} but its specific powers to be bounded.

1.4 An Orthonormal Basis

The obvious choice of an orthonormal system is a set of orthogonal polynomials—unless we use (possibly shifted) Legendre polynomials, this means replacing the L_2 inner product by another, defined by the orthogonality weight function—but it is clear that this produces a lower triangular \mathcal{D} . While there are nontrivial ways round it [22], there is strong motivation to consider alternative orthonormal systems.

The periodic case—without loss of generality, $\Omega = [-\pi, \pi]$ with periodic boundary conditions—is obvious: we let $\Phi = \{e^{inx}\}_{n \in \mathbb{Z}}$, the Fourier basis. An added bonus is fast expansion by means of a Fast Fourier Transform of any $L_2[-\pi, \pi] \cap C_{\text{per}}[-\pi, \pi]$ function in the underlying basis. This is the paradigmatic case whereby a spectral method has few competitors.

The Cauchy case $\Omega = (-\infty, \infty)$ has been the subject of an extensive recent study [11–14]. In particular, all orthonormal systems Φ such that \mathcal{D} is skew Hermitian and tridiagonal have been completely characterised. Specifically, they are in a one-to-one relationship with Borel measures, supported on the entire real line. Let $d\mu(x) = w(x) dx$ be such a measure (w might be a generalised function) and $\mathcal{P} = \{p_n\}_{n \in \mathbb{Z}_+}$ the underlying set of orthonormal polynomials. It is elementary that \mathcal{P} obeys a three-term recurrence relation

$$\beta_n p_{n+1}(x) = (x + \alpha_n) p_n(x) - \beta_{n-1} p_{n-1}(x), \quad n \in \mathbb{Z}_+, \tag{1.3}$$

where $\beta_{-1} = 0$, $\alpha_n \in \mathbb{R}$ and $\beta_n > 0$ for $n \in \mathbb{Z}_+$. Inverse Fourier transforming $\{w^{1/2} p_n\}_{n \in \mathbb{Z}_+}$ and multiplying the n term by i^n , we obtain an orthonormal set Ψ , dense in $L_2(\mathbb{R})$ and such that

$$\psi'_n = -\beta_{n-1} \psi_{n-1} + i\alpha_n \psi_n + \beta_n \psi_{n+1}, \quad n \in \mathbb{Z}_+.$$

In other words, $\psi_n = i^n \mathcal{F}^{-1}(w^{1/2} p_n)$. Therefore, \mathcal{D} is skew Hermitian (skew symmetric if $\alpha_n \equiv 0$, which is the case once w is an even function) and tridiagonal, while orthonormality follows from the Plancherel theorem. Tridiagonality is a valuable feature because it is easy to manipulate \mathcal{D} (e.g. multiply \mathcal{D}_N by a vector or approximate $\exp(t\mathcal{D}_N)$) and the powers of the infinite-dimensional matrix \mathcal{D} (approximating higher derivatives) remain bounded.

This leaves us with the third—and most difficult—case, namely zero Dirichlet boundary conditions.¹ This is the subject of this paper.

A natural inclination is to extend the Fourier transform-based theory from $(-\infty, \infty)$ to, say, $(-1, 1)$. This can be done in one of two obvious ways and, unfortunately, both fail. The first is to choose a measure $d\mu$ supported by $(-1, 1)$, but this leads again to Φ supported on the entire real line, the only difference being that in this case the closure of Φ is not $L_2(-\infty, \infty)$ but a Paley–Wiener space [11]. Another possibility is to abandon altogether the Fourier route and alternatively commence by specifying a φ_0 , subsequently determining the φ_n s for $n \in \mathbb{N}$ and the matrix \mathcal{D} consistently

¹ It is elementary to reduce nonzero Dirichlet conditions to zero ones by reformulating the PDE for $u^{\text{new}} = u - \rho$, where ρ is any sufficiently regular function that obeys the correct Dirichlet boundary conditions on the boundary.

with both orthogonality and tridiagonality. A forthcoming paper demonstrates how to do this algorithmically. However, given $\varphi' = \mathcal{D}\varphi$, it follows by induction that $\varphi^{(s)} = \mathcal{D}^s\varphi$ for all $s \in \mathbb{Z}_+$. Consistency with zero Dirichlet boundary conditions, though, requires $\varphi_n(\pm 1) \equiv 0$, and this implies that $\varphi_n^{(s)}(\pm 1) \equiv 0$ for all $n, s \in \mathbb{Z}_+$. If φ_0 is analytic in $(-1, 1)$, this means that it necessarily must have an essential singularity at the endpoints. Intuitively, this is bad news, and this is confirmed by numerical experiments that indicate that the φ_n s develop boundary layers and wild oscillations near ± 1 and their approximation power is nil.

Both ideas above fall short and the current paper embarks on an altogether different approach, abandoning tridiagonality and the Fourier route altogether. Note that the existence of an essential singularity at the endpoints hinged on the fact that all powers of the infinite matrix \mathcal{D} are bounded. This is obvious once \mathcal{D} is tridiagonal (or, more generally, banded); hence, our main idea is to choose an orthonormal set \mathcal{D} such that \mathcal{D}^{s+1} blows up for some $s \in \mathbb{N}$. At the same time, we wish to retain a major blessing of tridiagonality, namely that $\mathcal{D}_N w$ can be computed in $\mathcal{O}(N)$ operations for any $w \in \mathbb{C}^{N+1}$ and, more generally, that \mathcal{D}_N is amenable for fast linear algebra.

1.5 Plan of this Paper

The main idea underlying this paper is exceedingly simple: given a measure $d\mu = w dx$, where $w \in \mathbb{C}^1(a, b)$, and an underlying set $\mathcal{P} = \{p_n\}_{n \in \mathbb{Z}_+}$ of orthonormal polynomials, we set

$$\varphi_n(x) = \sqrt{w(x)}p_n(x), \quad x \in (a, b). \tag{1.4}$$

It follows at once that Φ is orthonormal with respect to $L_2(a, b)$ and it is easy to determine conditions so that $\varphi_n(a) = \varphi_n(b) = 0$ for all $n \in \mathbb{Z}_+$. It is not difficult to specify the conditions on w for skew symmetry of \mathcal{D} . However, the narrative becomes more complicated once we seek a system such that \mathcal{D}^k is bounded for $k = 1, \dots, s$ and blows up for $k = s + 1$. Likewise, it is considerably more challenging to identify systems Φ that allow for fast computation of $\mathcal{D}_N w$ for $w \in \mathbb{C}^{N+1}$.

In Sect. 2, we introduce the functions (1.4) in a more rigorous setting of Sobolev spaces and explore general properties of their differentiation matrices. Section 3 is devoted to two families of weight functions, namely the *Laguerre family* $w_\alpha(x) = x^\alpha e^{-x} \chi_{(0, \infty)}(x)$ and the *ultraspherical family* $w_\alpha(x) = (1 - x^2)^\alpha \chi_{(-1, 1)}(x)$. We prove that both families have a *separable differentiation matrix*. This feature (put to a good use in Sect. 4) is very special—indeed, there are good reasons to conjecture that these two families are the only weights with this feature. We present a detailed example of two orthogonal families, *generalised Hermite weights* $w_\mu(x) = |x|^{2\mu} e^{-x^2}$ and *Konoplev weights* $w_{\alpha, \beta}(x) = |x|^{2\beta+1} (1 - x^2)^\alpha \chi_{(-1, 1)}(x)$, and prove that their differentiation matrices cannot be separable unless (for Konoplev weights) $\beta = -\frac{1}{2}$ and the weight reduces to ultraspherical. Note, however, that separability is just one feature lending itself to fast numerical algebra and one cannot rule out that other weights might lead to differentiation matrices which, while non-separable, are amenable to fast computation.

Finally, in Sect. 4 we demonstrate how separability of the differentiation matrix can be utilised for fast multiplication of $\mathcal{D}_N w$, $w \in \mathbb{R}^{N+1}$, in $\mathcal{O}(N)$ operations.

The original idea to consider functions of the form (1.4) in the specific case of *Freud weights* has been considered first by Luong [19], who demonstrated that in this specific case \mathcal{D} is a skew-symmetric, banded matrix with bandwidth seven. This was a serendipitous choice: in Sect. 2, we prove that the only weights that produce a banded matrix in the setting of (1.4) are generalised Freud weights!

2 W-Functions

2.1 The Definition and Some of Its Consequences

Let (a, b) be a non-empty real interval, $-\infty \leq a < b \leq \infty$, and $s \in \mathbb{N} \cup \{\infty\}$. We denote by $\mathring{H}_2^s[a, b]$ the Sobolev space of $H_2^s[a, b]$ functions f such that

$$f^{(k)}(a) = f^{(k)}(b) = 0, \quad k = 0, \dots, s - 1,$$

(note that $C^{s-1}[a, b] \subset H_2^s[a, b]$, therefore the derivatives are well defined) equipped with the inner product

$$\langle f, g \rangle_s = \sum_{k=0}^s \int_a^b f^{(k)}(x)g^{(k)}(x) \, dx.$$

A *weight function* $w \in L_2(a, b) \cap C^1(a, b)$ is a positive function with all its moments

$$\mu_k = \int_a^b x^k w(x) \, dx, \quad k \in \mathbb{Z}_+,$$

bounded. Given a weight functions, we can define (e.g. using a Gram–Schmidt process) a set of orthonormal polynomials $\mathcal{P} = \{p_n\}_{n \in \mathbb{Z}_+}$ such that

$$\int_a^b p_m(x)p_n(x)w(x) \, dx = \delta_{m,n}, \quad m, n \in \mathbb{Z}_+. \tag{2.1}$$

Such a set is unique once we require, for example, that the coefficient of x^n in p_n is always positive. We say that φ_n is the *n*th *W-function* once

$$\varphi_n(x) = \sqrt{w(x)}p_n(x), \quad n \in \mathbb{Z}_+,$$

and let $\Phi = \{\varphi_n\}_{n \in \mathbb{Z}_+}$. It follows at once from (2.1) that Φ is an orthonormal set with respect to the standard L_2 inner product.

Remark 1 The functions φ_n inherit some features of orthonormal polynomials; in particular, they obey the same three-term recurrence relation. However, the expansion coefficients of an arbitrary function are different:

$$f \sim \sum_{n=0}^{\infty} \hat{f}_n^P p_n, \quad \hat{f}_n^P = \int_a^b f(x) p_n(x) w(x) dx, \quad f \in L_2((a, b), w dx),$$

$$f \sim \sum_{n=0}^{\infty} \hat{f}_n^\Phi \varphi_n, \quad \hat{f}_n^\Phi = \int_a^b f(x) p_n(x) \sqrt{w(x)} dx, \quad f \in L_2(a, b).$$

Remark 2 An important difference between \mathcal{P} and Φ is that while the p_n are polynomials, hence analytic functions, the W-functions carry over potential singularities of the weight function. For example, for the Chebyshev weight function $w(x) = (1 - x^2)^{-1/2} \chi_{(-1,1)}(x)$ the φ_n s have weak singularity at the endpoints ± 1 , while their derivatives possess strong singularity there.

We let \mathcal{D} stand for the infinite-dimensional *differentiation matrix*

$$\mathcal{D}_{m,n} = \int_a^b \varphi'_m(x) \varphi_n(x) dx, \quad m, n \in \mathbb{Z}_+. \tag{2.2}$$

We say that w is of *index* $s \in \mathbb{N} \cup \{\infty\}$ and denote this by $\text{ind } w = s$ if \mathcal{D}^k is bounded for $k = 1, \dots, s$, while \mathcal{D}^{s+1} is unbounded.

Lemma 1 \mathcal{D} is skew symmetric if and only if $w(a) = w(b) = 0$.

Proof Assume first that $-\infty < a < b < \infty$ and note that \mathcal{D} is skew symmetric if and only if $\mathcal{D}_{m,n} + \mathcal{D}_{n,m} = 0, m, n \in \mathbb{Z}_+$. Since

$$\varphi'_n = \sum_{k=0}^{\infty} \mathcal{D}_{n,k} \varphi_k, \quad n \in \mathbb{Z}_+,$$

it follows from (2.2) and the orthonormality of Φ that \mathcal{D} is skew symmetric if

$$\int_a^b \frac{d\sqrt{w(x)} p_m(x)}{dx} \sqrt{w(x)} p_n(x) dx + \int_a^b \sqrt{w(x)} p_m(x) \frac{d\sqrt{w(x)} p_n(x)}{dx} dx = 0$$

for $m, n \in \mathbb{Z}_+$. The latter is equivalent, for every $m, n \in \mathbb{Z}_+$, to

$$\begin{aligned} & \int_a^b \left[\frac{w'(x)}{2\sqrt{w(x)}} p_m(x) + \sqrt{w(x)} p'_m(x) \right] \sqrt{w(x)} p_n(x) dx \\ & + \int_a^b \sqrt{w(x)} p_m(x) \left[\frac{w'(x)}{2\sqrt{w(x)}} p_n(x) + \sqrt{w(x)} p'_n(x) \right] dx = 0 \\ \Leftrightarrow & \int_a^b [w'(x) p_m(x) p_n(x) + w(x) p'_m(x) p_n(x) + w(x) p_m(x) p'_n(x)] dx = 0 \end{aligned}$$

$$\begin{aligned} &\Leftrightarrow \int_a^b \left[w'(x)p_m(x)p_n(x) + w(x) \frac{d p_m(x)p_n(x)}{dx} \right] dx = 0 \\ &\Leftrightarrow \int_a^b w'(x)p_m(x)p_n(x) dx + w(x)p_m(x)p_n(x) \Big|_a^b - \int_a^b w'(x)p_m(x)p_n(x) dx = 0 \\ &\Leftrightarrow w(b)p_m(b)p_n(b) = w(a)p_m(a)p_n(a). \end{aligned}$$

All the zeros of orthogonal polynomials reside in (a, b) . Therefore, they cannot vanish at the endpoints and $(-1)^k p_k(a)p_k(b) > 0, k \in \mathbb{N}$; hence, $p_m(a)p_n(a)$ cannot equal $p_m(b)p_n(b)$ for all $m, n \in \mathbb{N}$. We deduce that \mathcal{D} is skew symmetric if and only if $w(a) = w(b) = 0$.

The proof is similar—in fact, somewhat simpler—once either $b = \infty$ or $a = -\infty$. If $(a, b) = (-\infty, \infty)$, then L_2 boundedness and continuity imply $w(\pm\infty) = 0$, \mathcal{D} is skew symmetric, and there is nothing to prove. \square

Consequently, we impose an additional condition on the weight function, namely that it vanishes at the endpoints. Note that this is automatically true once an endpoint is infinite.

A quintessential example of a family of W-functions is *Hermite functions*

$$\varphi_n(x) = \frac{e^{-x^2/2}}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x), \quad n \in \mathbb{Z}_+,$$

where the H_n s are standard Hermite polynomials. Hermite functions are well known in mathematical physics, because they are eigenfunctions of the free Schrödinger operator. They can be derived from orthonormalised Hermite polynomials via the Fourier transform route, as mentioned in introduction and e.g. in Arieh Iserles and Marcus Webb [11]; hence, their differentiation matrix is tridiagonal. On the other hand, they are W-functions with $w(x) = e^{-x^2}$. Tridiagonality implies that $\text{ind } w = \infty$. More generally, $\text{ind } w = \infty$ once \mathcal{D} is a banded matrix and it is interesting to characterise all weight functions with this feature.

Theorem 2 *The differentiation matrix \mathcal{D} of a system of W-functions is banded if and only if $w(x) = e^{-c(x)}$, $x \in \mathbb{R}$, where c is an even-degree polynomial whose highest degree coefficient is strictly positive.*

Proof Letting $m \leq n - 1$, orthogonality implies that

$$\mathcal{D}_{m,n} = \int_a^b w \left(\frac{1}{2} \frac{w'}{w} p_m + p'_m \right) p_n dx = \frac{1}{2} \int_a^b w' p_m p_n dx$$

while for $m \geq n + 1$ skew symmetry yields

$$\mathcal{D}_{m,n} = -\frac{1}{2} \int_a^b w' p_m p_n dx. \tag{2.3}$$

Assume that \mathcal{D} has bandwidth $2L + 1$, in other words that $\mathcal{D}_{m,n} = 0$ for $|m - n| \geq L + 1$. It follows for $m \leq n - 1$ that

$$\int_a^b w X_m p_n \, dx = 0, \quad n \geq m + L + 1 \quad \text{where} \quad X_m = \frac{w'}{w} p_m.$$

If in addition $n \geq L$, expanding X_m in the basis \mathcal{P} it follows that X_m is a polynomial of degree $m + L + 1$. However,

$$w' = \frac{X_m}{p_m} w \quad \rightarrow \quad w(x) = w(x_0) \exp\left(\int_{x_0}^x \frac{X_m(y)}{p_m(y)} \, dy\right)$$

for some x_0 . Since w is independent of m , necessarily p_m divides X_m and the remainder c is a polynomial independent of m . Therefore, without loss of generality $w(x) = e^{-c(x)}$ where c is a polynomial of degree $L + 1$. w being integrable and $w(a) = w(b) = 0$, necessarily $a = -\infty, b = \infty$,² L is odd and c is an even-degree polynomial with strictly positive leading-degree coefficient. \square

We have recovered *precisely* the W-functions associated with generalised Freud polynomials that have been originally considered in Luong [19]. However, such W-functions are of little interest within the context of this paper, since we seek weight functions of finite index.

This is the point to note the expression (2.3) for the elements of \mathcal{D} such that $m \geq n + 1$. (If $m \leq n - 1$ we need to flip the sign.) We will make much use of it in the sequel.

2.2 The Boundedness of \mathcal{D}^s

We assume in this section that the weight w is strictly positive in (a, b) , as smooth in $[a, b]$ as needed in our construction, and set

$$q_j(x) = \frac{d^j \sqrt{w(x)}}{dx^j}, \quad j \in \mathbb{Z}_+.$$

Therefore,

$$\varphi_m^{(\ell)} = \sum_{j=0}^{\ell} \binom{\ell}{j} q_j p_m^{(\ell-j)}, \quad \ell, m \in \mathbb{Z}_+.$$

As long as \mathcal{D}^s is bounded, we have

$$(\mathcal{D}^s)_{m,n} = \int_a^b \varphi_m^{(s)} \varphi_n \, dx = \sum_{j=0}^s \binom{s}{j} \int_a^b \sqrt{w} q_j p_m^{(s-j)} p_n \, dx. \tag{2.4}$$

² In a finite interval, we would have had an essential singularity at an endpoint.

It is trivial to prove that

$$q_r = \sum_{j=1}^r \frac{U_{r,j}}{w^{j-\frac{1}{2}}},$$

where each $U_{r,j}$ is a linear combination of products of the form $\prod_i w^{(\ell_i)}$ such that $\ell_i \geq 1$ and $\sum_i \ell_i = r$: for example,

$$U_{4,1} = \frac{1}{2}w^{(4)}, \quad U_{4,2} = -\frac{3}{4}w''^2 - w'w''', \quad U_{4,3} = \frac{9}{4}w'^2w'', \quad U_{4,4} = -\frac{15}{16}w'^4.$$

In general, $U_{r,r} = (-1)^{r-1}(2r)!w''^r/(4^r r!)$, $r \in \mathbb{N}$.

Since $w(x) > 0$ in (a, b) , the only possible source of singularity in (2.4) is that $\varphi_m^{(s)}$ is non-integrable at an endpoint. Recalling that $w(a) = w(b) = 0$, integrability is lost exclusively when dividing by a power of w , and the larger the power, the more significant the singularity. In other words, $\varphi_m^{(s)}$ is bounded for all $m \in \mathbb{Z}_+$ only if the integral

$$\int_a^b \sqrt{w}q_s p \, dx$$

is bounded for any polynomial p , and this is contingent on $\tilde{w}_s = w^s/w^{s-1}$ being a *signed* weight function, i.e. all its moments exist and $\tilde{w}_s \not\equiv 0$.³ The following theorem is thereby true.

Theorem 3 *A necessary condition for $\text{ind } w \geq s$ is that \tilde{w}_r , $r = 2, \dots, s$, are signed weight functions.*

Let $-\infty < a < b < \infty$. Regularity and $w(a) = w(b) = 0$ imply that

$$w(x) = (x - a)^\alpha (b - x)^\beta v(x), \quad x \in [a, b], \quad v(a), v(b) \neq 0. \quad (2.5)$$

Therefore, after elementary algebra,

$$\tilde{w}_s = (x - a)^{\alpha-s} (b - x)^{s-\beta} v \left[(\alpha\beta + \beta\alpha) - (\alpha + \beta)x + (x - a)(b - x) \frac{v'}{v} \right]^s.$$

Theorem 4 *A necessary condition for $\text{ind } w \geq s$ in a finite interval (a, b) is that $\alpha, \beta > s - 1$. Likewise, once $b = \infty$, we need $\alpha > s - 1$ and for $a = -\infty$ the condition is $\beta > s - 1$.*

Proof Consistently with our assumptions, $v \neq 0$ in $[a, b]$; therefore, the only source of singularity may come from $(x - a)^{\alpha-s}$ and $(b - x)^{\beta-s}$. We conclude that, for \mathcal{D}^s to be bounded, we need $\alpha, \beta > s - 1$. The semi-infinite cases follow in an identical (and simpler!) manner. □

³ \tilde{w}_1 cannot be a ‘true’ weight function because $w(a) = w(b) = 0$; hence, w' changes sign in (a, b) .

In the special case $s = 2$, we can complement Theorem 3 with a sufficient condition.

Theorem 5 *A necessary and sufficient condition for $\text{ind } w \geq 2$ is that $\tilde{w}_2 = w'^2/w$ is a signed measure.*

Proof We compute \mathcal{D}^2 directly. Using skew symmetry,

$$\mathcal{D}_{m,n}^2 = \sum_{\ell=0}^{\infty} \mathcal{D}_{m,\ell} \mathcal{D}_{\ell,n} = - \sum_{\ell=0}^{n-1} \mathcal{D}_{m,\ell} \mathcal{D}_{n,\ell} + \sum_{\ell=n+1}^{m-1} \mathcal{D}_{m,\ell} \mathcal{D}_{\ell,n} - \sum_{\ell=m+1}^{\infty} \mathcal{D}_{\ell,m} \mathcal{D}_{\ell,n}.$$

Recalling (2.3), let us consider the infinite sum

$$\begin{aligned} - \sum_{\ell=m+1}^{\infty} \mathcal{D}_{\ell,m} \mathcal{D}_{\ell,n} &= -\frac{1}{4} \int_a^b \int_a^b w'(x)w'(y) p_m(x)p_n(y) \sum_{\ell=m+1}^{\infty} p_{\ell}(x)p_{\ell}(y) \, dx \, dy \\ &= -\frac{1}{4} \int_a^b \int_a^b w'(x)w'(y) p_m(x)p_n(y) \sum_{\ell=0}^{\infty} p_{\ell}(x)p_{\ell}(y) \, dx \, dy \\ &\quad + \frac{1}{4} \sum_{\ell=0}^m \int_a^b w'(x) p_m(x) p_{\ell}(x) \, dx \int_a^b w'(y) p_n(y) p_{\ell}(y) \, dy \\ &= -\frac{1}{4} \int_a^b \int_a^b w'(x)w'(y) p_m(x)p_n(y) \sum_{\ell=0}^{\infty} p_{\ell}(x)p_{\ell}(y) \, dx \, dy \\ &\quad + \sum_{\ell=0}^{n-1} \mathcal{D}_{m,\ell} \mathcal{D}_{n,\ell} - \sum_{\ell=m+1}^{n-1} \mathcal{D}_{m,\ell} \mathcal{D}_{\ell,n}. \end{aligned}$$

Therefore,

$$\mathcal{D}_{m,n}^2 = -\frac{1}{4} \int_a^b \int_a^b w'(x)w'(y) p_m(x)p_n(y) \sum_{\ell=0}^{\infty} p_{\ell}(x)p_{\ell}(y) \, dx \, dy. \tag{2.6}$$

Let \mathcal{P} be orthonormal and complete in $L_2((a, b), w \, dx)$ and $f \in L_2((a, b), w \, dx)$. Then

$$f(x) = \sum_{m=0}^{\infty} \hat{f}_m p_m(x), \quad \text{where} \quad \hat{f}_m = \int_a^b w(x) f(x) p_m(x) \, dx.$$

Moreover, by the Parseval theorem,

$$\int_a^b w(x) |f(x)|^2 \, dx = \|f\|^2 = \sum_{m=0}^{\infty} |\hat{f}_m|^2. \tag{2.7}$$

Since

$$|\hat{f}_m|^2 = \int_a^b \int_a^b w(x)w(y)f(x)\overline{f(y)}p_m(x)p_m(y) \, dx \, dy,$$

exchanging summation and integration we have

$$\|f\|^2 = \int_a^b \int_a^b w(x)w(y)f(x)\overline{f(y)} \sum_{m=0}^{\infty} p_m(x)p_m(y) \, dx \, dy.$$

Let

$$K(x, y) = \sqrt{w(x)w(y)} \sum_{m=0}^{\infty} p_m(x)p_m(y)$$

be the *Christoffel–Darboux kernel*. It now follows from (2.7) that for every $f \in L_2((a, b), w \, dx)$ it is true that

$$\int_a^b w(x)|f(x)|^2 \, dx = \int_a^b \int_a^b \sqrt{w(x)w(y)}f(x)\overline{f(y)}K(x, y) \, dx \, dy$$

and we deduce that

$$K(x, y) = \delta_{x-y}. \tag{2.8}$$

In other words, K is a reproducing kernel.⁴

We now return to (2.6), deducing that

$$\begin{aligned} \mathcal{D}_{m,n}^2 &= -\frac{1}{4} \int_a^b \int_a^b \frac{w'(x)w'(y)}{\sqrt{w(x)w(y)}} p_m(x)p_n(y)K(x, y) \, dx \, dy \\ &= -\frac{1}{4} \int_a^b \frac{w'^2(x)}{w(x)} p_m(x)p_n(x) \, dx = -\frac{1}{4} \int_a^b \tilde{w}_2(x)p_m(x)p_n(x) \, dx. \end{aligned}$$

This is bounded because \tilde{w}_2 is a signed measure and $p_m p_n$ a polynomial, and we deduce that $\text{ind } w \geq 2$. The necessity of \tilde{w}_2 being a signed measure is obvious from the argument that led to Theorem 4. □

3 Separable Systems

In this section, we consider two families of weight functions that share a hugely beneficial attribute of *separability*, and we also provide two examples of weights that lack this feature.

⁴ While this is probably known, the author failed to find this result in literature, even in the encyclopaedic review of the Christoffel–Darboux kernel in Simon [25], see also Refs. [15, 18]—the reason might well be that the emphasis is usually on general Borel measures, rather than on $d\mu = w \, dx$ with $w \in C^1(a, b)$. One way or the other, the proof is included for completeness.

We say that a weight function w is *separable* if there exist real sequences $\mathbf{a} = \{a_n\}_{n \in \mathbb{Z}_+}$ and $\mathbf{b} = \{b_n\}_{n \in \mathbb{Z}_+}$ such that

$$\mathcal{D}_{m,n} = \begin{cases} a_m b_n, & m \geq n + 1, \\ 0, & m = n, \\ -a_n b_m, & m \leq n - 1, \end{cases} \quad m, n \in \mathbb{Z}_+ \tag{3.1}$$

and it is *symmetrically separable* subject to the existence of real sequences $\mathbf{a} = \{a_n\}_{n \in \mathbb{Z}_+}$ and $\mathbf{b} = \{b_n\}_{n \in \mathbb{Z}_+}$ such that

$$\mathcal{D}_{m,n} = \begin{cases} -a_m b_n & m + n \text{ odd}, m \geq n + 1, \\ 0, & m + n \text{ even}, \\ a_m b_n, & m + n \text{ odd}, m \leq n - 1, \end{cases} \quad m, n \in \mathbb{Z}_+ \tag{3.2}$$

It is demonstrated in Sect. 4 that separability or symmetric separability allows for very rapid computation of products of the form $\mathcal{D}_N \mathbf{v}$ for $\mathbf{v} \in \mathbb{R}^{N+1}$. This is intimately related to earlier results on fast computation of some structured algebraic systems in Refs. [6, 7].

In this section, we consider two families of measures, one separable and the other symmetrically separable: the *Laguerre weight* $w(x) = x^\alpha e^{-x} \chi_{(0,\infty)}(x)$ and the *ultraspherical weight* $(1-x^2)^\alpha \chi_{(-1,1)}(x)$, respectively. In a way, they are the most obvious measures in intervals of the form $(0, \infty)$ and $(-1, 1)$, respectively. Yet, interestingly, separability appears to be a very rare feature and we provide counterexamples further in this section.

In both Laguerre and ultraspherical cases, we are able to present comprehensive analysis, deriving the sequences \mathbf{a} , \mathbf{b} explicitly, determining $\text{ind } w$ and (in Sect. 4) discussing the optimal choice of the parameter α . While both Laguerre and ultraspherical polynomials have been comprehensively studied, the separability and the formulæ (3.1) and (3.2) are, to the best of author’s knowledge, new.

3.1 The Laguerre Family

Laguerre polynomials are orthogonal with respect to the Laguerre weight,

$$\int_0^\infty x^\alpha e^{-x} L_m^{(\alpha)}(x) L_n^{(\alpha)}(x) dx = \frac{\Gamma(n+1+\alpha)}{n!} \delta_{m,n}, \quad m, n \in \mathbb{Z}_+, \quad \alpha > -1.$$

[23, p. 206]. In our case, we consider just the case $\alpha > 0$, so that the weight function vanishes at the origin. We have

$$p_n(x) = \sqrt{\frac{n!}{\Gamma(n+1+\alpha)}} L_n^{(\alpha)}(x),$$

$$\varphi_n(x) = \sqrt{\frac{n!}{\Gamma(n+1+\alpha)}} x^{\alpha/2} e^{-x/2} L_n^{(\alpha)}(x), \quad n \in \mathbb{Z}_+.$$

In Theorem 12, we determine that the Laguerre weight is separable—the proof requires a fair bit of algebraic computation and is relegated to Appendix A. The separability coefficients are given in (A.8), which we repeat here for clarity,

$$a_m = \sqrt{\frac{m!}{2\Gamma(m+1+\alpha)}} \sim \frac{1}{m^{\alpha/2}}, \quad b_n = \sqrt{\frac{\Gamma(n+1+\alpha)}{2n!}} \sim n^{\alpha/2}, \quad m, n \in \mathbb{Z}_+. \tag{3.3}$$

Note that

$$a_m b_m \equiv \frac{1}{2}, \quad m \in \mathbb{Z}_+ \tag{3.4}$$

this will be important in the sequel.

Theorem 4 presents a necessary condition for $\text{ind } w \geq s$ for $s \geq 2$: for a Laguerre weight $w = w_\alpha$ it translates to $\alpha > s - 1$. In the remainder of this subsection, we wish to prove that for the Laguerre weight function this condition is also sufficient.

The matrix \mathcal{D}^s is *absolutely bounded* for $s \geq 0$ if

$$\sum_{k_1=0}^\infty \sum_{k_2=0}^\infty \cdots \sum_{k_{s-1}=0}^\infty |\mathcal{D}_{m,k_1} \mathcal{D}_{k_1,k_2} \cdots \mathcal{D}_{k_{s-2},k_{s-1}} \mathcal{D}_{k_{s-1},n}| < \infty, \quad m, n \in \mathbb{Z}_+. \tag{3.5}$$

It is clear that absolute boundedness implies boundedness.

We assume that $m \geq n + 1$ and observe that everything depends on the interplay of the relative sizes of $k_0 = m, k_1, k_2, \dots, k_{s-1}, k_s = n$ because, for example,

$$k_j > k_{j+1} \Rightarrow |\mathcal{D}_{k_j,k_{j+1}}| = a_{k_j} b_{k_{j+1}}, \quad k_j < k_{j+1} \Rightarrow |\mathcal{D}_{k_j,k_{j+1}}| = a_{k_{j+1}} b_{k_j}.$$

We can disregard the case $k_j = k_{j+1}$ because then $\mathcal{D}_{k_j,k_{j+1}} = 0$ and the entire product vanishes; hence, we assume that always $k_j \neq k_{j+1}$. We use the shorthand \searrow for $k_j > k_{j+1}$ and \nearrow for $k_j < k_{j+1}$. Note that once s is even then \mathcal{D}^s is symmetric and diagonal elements no longer vanish: in that case we need to consider also the case $m = n$ but the proof is identical.

To illustrate our argument, for $s = 4$ we have eight options:

$\searrow \searrow \searrow = \searrow^3:$	$k_0 > k_1 > k_2 > k_3,$	$\searrow \nearrow \searrow:$	$k_0 > k_1 < k_2 > k_3,$
$\searrow \searrow \nearrow = \searrow^2 \nearrow:$	$k_0 > k_1 > k_2 < k_3,$	$\searrow \nearrow \nearrow = \searrow \nearrow^2:$	$k_0 > k_1 < k_2 < k_3,$
$\nearrow \searrow \searrow = \nearrow \searrow^2:$	$k_0 < k_1 > k_2 > k_3,$	$\nearrow \nearrow \searrow = \nearrow^2 \searrow:$	$k_0 < k_1 < k_2 > k_3,$
$\nearrow \searrow \nearrow:$	$k_0 < k_1 > k_2 < k_3,$	$\nearrow \nearrow \nearrow = \nearrow^3:$	$k_0 < k_1 < k_2 < k_3,$

except that \nearrow^3 is impossible because $k_0 = m > n = k_2$.

We let \mathcal{Q}_N stand for a generic polynomial of degree exactly N and note for further use the following technical result with a straightforward proof.

Proposition 6 *The sum*

$$\sum_{k=1}^K \frac{\mathcal{Q}_N(k)}{k^\alpha} \sim c K^{N-\alpha+1}, \quad K \gg 1, \quad N \neq \alpha - 1,$$

converges as $K \rightarrow \infty$ if and only if $\alpha > N + 1$. Here c is a constant, while “ \sim ” means that we are disregarding lower-order terms.

In general, the main idea is to write a sequence of \nearrow s and \searrow s in the form

$$\nearrow^{i_1} \searrow^{j_1} \nearrow^{i_2} \searrow^{j_2} \dots \nearrow^{i_r} \searrow^{j_r},$$

where $i_k, j_k \geq 0$ and $\sum_{k=1}^r (i_k + j_k) = s$. We call \nearrow^r a \nearrow -pre-chain of length r and \searrow^r a \searrow -pre-chain of length r ; in other words, we decompose each product in (3.5) into a sequence of alternating pre-chains.

Consider first an \nearrow -pre-chain of length $r \geq 1$. Because of (3.4), it equals

$$\begin{aligned} & \sum_{k_\ell=0}^{k_{\ell-1}-1} \sum_{k_{\ell+1}=0}^{k_\ell-1} \dots \sum_{k_{\ell+r-1}=0}^{k_{\ell+r-2}-1} |\mathcal{D}_{k_{\ell-1}, k_\ell} \mathcal{D}_{k_\ell, k_{\ell+1}} \dots \mathcal{D}_{k_{\ell+r-2}, k_{\ell+r-1}}| \\ &= \sum_{k_\ell=0}^{k_{\ell-1}-1} \sum_{k_{\ell+1}=0}^{k_\ell-1} \dots \sum_{k_{\ell+r-1}=0}^{k_{\ell+r-2}-1} \prod_{j=\ell-1}^{\ell+r-2} \mathbf{a}_{k_j} \mathbf{b}_{k_{j+1}} = \frac{\mathbf{a}_{k_{\ell-1}}}{2^{r-1}} \sum_{k_\ell=0}^{k_{\ell-1}-1} \sum_{k_{\ell+1}=0}^{k_\ell-1} \dots \sum_{k_{\ell+r-1}=0}^{k_{\ell+r-2}-1} \mathbf{b}_{k_{\ell+r-1}}. \end{aligned}$$

We say that $\mathbf{a}_{k_{\ell-1}}$ and $\mathbf{b}_{k_{\ell+r-1}}$ are the *head* and the *tail* of the pre-chain, respectively.

Likewise, for an \searrow -pre-chain of length $r \geq 1$ we have

$$\begin{aligned} & \sum_{k_\ell=k_{\ell-1}+1}^{\infty} \sum_{k_{\ell+1}=k_\ell+1}^{\infty} \dots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty} |\mathcal{D}_{k_{\ell-1}, k_\ell} \mathcal{D}_{k_\ell, k_{\ell+1}} \dots \mathcal{D}_{k_{\ell+r-2}, k_{\ell+r-1}}| \\ &= \frac{\mathbf{b}_{k_{\ell-1}}}{2^{r-1}} \sum_{k_\ell=k_{\ell-1}+1}^{\infty} \sum_{k_{\ell+1}=k_\ell+1}^{\infty} \dots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty} \mathbf{a}_{k_{\ell+r-1}}. \end{aligned}$$

Now $\mathbf{b}_{k_{\ell-1}}$ and $\mathbf{a}_{k_{\ell+r-1}}$ are the head and the tail of the pre-chain, respectively.

Except for $\ell = 0$ and $\ell + r = s$, we join the tail of a pre-chain to the head of the succeeding pre-chain. The outcome is \nearrow -chains and \searrow -chains. Note thus that a chain has no head, while its tail is multiplied by the head of its successor pre-chain. (In this procedure, we lose the head of the leading pre-chain and the tail of the last pre-chain, but this makes no difference to the finiteness—or otherwise—of the sum)

An \nearrow -chain of length r is of the form

$$\frac{1}{2^{r-1}} \sum_{k_\ell=0}^{k_{\ell-1}-1} \sum_{k_{\ell+1}=0}^{k_\ell-1} \dots \sum_{k_{\ell+r-1}=0}^{k_{\ell+r-2}-1} \mathbf{b}_{k_{\ell+r-1}}^2 = \frac{1}{2^r} \sum_{k_\ell=0}^{k_{\ell-1}-1} \sum_{k_{\ell+1}=0}^{k_\ell-1} \dots \sum_{k_{\ell+r-1}=0}^{k_{\ell+r-2}-1} \frac{\Gamma(k_\ell + r + \alpha)}{(k_\ell + r - 1)!},$$

a finite sum. Hence, it cannot be a source for unboundedness of the sum (3.5). Matters are different, though, with an \searrow -chain of length r : straightforward algebra and Proposition 6 imply that

$$\sum_{k_\ell=k_{\ell-1}+1}^{\infty} \sum_{k_{\ell+1}=k_\ell+1}^{\infty} \dots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty} \mathbf{a}_{k_{\ell+r-1}}^2$$

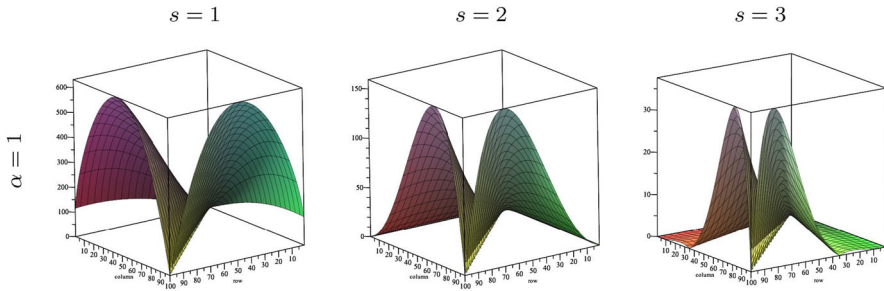


Fig. 1 Laguerre W-functions: The magnitude of \mathcal{D}^s for different values of α and $1 \leq s \leq 3$

$$\begin{aligned}
 &= \sum_{k_{\ell+r-1}=k_{\ell-1}+r}^{\infty} a_{k_{\ell+r-1}}^2 \sum_{k_{\ell}=k_{\ell-1}+r-1}^{k_{\ell+r-1}-r+1} \sum_{k_{\ell+1}=k_{\ell-1}+r-2}^{k_{\ell+r-1}-r+2} \cdots \sum_{k_{\ell+r-2}=k_{\ell-1}+1}^{k_{\ell+r-1}-1} 1 \\
 &= \sum_{k_{\ell+r-1}=k_{\ell-1}+r}^{\infty} a_{k_{\ell+r-1}}^2 Q_{r-1}(k_{\ell+r-1}) \sim \sum_{\ell=k_{\ell-1}+r}^{\infty} \frac{1}{\ell^{\alpha-r}}.
 \end{aligned}$$

Therefore, boundedness takes place if $\alpha - r > 1$.

Since the length of any chain is at most $s - 1$ and \searrow^{s-1} is impossible (recall, $k_0 > k_s$), the maximal length of an \searrow -chain is $s - 2$. We thus deduce that $\alpha > s - 1$.

Theorem 7 *ind $w_\alpha \geq s$ for the Laguerre weight if and only if $\alpha > s - 1$.*

Proof The necessity is proved in Theorem 4, while sufficiency follows because absolute boundedness in (3.5) implies boundedness. □

In Fig. 1, we display the absolute values of the entries of \mathcal{D}^s for different values of α and s . The computation involves infinite matrices, hence infinite products which need be truncated in computation. Thus, we compute 300×300 matrices and their powers, while displaying just their 100×100 section, since this minimises the truncation effects. For $s = 1$, all differentiation matrices are bounded and of a moderate size; the sole difference is that as α grows, the matrix becomes more ‘centred’ about the diagonal. However, already for $s = 2$ the difference is discernible. For $\alpha = 1$, we are right on the boundary of $\alpha > s - 1$ (on its wrong side!) and the size of \mathcal{D}^2 grows rapidly: had we displayed a section of an $M \times M$ matrix for $M \gg 300$, the magnitude would have grown at a logarithmic rate, as indicated by the proof of absolute boundedness. Once $\alpha > 1$, the magnitude grows at a slower rate and would remain bounded for $M \rightarrow \infty$. Finally, for $s = 3$ the cases $\alpha = 1$ and $\alpha = 2$ correspond to polynomial and logarithmic growth, respectively, and this is apparent in the figure. Finally, for $\alpha = 4$ the rate of growth slows down and it is persuasive that the magnitude remains bounded as $M \rightarrow \infty$. Note that even in a ‘good’ α regime the magnitude, while decaying along rows and columns, grows along diagonals. We refer to the discussion following Fig. 2 for an explanation of this behaviour, commenting here that this phenomenon follows from $a_m \sim m^{-\alpha/2}$ and $b_n \sim n^{\alpha/2}$.

It follows from Theorem 7 that once we approximate functions in $\mathring{H}_2^s(0, \infty)$, we need to choose $\alpha > s - 1$. However, there is much more to the choice of a good α and

we defer its discussion to Section 4. As it turns out, the quality of approximation is exceedingly sensitive to the right choice and numerical results indicate that there exists a ‘sweet spot’ that brings about substantially improved quality of approximation.

3.2 The Ultraspherical Family

The ultraspherical weight⁵ a special case of the *Jacobi weight*, is $w_\alpha(x) = (1 - x^2)^\alpha \chi_{(-1,1)}(x)$, $\alpha > 1$ —in our case the requirement $w_\alpha(\pm 1) = 0$ restricts α to the range $(0, \infty)$. We have

$$p_n(x) = g_n^\alpha P_n^{(\alpha,\alpha)}(x),$$

$$\varphi_n(x) = g_n^\alpha (1 - x^2)^{\alpha/2} P_n^{(\alpha,\alpha)}(x), \quad n \in \mathbb{Z}_+,$$

where the constant

$$g_n^\alpha = \frac{\sqrt{\frac{1}{2}n!(2n + 2\alpha + 1)\Gamma(n + 2\alpha + 1)}}{2^\alpha \Gamma(n + \alpha + 1)}$$

orthonormalises an ultraspherical polynomial [23, p. 260]. Recalling the identity (2.3), we let

$$\mathcal{E}_{m,n} = \frac{1}{\alpha g_m^\alpha g_n^\alpha} \mathcal{D}_{m,n} = \int_{-1}^1 (1 - x^2)^{\alpha-1} x P_m^{(\alpha,\alpha)}(x) P_n^{(\alpha,\alpha)}(x) dx.$$

It is sufficient to derive the $\mathcal{E}_{m,n}$ s explicitly and prove that \mathcal{E} is symmetrically separable. We recall that our interest is in odd values of $m+n$ and assume without loss of generality that $m \geq n + 1$

Let

$$S_{m,n}^\alpha = \int_{-1}^1 (1 - x^2)^{\alpha-1} P_m^{(\alpha,\alpha)}(x) P_n^{(\alpha,\alpha)}(x) dx,$$

noting that $S_{m,n}^\alpha = 0$ if $m+n$ is odd. The three-term recurrence relation for orthonormal ultraspherical polynomials is

$$x P_m^{(\alpha,\alpha)}(x) = \frac{(m + 1)(m + 2\alpha + 1)}{(m + 1 + \alpha)(2m + 2\alpha + 1)} P_{m+1}^{(\alpha,\alpha)}(x) + \frac{m + \alpha}{2m + 2\alpha + 1} P_{m-1}^{(\alpha,\alpha)}(x), \quad (3.6)$$

as can be easily confirmed from Rainville [23, p. 263]. Therefore,

$$\begin{aligned} \mathcal{E}_{m,n} &= \int_{-1}^1 (1 - x^2)^{\alpha-1} \left[\frac{(m + 1)(m + 2\alpha + 1)}{(m + 1 + \alpha)(2m + 2\alpha + 1)} P_{m+1}^{(\alpha,\alpha)} + \frac{m + \alpha}{2m + 2\alpha + 1} P_{m-1}^{(\alpha,\alpha)} \right] P_n^{(\alpha,\alpha)} dx \end{aligned}$$

⁵ Also known, subject to different scaling, as the Gegenbauer weight [23, p. 276].

$$= \frac{(m + 1)(m + 2\alpha + 1)}{(m + 1 + \alpha)(2m + 2\alpha + 1)} S_{m+1,n}^\alpha + \frac{m + \alpha}{2m + 2\alpha + 1} S_{m-1,n}^\alpha. \tag{3.7}$$

Our next task is determining the explicit form of $S_{m,n}^\alpha$ for even $m + n$ and, without loss of generality, $m \geq n$. This is accomplished in Appendix B and results in

$$S_{m,n}^\alpha = \frac{4^\alpha \Gamma(m + 1 + \alpha)\Gamma(n + 1 + \alpha)}{\alpha n!\Gamma(m + 1 + 2\alpha)}, \quad m \geq n, \quad m + n \text{ even.}$$

We conclude from (3.7) that

$$\mathcal{E}_{m,n} = \frac{4^\alpha \Gamma(m + 1 + \alpha)\Gamma(n + 1 + \alpha)}{\alpha n!\Gamma(m + 1 + 2\alpha)}, \quad m \geq n, \quad m + n \text{ odd}$$

and

$$\mathcal{D}_{m,n} = \alpha g_m^\alpha g_n^\alpha \mathcal{E}_{m,n} = \frac{1}{2} \sqrt{\frac{m!(2m + 2\alpha + 1)(2n + 2\alpha + 1)\Gamma(n + 1 + 2\alpha)}{n!\Gamma(m + 1 + 2\alpha)}} \tag{3.8}$$

is valid for all odd $m + n$, $m \geq n + 1$ —once $n \geq m + 1$, we need to swap m and n and invert the sign. (Of course, $\mathcal{D}_{m,n} = 0$ once $m + n$ is even.)

Our first conclusion is that the measure is symmetrically separable with

$$a_m = \sqrt{\frac{m!(2m + 2\alpha + 1)}{2\Gamma(m + 1 + 2\alpha)}}, \quad b_n = \sqrt{\frac{(2n + 2\alpha + 1)\Gamma(n + 1 + 2\alpha)}{2n!}}. \tag{3.9}$$

The next conclusion is that the rate of growth (or decay) is dramatically different along the rows and the columns of \mathcal{D} . It follows from (3.8) and the standard Stirling formula [21, 5.11.3] that

$$\mathcal{D}_{m,n} \sim \frac{n^{\alpha+\frac{1}{2}}}{m^{\alpha-\frac{1}{2}}}, \quad m, n \gg 1, \quad m \geq n + 1.$$

Therefore, the elements of the differentiation matrix decay geometrically (at any rate, for $\alpha > \frac{1}{2}$) along rows (and, because of skew symmetry, columns) and increase geometrically along diagonals. Note that, forming powers of \mathcal{D} , it is the decay along rows and columns that allows for boundedness. (Incidentally, it can be proved using special functions that $(\mathcal{D}^2)_{0,0} = \alpha(2\alpha + 1)/[4(\alpha - 1)]$, driving home the fact, already known from Theorem 5, that $\alpha > 1$ is necessary and sufficient for boundedness. We leave the proof, which plays no further role in our narrative, as an exercise for the reader.)

Figure 2 displays the magnitude of the powers of differentiation matrix for ultraspherical weights and different values of α and s , using the same rules of engagement as in Fig. 1. Two trends are discernible, both following from our discussion. Firstly, the decay along rows accelerates as α grows and \mathcal{D}^s is more concentrated near the diagonal. Secondly, the terms along the diagonal (of course, with $m + n$ of the right

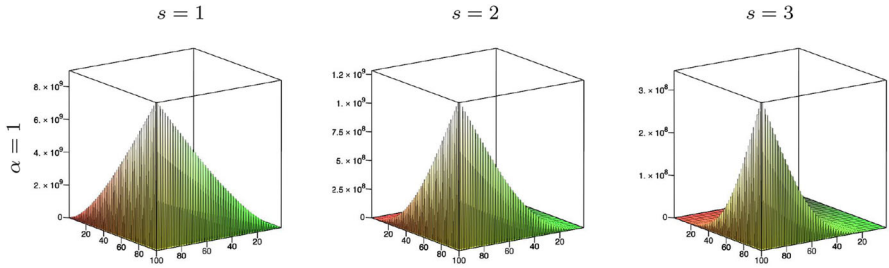


Fig. 2 Ultraspherical W-functions: The magnitude of \mathcal{D}^s for different values of α and $1 \leq s \leq 3$

parity) grow the fastest. Their rate of growth is rapid (and grows with α), but this need not be a problem, at any rate once we approximate sufficiently smooth functions. In that case \mathcal{D} , its powers and possibly functions (e.g. $\exp(h\mathcal{D})$) act on the expansion coefficients of functions in the underlying basis Φ . Provided these functions are sufficiently smooth, it is plausible that these coefficients decay very rapidly and, for analytic functions, at an exponential rate. (We defer to Sect. 4 for more substantive discussion of convergence.) Thus, large terms along the diagonal will multiply small terms in a vector of expansion coefficients—something that might conceivably cause loss of accuracy for truly huge matrices but which is probably negligible in practice.

Similar to Laguerre weights, we now seek to prove that Theorem 4 provides also a sufficient condition for $\text{ind } w_\alpha \geq s$ for ultraspherical weights, i.e. that $\alpha > s - 1$ implies that \mathcal{D}^s is bounded. Our method of proof is similar to that of Theorem 7, except that we need to account for a number of differences: firstly $\mathcal{D}_{m,n}$ can be nonzero only when $m + n$ is even, secondly, we have symmetric separability in place of separability, and thirdly, (3.4) is no longer true and needs to be replaced by

$$a_m b_m = m + \alpha + \frac{1}{2}. \tag{3.10}$$

Letting again $k_0 = m, k_s = n$, where $m \geq n + 1$ (or $m \geq n$ once s is even), we need to replace (3.5) by

$$\sum_{k_1=0}^{\infty} \star \sum_{k_2=0}^{\infty} \star \cdots \sum_{k_{s-1}=0}^{\infty} \star |\mathcal{D}_{k_0,k_1} \mathcal{D}_{k_1,k_2} \cdots \mathcal{D}_{k_{s-2},k_{s-1}} \mathcal{D}_{k_{s-1},k_s}| < \infty,$$

where the star means that we sum only over pairs (k_{i-1}, k_i) such that $k_{i-1} + k_i$ is odd. Note that Proposition 6 remains true for the ‘starred sum’ except that the constant (of which we care little) is different.

We again commence with \nearrow - and \searrow -pre-chains. Little changes for a \searrow chain, since the sum remains finite. The only possible challenge to boundedness may originate in a \nearrow chain. We analyse a \nearrow -pre-chain using (3.10),

$$\sum_{k_\ell=k_{\ell-1}+1}^{\infty} \star \sum_{k_{\ell+1}=k_\ell+1}^{\infty} \star \cdots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty} \star |\mathcal{D}_{k_\ell,k_{\ell-1}} \mathcal{D}_{k_{\ell+1},k_\ell} \cdots \mathcal{D}_{k_{\ell+r-1},k_{\ell+r-2}}|$$

$$\begin{aligned}
 &= \sum_{k_\ell=k_{\ell-1}+1}^{\infty^*} \sum_{k_{\ell+1}=k_\ell+1}^{\infty^*} \cdots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty^*} \prod_{j=\ell}^{\ell+r-2} a_{k_j} b_{k_{j-1}} \\
 &= b_{k_{\ell-1}} \sum_{k_\ell=k_{\ell-1}+1}^{\infty^*} \left(k_\ell + \alpha + \frac{1}{2}\right) \sum_{k_{\ell+1}=k_\ell+1}^{\infty^*} \left(k_{\ell+1} + \alpha + \frac{1}{2}\right) \cdots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty^*} a_{k_{\ell+r-1}},
 \end{aligned}$$

and, using Proposition 6, transition seamlessly to a \nearrow -chain, while disregarding lower-order terms,

$$\begin{aligned}
 &\sum_{k_\ell=k_{\ell-1}+1}^{\infty^*} \left(k_\ell + \alpha + \frac{1}{2}\right) \sum_{k_{\ell+1}=k_\ell+1}^{\infty^*} \left(k_{\ell+1} + \alpha + \frac{1}{2}\right) \cdots \sum_{k_{\ell+r-1}=k_{\ell+r-2}+1}^{\infty^*} a_{k_{\ell+r-1}}^2 \\
 &\sim \sum_{k_{\ell+r-1}=k_{\ell-1}+r}^{\infty^*} a_{k_{\ell+r-1}}^2 \sum_{k_\ell=k_{\ell-1}+r-1}^{k_{\ell+r-1}-r+1} k_\ell \sum_{k_{\ell+1}=k_{\ell-1}+r-2}^{k_{\ell+r-1}-r+2} k_{\ell+1} \cdots \sum_{k_{\ell+r-2}=k_{\ell-1}+1}^{k_{\ell+r-1}-1} k_{\ell+r-2} \\
 &\sim \sum_{k_{\ell+r-1}=k_{\ell-1}+r}^{\infty^*} a_{k_{\ell+r-1}}^2 Q_{2r-2}(k_{\ell+r-1}) \sim \sum_{k=k_{\ell-1}+r}^{\infty^*} \frac{1}{k^{2\alpha-2r+1}}.
 \end{aligned}$$

Consequently, $\alpha > r$ is necessary and sufficient for convergence for each \nearrow -chain. Since the length of an \nearrow -chain is at most $s - 1$, we deduce, similar to Theorem 7, that

Theorem 8 *ind $w_\alpha \geq s$ for the ultraspherical weight if and only if $\alpha > s - 1$.*

Both Theorems 7 and 8 present the same inequality. This is not surprising since, for both weights, α measures the ‘strength’ of a zero at the endpoint(s).

We conclude this subsection with plots displaying the ℓ_2 norms of truncated $N \times N$ differentiation matrices for both Laguerre and ultraspherical cases for $\alpha \in \{1, 2, 3, 4\}$.

As can be seen from Fig. 3, the ℓ_2 norm of an $N \times N$ principal minor of \mathcal{D} increases linearly with N for Laguerre, quadratically for the ultraspherical weight. This has obvious implications, *inter alia* to the implementation of Krylov subspace methods in the manipulation of differentiation matrices. Having said so, Krylov subspace methods may be problematic in this context [10]. In Sect. 4, we outline alternative approaches to numerical algebra of separable differentiation matrices.

3.3 Counterexamples: generalised Hermite and Konoplev weights

Ultraspherical and Laguerre weights are the obvious and most elementary choice in the intervals $(-1, 1)$ and $(0, \infty)$, respectively, and they are both separable in the sense of this paper. This might lead to an impression that separability is ubiquitous: this would be highly misleading.

Lemma 9 *Let*

$$\iota_{m,n} = \mathcal{D}_{m,n} \mathcal{D}_{m+1,n+1} - \mathcal{D}_{m+1,n} \mathcal{D}_{m,n+1}, \tag{3.11}$$

$$\check{\iota}_{m,n} = \mathcal{D}_{m,n} \mathcal{D}_{m+2,n+2} - \mathcal{D}_{m+2,n} \mathcal{D}_{m,n+2}. \tag{3.12}$$

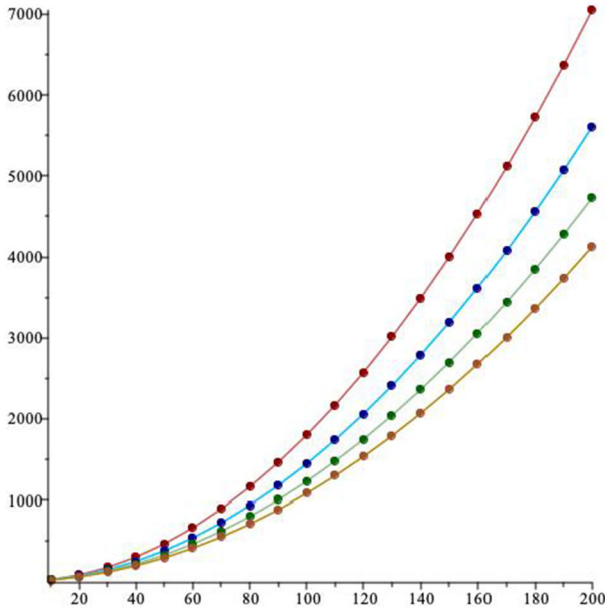


Fig. 3 l_2 norms of $N \times N$ differentiations matrices as N grows in multiples of ten. The top curve corresponds to $\alpha = 1$, underneath $\alpha = 2$, then $\alpha = 3$, and the bottom curve corresponds to $\alpha = 4$

Separability implies that $t_{m,n} = 0$ for all $m \geq n + 2$, while symmetric separability implies that $\check{t}_{m,n} = 0$ for all $m + n$ odd, $m \geq n + 2$.

Proof Follows at once from the definition of (symmetric) separability. □

Note that neither (3.11) nor (3.12) are sufficient. Thus, a skew-symmetric \mathcal{D} such that $\mathcal{D}_{2m+1,n} = 0$ for all $m \in \mathbb{Z}_+$ and $n \leq 2m - 1$ satisfies (3.11), but in general is not separable. Likewise, a tridiagonal skew-symmetric matrix obeys (3.12) but is not symmetrically separable—this is the case with the differentiation matrix associated with the Hermite weight, for example. Trying weights at random and computing, say, $t_{2,0}$ leads time and again to non-separable weights.

To explore further the (non)existence of separable weights, we examine two weights, generalisations of Hermite and ultraspherical weights, respectively, but endowed with an additional parameter: the *generalised Hermite* and *Konoplev* weights.

3.3.1 Generalised Hermite weights

Letting $\mu > -\frac{1}{2}$, we examine the weight

$$w_\mu(x) = |x|^{2\mu} e^{-x^2}, \quad x \in \mathbb{R} \tag{3.13}$$

[5, p. 156], originally considered by Szegő.⁶ It can be easily deduced from Chihara [5, p. 156–7] that the underlying W-functions are

$$\begin{aligned} \varphi_{2n}(x) &= (-1)^n 2^n \sqrt{\frac{n!}{\Gamma(n + \mu + \frac{1}{2})}} L_n^{(\mu - \frac{1}{2})}(x^2) |x|^\mu e^{-x^2/2}, \\ \varphi_{2n+1}(x) &= (-1)^n 2^{n+1} \sqrt{\frac{n!}{2\Gamma(n + \mu + \frac{3}{2})}} x L_n^{(\mu + \frac{1}{2})}(x^2) |x|^\mu e^{-x^2/2}, \quad n \in \mathbb{Z}_+. \end{aligned}$$

Generalised Hermite weights are of marginal importance to the work of this paper, and although their differentiation matrix can be derived explicitly,

$$\begin{aligned} \mathcal{D}_{2n,2n-1} &= \sqrt{n}, \quad \mathcal{D}_{2m,2n-1} = 0, \quad m \geq n + 1, \\ \mathcal{D}_{2n+1,2n} &= \frac{2n + 1}{2\sqrt{n + \mu + \frac{1}{2}}}, \quad \mathcal{D}_{2m+1,2n} = (-1)^{m+n-1} \mu \sqrt{\frac{m!}{(n + \mu + \frac{1}{2})_{m+1-n}}}, \end{aligned}$$

where $(z)_n = z(z + 1) \cdots (z + n - 1)$ is the Pochhammer symbol, with skew-symmetric complement, we will not present here a formal (and lengthy) algebra. Instead, a reader might use a symbolic algebra package to compute the first few elements, enough to evaluate $\iota_{2,0}$ and $\check{\iota}_{3,0}$ and check that they are both nonzero—in light of Lemma 9 this is sufficient to rule out separability and symmetric separability, respectively.

As a matter of fact, \mathcal{D} has an interesting shape: its $(2m + 1)$ st columns (hence also the $(2n + 1)$ st rows) are consistent with a tridiagonal matrix, more specifically with the differentiation matrix corresponding to the standard Hermite weight (i.e. with $\mu = 0$). More specifically,

$$\iota_{2n,2n-1} = \left(n + \frac{1}{2}\right) \sqrt{\frac{n}{n + \mu + \frac{1}{2}}} \neq 0, \quad \iota_{2n+1,2n} = \left(n + \frac{1}{2}\right) \sqrt{\frac{n + 1}{n + \mu + \frac{1}{2}}} \neq 0,$$

otherwise $\iota_{m,n} = 0$ for $m \geq n + 2$, while

$$\begin{aligned} \check{\iota}_{2n+3,2n} &= \frac{\mu \sqrt{(n + 2)!} [\mu \sqrt{(n + 1)!} + n + \frac{3}{2}]}{(n + \mu + \frac{3}{2}) \sqrt{(n + \mu + \frac{1}{2})(n + \mu + \frac{5}{2})}} \neq 0, \\ \check{\iota}_{2m+4,2n+1} &= 0, \quad m \geq n + 2 \end{aligned}$$

In each case, the separability tests (3.11) and (3.12) fail only marginally—but fail nonetheless.

⁶ We resist calling them “Szegő polynomials” since the name is reserved for another type of polynomials, orthogonal in the complex unit circle [24, 26].

3.3.2 Konoplev weights

Letting $\alpha, \gamma > -1$, we set

$$w_{\alpha,\gamma}(x) = |x|^{2\gamma+1}(1-x^2)^\alpha, \quad x \in (-1, 1). \tag{3.14}$$

The weight (3.14), which has been considered in [16, 17] and described in Chihara [5, p. 155], generalises ultraspherical weights by adding the possible weakly singular factor $|x|^{2\gamma+1}$. Specifically, $w_{\alpha,\gamma} \in C^s(-1, 1)$ if and only if

$$\text{either } \gamma \in \left\{ \frac{k}{2} : k \in \{-1, 0, \dots, s-2\} \right\} \quad \text{or} \quad \gamma > \frac{s-1}{2}.$$

The underlying orthogonal polynomial system is

$$S_{2n}(x) = P_n^{(\alpha,\gamma)}(2x^2 - 1), \quad S_{2n+1}(x) = xP_n^{(\alpha,\gamma+1)}(2x^2 - 1), \quad n \in \mathbb{Z}_+,$$

and the monic polynomials obey the three-term recurrence relation

$$\hat{S}_{n+1}(x) = x\hat{S}_n(x) - c_n\hat{S}_{n-1}(x),$$

where

$$c_{2n} = \frac{n(n+\alpha)}{(2n+\alpha+\gamma)(2n+1+\alpha+\gamma)}, \quad c_{2n+1} = \frac{(n+1+\gamma)(n+1+\alpha+\gamma)}{(2n+\alpha+\gamma)(2n+1+\alpha+\gamma)}.$$

Replacing Jacobi polynomials by their orthonormal counterparts and using a formula from Rainville [23, p. 260], easy algebra confirms that

$$\begin{aligned} \kappa_{2m}^{\alpha,\gamma} &= \int_{-1}^1 w_{\alpha,\gamma}(x) S_{2m}^2(x) dx = \frac{(m+1+\alpha+\gamma)_m \Gamma(m+1+\alpha) \Gamma(m+1+\gamma)}{m! \Gamma(2m+2+\alpha+\gamma)}, \\ \kappa_{2m+1}^{\alpha,\gamma} &= \int_{-1}^1 w_{\alpha,\gamma}(x) S_{2m+1}^2(x) dx = \frac{(m+2+\alpha+\gamma)_m \Gamma(m+1+\alpha) \Gamma(m+2+\gamma)}{m! \Gamma(2m+3+\alpha+\gamma)}, \end{aligned}$$

therefore

$$\begin{aligned} \varphi_{2m}(x) &= \frac{|x|^{\gamma+\frac{1}{2}}(1-x^2)^{\alpha/2}}{\sqrt{\kappa_{2m}^{\alpha,\gamma}}} P_m^{(\alpha,\gamma)}(2x^2 - 1), \\ \varphi_{2m+1}(x) &= \frac{x|x|^{\gamma+\frac{1}{2}}(1-x^2)^{\alpha/2}}{\sqrt{\kappa_{2m+1}^{\alpha,\gamma}}} P_m^{(\alpha,\gamma+1)}(2x^2 - 1). \end{aligned}$$

The weights (3.14) are symmetric; thus, we examine the possibility of symmetric separability. A brute force computation yields

$$\check{\iota}_{3,0} = (5 + \alpha + \gamma)(2\gamma + 1) \sqrt{\frac{(4 + \alpha + \gamma)(6 + \alpha + \gamma)}{2(1 + \alpha)(2 + \alpha)(1 + \gamma)(3 + \gamma)}},$$

ruling out symmetric separability except for the case $\gamma = -\frac{1}{2}$, which corresponds to the ultraspherical weight.

3.3.3 A Limiting Behaviour of the $\iota_{m,n}$ s

While separability, hence $\iota_{m,n} = 0$ for $m \geq n + 2$, appears to be exceedingly rare, we claim that the latter holds more broadly in a much weaker, asymptotic form.

Let w be a weight in (a, b) , $w(a) = w(b) = 0$, with the underlying orthonormal polynomials $\{p_n\}_{n=0}^\infty$, where the coefficient of x^n in p_n is $k_n > 0$. Comparing the coefficients of x^{n+1} in the three-term recurrence relation (1.3), we deduce at once that $k_{n+1}/k_n = \beta_n^{-1}$.

Theorem 10 *Assuming that $\beta_n \geq \beta^* > 0$ and $\tilde{w}(x) = [w'(x)]^2/w(x)$ is itself a weight function in (a, b) , it is true that*

$$\lim_{m \rightarrow \infty} \iota_{m,n} = 0, \quad n \in \mathbb{Z}_+. \tag{3.15}$$

Proof Letting $m \geq n + 2$, (2.3) yields

$$\begin{aligned} \iota_{m,n} &= \frac{1}{4} \int_a^b w'(y) p_m(y) p_n(y) dy \int_a^b w'(x) p_{m+1}(x) p_{n+1}(x) dx \\ &\quad - \frac{1}{4} \int_a^b w'(x) p_{m+1}(x) p_n(x) dx \int_a^b w'(y) p_m(y) p_{n+1}(y) dy \\ &= \frac{1}{4} \int_a^b \int_a^b w'(x) w'(y) p_{m+1}(x) p_m(y) [p_{n+1}(x) p_n(y) - p_n(x) p_{n+1}(y)] dx dy. \end{aligned}$$

We recall the *Christoffel–Darboux formula*,

$$\sum_{\ell=0}^n p_\ell(x) p_\ell(y) = \frac{k_n}{k_{n+1}} \frac{p_{n+1}(x) p_n(y) - p_n(x) p_{n+1}(y)}{x - y},$$

where $k_n > 0$ is the coefficient of x^n in p_n [5, p. 153]. Therefore,

$$\begin{aligned} \iota_{m,n} &= \frac{k_{n+1}}{k_n} \int_a^b \int_a^b w'(x) w'(y) p_{m+1}(x) p_m(y) (x - y) \sum_{\ell=0}^n p_\ell(x) p_\ell(y) dx dy \\ &\leq \frac{1}{\beta^*} \left| \int_a^b \int_a^b w'(x) w'(y) p_{m+1}(x) p_m(y) (x - y) \sum_{\ell=0}^n p_\ell(x) p_\ell(y) dx dy \right|, \tag{3.16} \end{aligned}$$

because $\beta_n^{-1} \leq \beta^{*-1}$. Letting $n \rightarrow \infty$ in (3.16), we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} t_{m,n} &\leq \frac{1}{\beta^*} \left| \int_a^b \int_a^b w'(x)w'(y)p_{m+1}(x)p_m(x)(x-y) \sum_{\ell=0}^{\infty} p_\ell(x)p_\ell(y) dx dy \right| \\ &= \frac{1}{\beta^*} \left| \int_a^b \int_a^b \frac{w'(x)w'(y)}{\sqrt{w'(x)w'(y)}} p_{m+1}(x)p_m(x)(x-y)K(x,y) dx dy \right|, \end{aligned}$$

where K is the Christoffel–Darbeaux kernel from the proof of Theorem 5. According to (2.8), it is a reproducing kernel and it follows at once that the double integral vanishes. \square

The condition $\beta_n > \beta^*$, $n \in \mathbb{Z}_+$, is very weak: we already know that $\beta_n > 0$, all the condition says is that, in addition, the β_n s are bounded away from zero.

4 Computational Aspects

4.1 A Product of \mathcal{D} by a Vector

Practical implementation of the ideas of this paper requires manipulation of expressions involving a matrix \mathcal{D} which is either separable or symmetrically separable: formation of products of the form $\mathcal{D}^r \mathbf{f}$ for $r \in \mathbb{N}$, the solution of algebraic linear systems of the form $p(\mathcal{D})\mathbf{y} = \mathbf{x}$, where p is a polynomial and the computation of $e^{h\mathcal{D}}\mathbf{u}$. Tridiagonal differentiation matrices, of the form considered in [11], enjoy substantial advantage in this context. Yet, once a weight function is separable (or symmetrically separable), all these objectives can be attained using fast algorithms. The matrix \mathcal{D} is a special case of a *semiseparable matrix* [7, p. 50]: all its minors located either wholly above the diagonal or wholly beneath it are of rank 1. This allows for fast products and fast computation of linear systems [6, 7, 28] and, using the Cauchy–Dunford integral formula (also known as the Dunford–Taylor or Riesz–Fantappi  formula, [4]), compute $e^{h\mathcal{D}}\mathbf{u}$.

In this subsection, we examine in detail the formation of products of the form $\mathbf{h} = \mathcal{D}\mathbf{f}$, where \mathbf{f} is a (real or complex) infinite-dimensional vector, while \mathcal{D} is either separable or symmetrically separable. While the main idea is not new—cf. for example, [6]—there is merit in presenting it for the convenience of the reader, as an elementary example of more advanced numerical algebra computations in [6, 7, 28]. As a matter of fact, our algorithm is somewhat more general, because it is based on infinite-dimensional computations.

Consider a separable weight function, e.g. a Laguerre weight. The starting point is an integer N , typically much larger than M , such that $|f_m|$ is negligible (in practical terms, smaller than a user-provided error tolerance) for $m > N$, and we wish to form

$$h_m = \sum_{n=0}^N \mathcal{D}_{m,n} f_n, \quad m = 0, \dots, M. \tag{4.1}$$

We commence by assuming that a weight is separable, whereby (3.1) yields

$$h_m = -\mathbf{a}_m \sum_{n=0}^{m-1} \mathbf{b}_n f_n + \mathbf{b}_m \sum_{n=m+1}^N \mathbf{a}_n f_n = \sigma_m + \rho_m, \quad m = 0, \dots, M,$$

where

$$\sigma_m = \sum_{n=0}^{m-1} \mathbf{b}_n f_n, \quad \rho_m = \sum_{n=m+1}^N \mathbf{a}_n f_n, \quad m = 0, \dots, M.$$

Then

$$\begin{aligned} h_0 &= \mathbf{b}_0 \rho_0, \\ h_m &= -\mathbf{a}_m \sigma_m + \mathbf{b}_m \rho_m, \quad m = 1, \dots, M, \\ &\text{where } \sigma_m = \sigma_{m-1} + \mathbf{b}_{m-1} f_{m-1}, \quad \rho_m = \rho_{m-1} - \mathbf{a}_m f_m. \end{aligned}$$

Assuming that the \mathbf{a}_m s and \mathbf{b}_n s have been precomputed (and this needs to be done only once, no matter how many products are required), the calculation (4.1) takes just $\approx N + 4M$ flops—and by the same token, computing the first $M + 1$ entries of $\mathcal{D}_N^r \mathbf{f}_N$ takes $\approx r(N + 4M)$ flops.

Similar operations count applies to symmetrically separable weight, whereby the entries of \mathcal{D} obey (3.2). Assuming that both M and N are even, we have

$$\begin{aligned} h_{2m} &= \sum_{n=0}^{N/2} \mathcal{D}_{2m, 2n+1} f_{2n+1}, \\ h_{2m+1} &= \sum_{n=0}^{N/2} \mathcal{D}_{2m+1, 2n} f_{2n}, \quad m = 0, \dots, \frac{M}{2}. \end{aligned}$$

Therefore

$$\begin{aligned} h_{2m} &= \mathbf{a}_{2m} \sum_{n=0}^{m-1} \mathbf{b}_{2n+1} f_{2n+1} - \mathbf{b}_{2m} \sum_{n=m+1}^{N/2} \mathbf{a}_{2n+1} f_{2n+1}, \\ h_{2m+1} &= \mathbf{a}_{2m+1} \sum_{n=0}^{m-1} \mathbf{b}_{2n} f_{2n} - \mathbf{b}_{2m+1} \sum_{n=m+1}^{N/2} \mathbf{a}_{2n} f_{2n}, \quad m = 0, \dots, \frac{M}{2}. \end{aligned}$$

Set

$$\sigma_m^E = \sum_{n=0}^{m-1} \mathbf{b}_{2n} f_{2n}, \quad \sigma_m^O = \sum_{n=0}^{m-1} \mathbf{b}_{2n+1} f_{2n+1},$$

$$\rho_m^E = \sum_{n=m+1}^{N/2} \alpha_{2n} f_{2n}, \quad \rho_m^O = \sum_{n=m+1}^{N/2} \alpha_{2n+1} f_{2n+1},$$

hence

$$h_{2m} = \alpha_{2m} \sigma_m^O - \mathfrak{b}_{2m} \rho_m^O, \quad h_{2m+1} = \alpha_{2m+1} \sigma_m^E - \mathfrak{b}_{2m+1} \rho_m^E, \quad m = 0, \dots, \frac{M}{2}.$$

However,

$$\sigma_0^E = \sigma_0^O = 0, \quad \rho_0^E = \sum_{m=1}^{N/2} \alpha_{2m} f_{2m}, \quad \rho_0^O = \sum_{m=1}^{N/2} \alpha_{2m+1} f_{2m+1},$$

and

$$\begin{aligned} \sigma_m^E &= \sigma_{m-1}^E + \mathfrak{b}_{2m-2} f_{2m-2}, & \sigma_m^O &= \sigma_{m-1}^O + \mathfrak{b}_{2m-1} f_{2m-1}, \\ \rho_m^E &= \rho_{m-1}^E - \alpha_{2m} f_{2m}, & \rho_m^O &= \rho_{m-1}^O - \alpha_{2m+1} f_{2m+1}, \end{aligned} \quad m = 1, \dots, \frac{M}{2}.$$

Thus, again, we need just $\approx 4M + N$ flops to compute the first $M + 1$ terms of $\mathcal{D}_N f_N$.

4.2 The Speed of Convergence

The convergence of orthogonal polynomials to ‘nice’ (in particular, analytic) functions is well understood, and this can be leveraged to the case of W-functions. It is beneficial first, though, to present some computational results, to highlight the importance of choosing the right value of α in the context of either Laguerre or ultraspherical weights, while comparing them to standard approximation by the underlying orthogonal polynomials.

It rapidly becomes apparent that we have a competition between different imperatives:

- *The number of zero boundary conditions:* This determines the value of α and, according to Theorems 7 and 8, we need $\alpha > s - 1$ in $\mathring{H}_2^s(a, b)$.
- *Regularity of approximating functions:* While \mathcal{P} consists of polynomials, hence analytic functions, this is not the case with Φ , whether in the context of ultraspherical or Laguerre weights: it all depends on the value of α . If α is an even integer, then the φ_n s are analytic; otherwise, analyticity fails at the endpoints.
- *The underlying function space:* Much depends on how the error is measured. Among the many possibilities, we single out two: the $\mathring{H}_2^p(a, b)$ norm for a suitable value of p (in particular, the $L_2(a, b)$ norm) and the $L_\infty[a, b]$ (and, more generally, $H_\infty^p[a, b]$) norm. The choice of a norm depends on the underlying application.

Preliminary numerical experimentation reveals a remarkable state of affairs. In Fig. 4, we let α be in $\{1, 2, 3, 4\}$. In this and all figures in this paper, we denote $\alpha = 1$ by a red, dotted line, $\alpha = 2$ by a magenta solid line, $\alpha = 3$ by a green dashed line and, finally, $\alpha = 4$ by a blue dash-dotted line. Because of the rapid decay of errors, we display

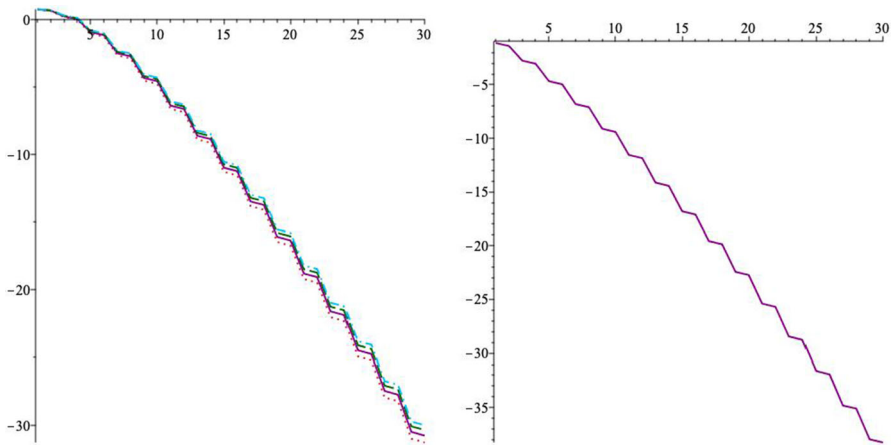


Fig. 4 Ultraspherical W-functions: The errors $\log_{10} \|F_N^P - f\|_\infty$ (top left), $\log_{10} \|^dF_N^P - f'\|_\infty$ (top right), $\log_{10} \|F_N^\Phi - f\|_\infty$ (bottom left) and $\log_{10} \|^dF_N^\Phi - f'\|_\infty$ (bottom right) for $\alpha = 1, 2, 3, 4$ and $N = 1, 2, \dots, 30$ (except that in the bottom-right plot only $\alpha = 2$ is displayed)

them all in a logarithmic scale to base 10—in other words, the y-axis displays the number of decimal digits. Given a function f and recalling the expansion coefficients \hat{f}_n^P and \hat{f}_n^Φ from Remark 1, corresponding to expansions in \mathcal{P} and Φ , respectively, we let

$$F_N^P(x) = \sum_{n=0}^N \hat{f}_n^P p_n(x), \quad F_N^\Phi(x) = \sum_{n=0}^N \hat{f}_n^\Phi \varphi_n(x), \quad N \in \mathbb{Z}_+.$$

Thus, $F_N^P - f$ and $F_N^\Phi - f$ are the (pointwise) errors with respect to the polynomial and the W-function basis, respectively, and we need to measure them in an appropriate norm. We denote by $^dF_N^P$ the derivative expansion, i.e. with p_n and f replaced by p'_n and f' , respectively, similarly for higher derivatives and for F_N^Φ .

4.2.1 Ultraspherical W-Functions

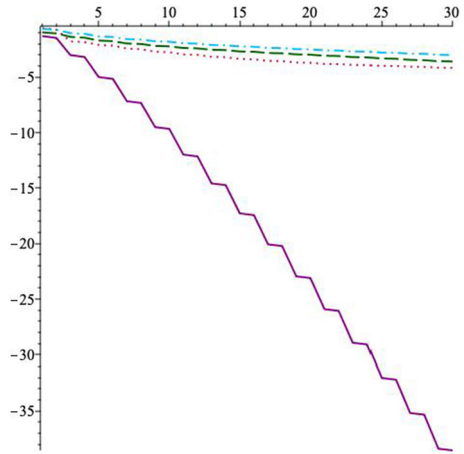
We commence from ultraspherical weights and consider

$$f(x) = (1 - 2x) \cos \frac{\pi x}{2} \in \mathring{H}_2^1(-1, 1). \tag{4.2}$$

In Fig. 4, we display in logarithmic norm the $L_\infty[-1, 1]$ error for polynomial approximation to f and its first derivative (top row) and for W-functions for the ultraspherical weight.⁷ Polynomial approximation—as can be expected from general theory and the analyticity of f —decays at an exponential speed and, for $N = 30$, we attain ≈ 32 significant digits. This is also the case with derivatives, with a very minor

⁷ Polynomial approximation, of course, leads to an unstable spectral method. Yet, its error and its comparison with the error committed by W-functions are of independent interest.

Fig. 5 Ultraspherical
 W-functions: The errors
 $\log_{10} \|F_N^P - f\|_2$ (left) and
 $\log_{10} \|F_N^\Phi - f\|_2$ (right) for
 $\alpha = 1, 2, 3, 4$



degradation in accuracy. The error for W-functions, though, is radically different. The errors for $\alpha \in \{1, 3, 4\}$ decay very slowly, at a polynomial rate, and for $N = 30$ we recover just ≈ 4 significant digits, an unacceptably large error. On the other hand, the error for $\alpha = 2$ at $N = 30$ is $\approx 3 \times 10^{-39}$, significantly better than polynomial approximation!

The reason for this ‘miraculous’ behaviour for $\alpha = 2$ bears some attention. Little surprise perhaps that $\alpha = 1$ behaves poorly because it is at the wrong end of the boundedness condition for \mathcal{D}^2 . However, as a matter of fact, we do not consider second derivatives in this particular instance and $\alpha \in \{3, 4\}$ are just as bad. The reasons are as follows. For $\alpha \in \{1, 3\}$, the φ_n s have a weak singularity along the boundary, while φ'_n becomes singular there. For $\alpha = 4$, on the other hand, $\varphi'_n(\pm 1) = 0$ mean that L_∞ convergence of derivatives is impossible unless also the derivatives of f vanish at the endpoints. (This is the reason why $\log_{10} \|F_N^\Phi - f'\|_\infty$ is displayed only for $\alpha = 2$.)

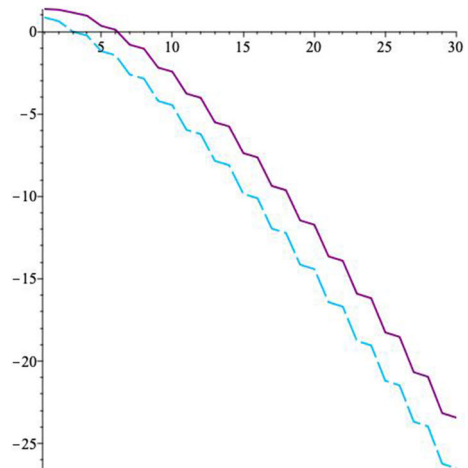
Not much changes if, instead of L_∞ , we compute an \dot{H}_2^p error, except that in general L_2 -like norms are more forgiving. In principle, neither singularities or excessive vanishing of derivatives at the endpoints need prevent convergence. Thus, in Fig. 5 we plot the $L_2(-1, 1)$ errors for example (4.2). The overall picture remains the same: polynomial approximation decays at exponential rate and we attain, regardless of the choice of α , about 34 significant digits for $N = 30$, while W-function approximation for $\alpha \in \{1, 3, 4\}$ is very poor yet, for $\alpha = 2$, we again hit the ‘sweet spot’ and recover ≈ 38 significant digits. W-functions are vastly superior for $\alpha = 2$ but fail dismally otherwise.

To explore further the error committed by ultraspherical W-functions, we consider

$$f(x) = (1 - 2x) \cos^2 \frac{\pi x}{2} \tag{4.3}$$

the only difference in this (not very imaginative!) choice is that now $f(\pm 1) = f'(\pm 1) = 0$. We display the L_∞ error for $f^{(i)}$, $i = 0, 1, 2$, in Fig. 6 for the W-functions. The error in polynomial approximation is roughly independent of α , and

Fig. 6 Ultraspherical W-functions: The errors $\log_{10} \|F_N^\Phi - f\|_2$ (left) $\log_{10} \|{}^d F_N^\Phi - f'\|_2$ (centre) and $\log_{10} \|{}^{dd} F_N^\Phi - f''\|_2$ (right) for $\alpha = 1, 2, 3, 4$ and the function (4.3)



for $N = 30$ we attain ≈ 24 decimal digits for f , ≈ 21 for f' and ≈ 19 for f'' . By this stage, we should not be surprised that $\alpha = 1$ and $\alpha = 3$ do badly in approximating f because of the weak singularity at the endpoints and they fail altogether approximating derivatives. For $\alpha \in \{2, 4\}$, the endpoints are analytic and indeed the underlying functions do very well indeed, definitely better than polynomial approximation. $\alpha = 4$ is a winner, unsurprisingly because $f \in \overset{\circ}{H}_2^2(-1, 1)$ and this is matched by Φ . However, $\alpha = 2$ does quite well, worse by perhaps two decimal digits but still beating polynomial approximation. The reason is that too few zero Dirichlet boundary conditions do not prevent L_∞ convergence of an orthogonal sequence, although they might slow it up to a modest extent. On the other hand, excessive zero Dirichlet boundary conditions prevent L_∞ convergence at the endpoints. Thus, *the interplay between the number of zero boundary conditions and the choice of α is not symmetric!* It is always better to err by choosing smaller α , as long as it is an even integer, consistent with the bound of Theorem 8.

4.2.2 Laguerre W-Functions

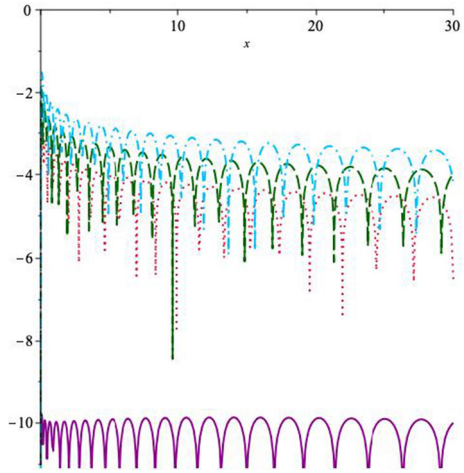
We are now concerned with the Laguerre weight and choose the model problem

$$f(x) = e^{-x} \sin x, \quad x \geq 0. \tag{4.4}$$

Note that $f(0) = 0, f'(0) \neq 0$.

An expansion in Laguerre (or any other) polynomials cannot be bounded in an infinite interval; hence, instead of plotting $\log_{10} \|F_N^P - f\|_2$ for increasing values of N , we choose $N = 40$ and plot the pointwise error in the interval $[0, 30]$. This is evident on the left of Fig. 7: the error is just about fine for small $x > 0$, subsequently growing rapidly (as a matter of fact, exponentially). On the other hand, as can be seen on the right of that figure, the error of W-functions is uniformly bounded. For $\alpha \in \{1, 3, 4\}$, it is fairly similar—and unacceptably large—while for $\alpha = 2$ we attain

Fig. 7 Laguerre W-functions:
 The errors
 $\log_{10} |F_{40}^P(x) - f(x)|$ (left) and
 $\log_{10} |F_{40}^\Phi(x) - f(x)|$ (right) for
 $x \in [0, 30]$ and $\alpha = 1, 2, 3, 4$



≈ 10 decimal digits of accuracy, apparently uniformly in $[0, \infty)$. Yet again we have the ‘sweet spot’ for $\alpha = 2$. This state of affairs remains true for the first few derivatives, and the deterioration in accuracy using W-functions is very mild indeed.

Finally, we consider

$$f(x) = e^{-x} \sin^2 x, \quad x \geq 0. \tag{4.5}$$

Now $f(0) = f'(0) = 0$ and $f''(0) \neq 0$. There is no need to display the $L_\infty[0, \infty)$ error committed by Laguerre polynomials since, again, it is unbounded.

In Fig. 8, we employ the same colour and style scheme to plot the errors committed in $[0, 30]$ for f, f' and f'' . Clearly, $\alpha = 2$ and $\alpha = 4$, the two values associated with analyticity at the origin, win insofar as approximating the function itself is concerned, although the margin is somewhat smaller than in our other examples. The approximation of the first and the second derivatives is more interesting: on the face of it, it is a dead heat between $\alpha = 2$ and $\alpha = 4$, but closer examination of the behaviour near the left endpoint unravels a crucial difference. For example, for $\eta = 10^{-10}$ we have (to four significant digits)

α	1	2	3	4
$ F_{60}^\Phi(\eta) - f(\eta) $	2.128 ₋₀₈	9.631 ₋₁₅	1.434 ₋₁₆	7.778 ₋₂₄
$ {}^dF_{60}^\Phi(\eta) - f'(\eta) $	1.064 ₊₀₂	9.631 ₋₀₅	2.151 ₋₀₆	9.555 ₋₁₄
$ {}^{dd}F_{60}^\Phi(\eta) - f''(\eta) $	5.319 ₊₁₁	4.092 ₋₀₃	1.075 ₊₀₅	9.555 ₋₀₄

The conclusion is clear. Once the inequality of Theorem 7 is breached, the approximation blows up at the origin: this happens with $\alpha = 1$ and *any* derivative. The error for $\alpha = 3$ decays for $N \gg 1$ for the function value and the first derivative, but it blows up for the second derivative, while for $\alpha = 2$ the progression to the correct boundary condition is considerably slower than for $\alpha = 4$. This is apparent from Fig. 9: $\alpha = 4$ wins, although by a small margin.

Fig. 8 Laguerre W-functions:
 The errors
 $\log_{10} |F_{60}^\Phi(x) - f(x)|$ (left)
 $\log_{10} |{}^d F_{60}^\Phi(x) - f'(x)|$ (right)
 and $\log_{10} |{}^{dd} F_{60}^\Phi(x) - f''(x)|$
 for the function (4.5),
 $x \in [0, 30]$ and $\alpha = 1, 2, 3, 4$

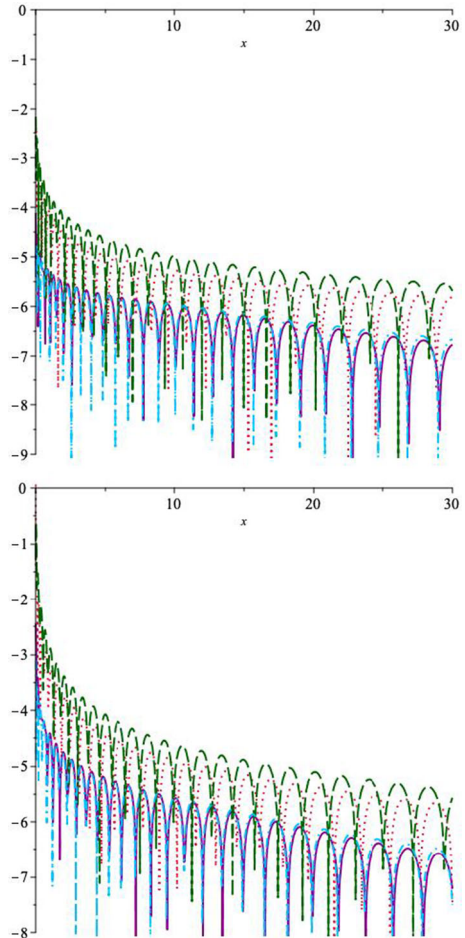
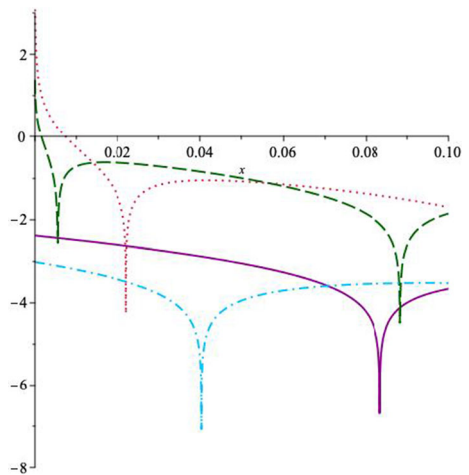


Fig. 9 Laguerre W-functions: A
 close-up of the bottom plot in
 Fig. 8 near the left endpoint



4.2.3 The Speed of Convergence Redux

As promised, we leverage standard theory on the convergence of orthogonal expansions of analytic functions to the setting of W-functions. For simplicity, we consider just the ultraspherical weight, supported in $[-1, 1]$, but our argument readily extends to all other settings. Thus, suppose that the function f is analytic in a Bernstein ellipse enveloping $[-1, 1]$ and $f(\pm 1) = 0$ and let $\tilde{f}(x) = f(x)/(1 - x^2)$. Note that \tilde{f} is also analytic in $[-1, 1]$. Expanding f in powers of an ultraspherical W-function yields the coefficients

$$\hat{f}_n = \int_{-1}^1 (1 - x^2)^{\alpha/2} f(x) P_n^{(\alpha, \alpha)}(x) dx = \int_{-1}^1 (1 - x^2)^{\alpha/2+1} \tilde{f}(x) P_n^{(\alpha, \alpha)}(x) dx, \quad n \in \mathbb{Z}_+$$

and, once $\alpha = 2$, we have standard expansion of \tilde{f} in the ultraspherical polynomial basis $\{P_n^{(2,2)}\}_{n \in \mathbb{Z}_+}$ and standard results for expansions in orthogonal polynomials (cf. for example [29]) apply. In particular,

$$\limsup_{n \rightarrow \infty} |\hat{f}_n|^{1/n} = \rho \in [0, 1)$$

and the rate of convergence in the L_2 norm is at least exponential. The same is true for any even integer $\alpha \geq 2$.

Similar reasoning applies to other weights and to higher-order zero boundary conditions.

Needless to say, choosing $\alpha \in 2\mathbb{N}$ is essential, but the ultimate choice depends on the number of zero boundary conditions, in the spirit of the previous two subsections.

4.3 Outstanding Computational and Theoretical Challenges

This is the first paper to consider W-functions in an organised way, although of course Hermite functions have been used and investigated extensively and W-functions associated with Freud weights (and which are special because of Theorem 2) have been introduced in [19]. Needless to say, this work neither resolves all the mathematical and computational issues associated with W-functions nor claims to do so. While there are important theoretical questions, e.g. to characterise all separable or symmetrically separable weight functions, perhaps the most urgent issues are related to the applications of W-functions to spectral methods. This concerns issues in approximation theory (speed of convergence in different function classes), as well as purely computational questions. The speed of approximation points out an imperfect duality between W-functions and the functions $\Psi = \{\psi_n\}_{n \in \mathbb{Z}_+}$ from Sect. 1. Recalling the \hat{f}_n^Φ , the n th expansion coefficient in \mathcal{P} and letting $\check{w}(x) = w(x)\chi_{(a,b)}(x)$, the Plancherel theorem yields at once for every $n \in \mathbb{Z}_+$

$$\hat{f}_n^\Phi = \int_a^b f(x)\phi_n(x) dx = \int_{-\infty}^{\infty} \sqrt{\check{w}(x)} f(x) p_n(x) dx = (-i)^n \int_{-\infty}^{\infty} \hat{f}(\xi) \overline{\psi_n(x)} dx,$$

and we recover an expansion in Ψ of the Fourier transform of f . This duality, though, is imperfect because, unless $(a, b) = \mathbb{R}$, it is valid (insofar as Ψ is concerned) only in the Paley–Wiener space $\mathcal{P}_{(a,b)}(\mathbb{R})$ rather than in $L_2(\mathbb{R})$ [11]. Moreover, comprehensive convergence theory for functions of the form Ψ is also lacking. Yet, even an imperfect duality might potentially lead to useful outcomes.

The final issue we wish to mention is fast computation of expansion coefficients in a W-function basis, similar perhaps to fast expansion algorithms in polynomial bases [22]. All this is a matter for future research.

Acknowledgements The author wishes to thank both referees, whose comprehensive and well-informed comments have led to a much better paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Separability Coefficients for Laguerre Weights

Our starting point is the generating function

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

[23, p. 202] and we recall that

$$\int_0^{\infty} [L_m^{(\alpha)}(x)]^2 x^\alpha e^{-x} dx = \frac{\Gamma(m+1+\alpha)}{m!} = \frac{\Gamma(1+\alpha)(1+\alpha)_m}{m!}. \tag{A.6}$$

Set

$$q_{m,n} = -\frac{1}{2} \int_0^{\infty} L_m^{(\alpha)}(x)L_n^{(\alpha)}(x) \frac{dx^\alpha e^{-x}}{dx} dx, \quad m, n \in \mathbb{Z}_+.$$

Because of

$$\tilde{p}_m(x) = \sqrt{\frac{m!}{\Gamma(m+1+\alpha)}} L_m^{(\alpha)}(x)$$

and (2.3), it follows that

$$\mathcal{D}_{m,n} = \sqrt{\frac{m!n!}{\Gamma(m+1+\alpha)\Gamma(n+1+\alpha)}} q_{m,n}, \quad m \geq n + 1.$$

We set

$$\begin{aligned}
 Q(s, t) &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} q_{m,n} s^m t^n \\
 &= -\frac{1}{2} \int_0^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} L_m^{(\alpha)}(x) L_n^{(\alpha)}(x) s^m t^n (-x^\alpha + \alpha x^{\alpha-1}) e^{-x} dx.
 \end{aligned}$$

Using (A.6) it follows that for $|s|, |t| < 1$,

$$\begin{aligned}
 Q(s, t) &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{\Gamma(m+1+\alpha)}{m!} (st)^m \\
 &\quad - \frac{\alpha}{2(1-s)^{1+\alpha}(1-t)^{1+\alpha}} \int_0^{\infty} x^{\alpha-1} \exp\left(-x + \frac{xt}{t-1} + \frac{xs}{s-1}\right) dx \\
 &= \frac{1}{2} \Gamma(1+\alpha) {}_1F_0\left[\begin{matrix} 1+\alpha; \\ -; \end{matrix} st\right] \\
 &\quad - \frac{\alpha}{2(1-s)^{1+\alpha}(1-t)^{1+\alpha}} \int_0^{\infty} x^{\alpha-1} \exp\left(-\frac{(1-ts)x}{(1-s)(1-t)}\right) dx \\
 &= \frac{\Gamma(1+\alpha)}{2} \frac{1}{(1-st)^{\alpha+1}} - \frac{\Gamma(1+\alpha)}{2(1-s)(1-t)(1-st)^\alpha}.
 \end{aligned}$$

(Cf. for example, [23] for the definition and basic facts on confluent hypergeometric functions.) We now expand: all it takes is elementary (but long) algebra:

$$\begin{aligned}
 \frac{1}{(1-st)^{1+\alpha}} &= \sum_{m=0}^{\infty} \frac{(1+\alpha)_m}{m!} (st)^m, \\
 \frac{1}{(1-s)(1-t)(1-st)^\alpha} &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \left[\sum_{k=0}^{\min\{m,n\}} \frac{(\alpha)_k}{k!} \right] s^m t^n.
 \end{aligned}$$

The following proposition can be trivially proved by induction.⁸

Proposition 11

$$\sum_{k=0}^j \frac{(\alpha)_k}{k!} = \frac{(1+\alpha)_j}{j!}.$$

We thus deduce from the definition of Q that $q_{n,n} = 0$ and

$$q_{m,n} = \frac{\Gamma(\alpha)}{2} \frac{(1+\alpha)_n}{n!}, \quad m \geq n + 1,$$

⁸ And it might well be already known.

and conclude that

$$\mathcal{D}_{m,n} = \frac{\Gamma(1 + \alpha)}{2} \sqrt{\frac{m!n!}{\Gamma(m + 1 + \alpha)\Gamma(n + 1 + \alpha)}} \frac{(1 + \alpha)_n}{n!} = \frac{1}{2} \sqrt{\frac{m!\Gamma(n + 1 + \alpha)}{\Gamma(m + 1 + \alpha)n!}} \tag{A.7}$$

for $m \geq n + 1$, with skew-symmetric completion for $m \leq n$.

We have just determined both separability and the explicit form of the sequences \mathfrak{a} and \mathfrak{b} .

Theorem 12 *The Laguerre weight is separable and*

$$\mathfrak{a}_m = \sqrt{\frac{m!}{2\Gamma(m + 1 + \alpha)}}, \quad \mathfrak{b}_n = \sqrt{\frac{\Gamma(n + 1 + \alpha)}{2n!}}, \quad m, n \in \mathbb{Z}_+. \tag{A.8}$$

B Symmetric Separability Coefficients for Ultraspherical Weights

We recall that

$$S_{m,n}^\alpha = \int_{-1}^1 (1 - x^2)^{\alpha-1} P_m^{(\alpha,\alpha)}(x) P_n^{(\alpha,\alpha)}(x) dx, \quad \alpha > 0,$$

and we are concerned with $m \geq n$ ($S_{m,n}$ is symmetric) and even $m + n$. We commence by dividing $P_n^{(\alpha,\alpha)}$ by $(1 - x^2)$ —it follows from the Euclidean algorithm that

$$P_n^{(\alpha,\alpha)}(x) = (1 - x^2)\gamma_n(x) + \delta_n(x),$$

where $\gamma_n \in \mathbb{P}_{n-1}$ and $\delta_n(x) = \delta_{n,0} + \delta_{n,1}x$ is linear. Because of parity, if n is even then $\delta_1 = 0$, while if it is odd, then $\delta_0 = 0$. The description of δ can be completed by considering $x = 1$,

$$\delta_{2n,0} = P_{2n}^{(\alpha,\alpha)}(1) = \frac{(1 + \alpha)_{2n}}{(2n)!}, \quad \delta_{2n+1,1} = P_{2n+1}^{(\alpha,\alpha)}(1) = \frac{(1 + \alpha)_{2n+1}}{(2n + 1)!};$$

therefore,

$$\delta_{2n} \equiv \frac{(1 + \alpha)_{2n}}{(2n)!}, \quad \delta_{2n+1}(x) = \frac{(1 + \alpha)_{2n+1}}{(2n + 1)!}x.$$

Since $\deg \gamma \leq n - 1 \leq m - 1$, it follows from orthogonality that

$$\begin{aligned} S_{m,n}^\alpha &= \int_{-1}^1 (1 - x^2)^{\alpha-1} P_m^{(\alpha,\alpha)}(x) [(1 - x^2)\gamma_n(x) + \delta_n(x)] dx \\ &= \int_{-1}^1 (1 - x^2)^{\alpha-1} P_m^{(\alpha,\alpha)}(x) \delta_n(x) dx. \end{aligned}$$

Letting

$$\epsilon_m = \int_{-1}^1 (1 - x^2)^{\alpha-1} P_{2m}^{(\alpha,\alpha)}(x) dx, \quad \mathfrak{o}_m = \int_{-1}^1 (1 - x^2)^{\alpha-1} x P_{2m+1}^{(\alpha,\alpha)}(x) dx,$$

we thus have

$$S_{2m,2n}^\alpha = \frac{(1 + \alpha)_{2n}}{(2n)!} \epsilon_m, \quad S_{2m+1,2n+1}^\alpha = \frac{(1 + \alpha)_{2n+1}}{(2n + 1)!} \mathfrak{o}_m.$$

We wish to prove that

$$\epsilon_m = \frac{4^\alpha \Gamma(2m + 1 + \alpha)\Gamma(1 + \alpha)}{\alpha \Gamma(2m + 1 + 2\alpha)}, \quad \mathfrak{o}_m = \frac{4^\alpha \Gamma(2m + 2 + \alpha)\Gamma(1 + \alpha)}{\alpha \Gamma(2m + 2 + 2\alpha)}. \tag{B.1}$$

To this end, it is helpful to rewrite (B.1) in the form

$$\epsilon_m = \frac{\sqrt{\pi} (m + 1 + \alpha)_m \Gamma(\alpha)}{4^m \Gamma(m + \alpha + \frac{1}{2})}, \quad \mathfrak{o}_m = \frac{\sqrt{\pi} (m + 1 + \alpha)_{m+1} \Gamma(\alpha)}{2 \cdot 4^m \Gamma(\alpha + m + \frac{3}{2})}. \tag{B.2}$$

To prove that (B.1) is identical to (B.2) for ϵ_m we commence from the latter, noting that it is the same as

$$\epsilon_m = \frac{\sqrt{\pi}}{4^m} \frac{\Gamma(2m + 1 + \alpha)\Gamma(\alpha)}{\Gamma(m + \alpha + \frac{1}{2})\Gamma(m + 1 + \alpha)}$$

and use the Gamma duplication formula

$$\Gamma(2z) = \pi^{-1/2} 2^{2z-1} \Gamma(z)\Gamma(z + \frac{1}{2}), \quad z \in \mathbb{C} \setminus -\mathbb{Z}_+$$

[21, 5.5.5]. Letting $z = m + \alpha + \frac{1}{2}$, we have

$$\Gamma(m + \alpha + \frac{1}{2})\Gamma(m + 1 + \alpha) = \frac{\sqrt{\pi}\Gamma(2m + 1 + 2\alpha)}{4^{m+\alpha}}.$$

and obtain (B.1) following elementary manipulation. An identical procedure applies to \mathfrak{o}_m .

Replacing m by $2m$ in (3.6) results in the recursion

$$\epsilon_m = \frac{1}{2} \frac{(2m + \alpha)(4m - 1 + 2\alpha)}{m + \alpha} \mathfrak{o}_{m-1} - \frac{1}{2} \frac{(2m - 1 + \alpha)(2m + \alpha)}{m + \alpha} \epsilon_{m-1}, \tag{B.3}$$

while replacing m by $2m + 1$ results in

$$\mathfrak{o}_m = \frac{(2m + 1 + \alpha)(4m + 1 + 2\alpha)}{2m + 1 + 2\alpha} \int_{-1}^1 (1 - x^2)^{\alpha-1} x^2 P_{2m}^{(\alpha,\alpha)}(x) dx$$

$$-\frac{(2m + \alpha)(2m + 1 + \alpha)}{2m + 1 + 2\alpha} \mathfrak{o}_{m-1}.$$

Replacing $x^2 = 1 - (1 - x^2)$ and using orthogonality,

$$\begin{aligned} \int_{-1}^1 (1 - x^2)^{\alpha-1} x^2 P_{2m}^{(\alpha, \alpha)}(x) dx &= \int_{-1}^1 (1 - x^2)^{\alpha-1} P_{2m}^{(\alpha, \alpha)}(x) dx \\ &\quad - \int_{-1}^1 (1 - x^2)^{\alpha} P_{2m}^{(\alpha, \alpha)}(x) dx = \mathfrak{e}_m \end{aligned}$$

for $m \in \mathbb{N}$. Thus,

$$\mathfrak{o}_m = \frac{(2m + 1 + \alpha)(4m + 1 + 2\alpha)}{2m + 1 + 2\alpha} \mathfrak{e}_m - \frac{(2m + \alpha)(2m + 1 + \alpha)}{2m + 1 + 2\alpha} \mathfrak{o}_{m-1}. \quad (\text{B.4})$$

We compute directly

$$\mathfrak{e}_0 = \frac{\sqrt{\pi}(1 + \alpha)\Gamma(\alpha)}{\Gamma(\alpha + \frac{1}{2})}, \quad \mathfrak{o}_0 = \frac{\sqrt{\pi}(1 + \alpha)\Gamma(\alpha)}{2\Gamma(\alpha + \frac{3}{2})}$$

(this is consistent with (B.2) for $m = 0$), whereby (B.2) follows from (B.3) and (B.4) by easy induction. We deduce that

$$S_{m,n}^{\alpha} = \frac{4^{\alpha}}{\alpha} \frac{\Gamma(m + 1 + \alpha)\Gamma(n + 1 + \alpha)}{n!\Gamma(m + 1 + 2\alpha)}, \quad m \geq n, \quad m + n \text{ even}. \quad (\text{B.5})$$

References

1. Philipp Bader, Arieh Iserles, Karolina Kropielnicka, and Pranav Singh. Effective approximation for the semiclassical Schrödinger equation. *Found. Comput. Math.*, 14(4):689–720, 2014.
2. Sergio Blanes and Vasile Gradinaru. High order efficient splittings for the semiclassical time-dependent Schrödinger equation. *J. Comput. Phys.*, 405:109157, 13, 2020.
3. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer-Verlag, Berlin, 2006.
4. Diego Caratelli, Ernesto Palini, and Paolo Emilio Ricci. Finite dimensional applications of the Dunford-Taylor integral. *Bull. TICMI*, 25(1):63–75, 2021.
5. T. S. Chihara. *An Introduction to Orthogonal Polynomials*. Mathematics and its Applications, Vol. 13. Gordon and Breach Science Publishers, New York-London-Paris, 1978.
6. Y. Eidelman and I. Gohberg. Algorithms for inversion of diagonal plus semiseparable operator matrices. *Integral Equations Operator Theory*, 44(2):172–211, 2002.
7. Wolfgang Hackbusch. *Hierarchical matrices: algorithms and analysis*, volume 49 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2015.
8. Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.
9. Jan S. Hesthaven, Sigal Gottlieb, and David Gottlieb. *Spectral methods for time-dependent problems*, volume 21 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2007.
10. Marlis Hochbruck and Christian Lubich. On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 34(5):1911–1925, 1997.

11. Arieh Iserles and Marcus Webb. Orthogonal systems with a skew-symmetric differentiation matrix. *Found. Comput. Math.*, 19(6):1191–1221, 2019.
12. Arieh Iserles and Marcus Webb. A family of orthogonal rational functions and other orthogonal systems with a skew-Hermitian differentiation matrix. *J. Fourier Anal. Appl.*, 26(1):Paper No. 19, 2020.
13. Arieh Iserles and Marcus Webb. A differential analogue of Favard's theorem. In *From operator theory to orthogonal polynomials, combinatorics, and number theory—a volume in honor of Lance Littlejohn's 70th birthday*, volume 285 of *Oper. Theory Adv. Appl.*, pages 239–263. Birkhäuser/Springer, Cham, 2021.
14. Arieh Iserles and Marcus Webb. Fast computation of orthogonal systems with a skew-symmetric differentiation matrix. *Comm. Pure Appl. Math.*, 74(3):478–506, 2021.
15. Mourad E. H. Ismail. *Classical and Quantum Orthogonal Polynomials in One Variable*, volume 98 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.
16. V. P. Konoplev. Polynomials orthogonal with respect to weight functions which are zero or infinite at isolated points of the interval of orthogonality. *Dokl. Akad. Nauk SSSR*, 141:781–784, 1961.
17. V. P. Konoplev. The asymptotic behaviour of orthogonal polynomials at one-sided singularities of weight functions (algebraic singularities). *Dokl. Akad. Nauk SSSR*, 160:997–1000, 1965.
18. Jean Bernard Lasserre, Edouard Pauwels, and Mihai Putinar. *The Christoffel–Darboux Kernel for Data Analysis*, volume 38 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2022.
19. Karen Min Luong. *Approximation of Wave Packets on the Real Line*. PhD thesis, University of Cambridge, 2023.
20. Robert I. McLachlan and G. Reinout W. Quispel. Splitting methods. *Acta Numer.*, 11:341–434, 2002.
21. Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions*. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010.
22. Sheehan Olver, Richard Mikaël Slevinsky, and Alex Townsend. Fast algorithms using orthogonal polynomials. *Acta Numer.*, 29:573–699, 2020.
23. Earl D. Rainville. *Special Functions*. The Macmillan Co., New York, 1960.
24. Barry Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005.
25. Barry Simon. The Christoffel–Darboux kernel. In *Perspectives in Partial Differential Equations, Harmonic Analysis and Applications*, volume 79 of *Proc. Sympos. Pure Math.*, pages 295–335. Amer. Math. Soc., Providence, RI, 2008.
26. Gábor Szegő. *Orthogonal polynomials*. American Mathematical Society Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, R.I., fourth edition, 1975.
27. Lloyd N. Trefethen. *Spectral methods in MATLAB*, volume 10 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.
28. Ellen Van Camp, Nicola Mastronardi, and Marc Van Barel. Two fast algorithms for solving diagonal-plus-semiseparable linear systems. In *Proceedings of the 10th International Congress on Computational and Applied Mathematics (ICCAM-2002)*, volume 164/165, pages 731–747, 2004.
29. J. L. Walsh. *Interpolation and approximation by rational functions in the complex domain*. American Mathematical Society Colloquium Publications, Vol. XX. American Mathematical Society, Providence, R.I., fourth edition, 1965.