On a filter for exponentially localized kernels based on Jacobi polynomials

F. Filbir*, H. N. Mhaskar†, and J. Prestin‡

Abstract

Let $\alpha, \beta \ge -1/2$, and for $k = 0, 1, \cdots$, ${p_k}^{(\alpha,\beta)}$ denote the orthonormalized Jacobi polynomial of degree k. We discuss the construction of a matrix H so that there exist positive constants c, c_1, depending only on H, α, and β such that

$$\left|\sum_{k=0}^{\infty} H_{k,n} {p_k}^{(\alpha,\beta)}(\cos\theta) {p_k}^{(\alpha,\beta)}(\cos\varphi)\right| \le c_1 n^{2\max(\alpha,\beta)+2} \exp(-cn(\theta-\varphi)^2), \quad \theta, \varphi \in [0,\pi],\ n = 1, 2 \cdots.$$

Specializing to the case of Chebyshev polynomials, $\alpha = \beta = -1/2$, we apply this theory to obtain a construction of an exponentially localized polynomial basis for the corresponding L^2 space.

1 Introduction

The problem of detection of singularities of a function from spectral data and the closely related problem of spectral approximation of piecewise smooth or analytic functions arise in many important applications, for example, computer tomography [14], nuclear magnetic resonance inversion [4], and conservation laws in differential equations [30]. These problems are studied by many authors; some recent references are [32, 31, 33], and references therein. Typically, one finds first the location of singularities using an appropriate filter, and then uses pseudo-spectral methods on the maximum intervals of smoothness to compute the approximation on these intervals. A *filter* is a bi–infinite matrix H, and the corresponding *mollifier* is given by

$$\Phi_n^\circ(H,\theta) = \sum_{k\in\mathbb{Z}} H_{k,n}\exp(ik\theta),\ \ \sigma_n^\circ(H,f,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varphi)\Phi_n^\circ(H,\theta-\varphi)d\varphi, \qquad \theta\in\mathbb{R}.$$

We note that in classical harmonic analysis parlance, Φ_n° (respectively, σ_n°) is called a summability kernel (respectively, summability operator). We have studied the construction of filters for the detection of jump discontinuities in high order derivatives of f as well as spectral approximation of piecwise smooth functions in several papers, for example, [24, 25, 22]. A relatively recent survey can be found in [26]. They are all of the form $H_{k,n} = h(|k|/n)$ for a suitable, compactly supported function h. The corresponding kernels typically satisfy [24] a localization condition of the form

$$|\Phi_n^\circ(H,\theta)| \le c(H,Q)\frac{n}{(1+n|\theta|)^Q}, \qquad \theta\in(-\pi,\pi], \tag{1.1}$$

*Institute of Biomathematics and Biometry, Helmholtz Center Munich, 85764 Neuherberg, Germany, email: filbir@helmholtz-muenchen.de

†Department of Mathematics, California State University, Los Angeles, California, 90032, USA, email: hmhaska@calstatela.edu. The research of this author was supported, in part, by grant DMS-0605209 from the National Science Foundation, grant W911NF-04-1-0339 from the U.S. Army Research Office, and a fellowship from AvH Foundation.

‡Institute of Mathematics, University of Lübeck, Wallstraße 40, 23560, Lübeck, Germany, email: prestin@math.uni-luebeck.de.

where Q is a positive number, depending usually on the smoothness of the function h. We can obtain an arbitrarily large value of Q by choosing h to be an infinitely differentiable function. This localization holds also in the very general context of orthonormal families on a metric measure space [21].

If the target function is piecewise analytic, then the localization of the form (1.1) is not enough to obtain spectral approximation; i.e., obtain a geometrically decreasing degree of approximation in the intervals of analyticity, while yielding a near best degree of approximation globally. One of the first results in this direction that we are aware of was given by Gaier [10]. In this section only, for a continuous function $f : [-1,1] \to \mathbb{R}$, let $\|f\|_\infty = \max_{x\in[-1,1]} |f(x)|$, and $E_{n,\infty}(f) := \min_{P\in\Pi_n} \|f - P\|_\infty$. In [10], Gaier constructed a sequence of linear operators $\mathcal{G}_n$ such that for each continuous $f : [-1,1] \to \mathbb{R}$, and integer $n \geq 1$, $\mathcal{G}_n(f) \in \Pi_n$, and satisfies the following conditions:

$$\max_{x\in[-1,1]} |f(x) - \mathcal{G}_n(f,x)| \leq M(f)e^{-cn} + c_1 E_{n/6,\infty}(f), \tag{1.2}$$

and if f is analytic in the complex neighborhood $|z - x_0| \leq d$ of a point $x_0 \in [-1,1]$, then

$$|f(x_0) - \mathcal{G}_n(f,x_0)| \leq M(f)d^{-4}\exp(-c_2 d^2 n),$$

where $M(f)$ is a positive constant depending only on f, and c, c_1, c_2 are absolute positive constants. Gaier's construction is based on the Fourier-Chebyshev coefficients of f and depends heavily on a resulting contour integral. A more recent construction based on Chebyshev polynomials is by Tanner [32], where a filter is proposed for spectral approximation of piecewise analytic functions. This filter, however, depends upon the point x where the approximation is desired. Theoretically, Tanner's construction requires an a priori knowledge of the location of the singularities of the target function.

In [27], we have given a very simple construction to solve both the problems of singularity detection and spectral approximation in one stroke. The idea is to take a reproducing summability kernel Φ_n (as defined in [27]), without regard to localization, and consider the kernel

$$\Phi_n^*(x,y) = \left(1 - \frac{(x-y)^2}{4}\right)^n \Phi_n(x,y). \tag{1.3}$$

Of course, the factor in front of Φ_n above is just a simple choice. The theory of fast decreasing polynomials developed by Ivanov, Saff, and Totik among others ([17, 28] and references therein) deals with the construction of polynomials S_n of degree at most n such that

$$S_n(0) = 1,\ |S_n(t)| \leq c_1 \exp(-n\phi(t)) \text{ for } |t| \leq 1, \tag{1.4}$$

for a suitable function ϕ and a positive constant c_1 independent of n and t. Necessary and sufficient conditions on ϕ to ensure the existence of such polynomials can be found in [28]. Therefore, given any such function ϕ, the polynomial $S_n((x-y)/2)$ will work in place of $\left(1 - \frac{(x-y)^2}{4}\right)^n$ in (1.3) to give different dependences on the distance between x and y.

In the Fourier domain, the kernel Φ_n^* can be described as a *mixed filter*; rather than modifying each of the Fourier coefficients $\hat{f}(k)$ separately, we take a linear combination of all the available spectral data at each frequency. A major advantage of this construction is that it is applicable to a wide class of orthogonal polynomial expansions, whereas every other method known so far utilizes only the Chebyshev expansions. In particular, the construction does not depend upon special function properties of the orthogonal polynomial system involved.

The purpose of this paper is to demonstrate the construction of a filter in the important case of Jacobi polynomials. Let $\alpha, \beta \geq -1/2$, and for $k = 0, 1, \cdots$, ${p_k}^{(\alpha,\beta)}$ denote the orthonormalized Jacobi polynomial of degree k. We discuss the construction of a matrix H so that

$$\left|\sum_{k=0}^{\infty} H_{k,n} {p_k}^{(\alpha,\beta)}(\cos\theta) {p_k}^{(\alpha,\beta)}(\cos\varphi)\right| \leq c_1 n^{2\max(\alpha,\beta)+2} \exp(-cn(\theta-\varphi)^2), \qquad \theta, \varphi \in [0,\pi].$$

Specializing to the case of Chebyshev polynomials, $\alpha = \beta = -1/2$, we apply this theory to obtain a construction of an exponentially localized polynomial basis for the corresponding Hilbert space $L^2(-1/2,-1/2)$ defined in Section 2.

We review certain basic facts about Jacobi polynomials in Section 2. The construction of the filter and the properties of the corresponding mollifier are described in Section 3. In Section 4, we discuss the numerical computation of the filter, and demonstrate the superior localization properties of the exponentially localized kernels over a similar kernel obtained by using an infinitely differentiable mask. The construction of Riesz basis in the special case of Chebyshev polynomials is described in Section 5.

2 Jacobi polynomials

In this section, we introduce some notation, and review some basic facts concerning Jacobi polynomials. This material is based on [1, 3, 29]. Let $\alpha, \beta \geq -1/2$, and

$$w_{\alpha,\beta}(x) := \begin{cases} (1-x)^\alpha(1+x)^\beta, & \text{if } -1 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

For $1 \leq p < \infty$ the space $L^p(\alpha,\beta)$ is defined as the space of (equivalence classes of) functions f with

$$\|f\|_{\alpha,\beta;p} := \left(\int_{-1}^1 |f(x)|^p w_{\alpha,\beta}(x)dx\right)^{1/p} < \infty.$$

To simplify the statements of our theorems, we will adopt the notation often used by Butzer and his collaborators: the symbol $X^p(\alpha,\beta)$ denotes $L^p(\alpha,\beta)$, if $1 \leq p < \infty$, and $C[-1,1]$, the space of continuous functions on $[-1,1]$ with the maximum norm $\|\circ\|_\infty$, if $p=\infty$.

There exists a unique system of (orthonormalized Jacobi) polynomials $\{{p_k}^{(\alpha,\beta)}(x) = \gamma_k(\alpha,\beta)x^k + \cdots\}$, $\gamma_k(\alpha,\beta) > 0$ such that for integer $k, \ell = 0, 1, \cdots$,

$$\int_{-1}^1 {p_k}^{(\alpha,\beta)}(x){p_\ell}^{(\alpha,\beta)}(x)w_{\alpha,\beta}(x)dx = \begin{cases} 1, & \text{if } k=\ell, \\ 0, & \text{otherwise.} \end{cases}$$

The uniqueness of the system implies that ${p_k}^{(\beta,\alpha)}(x) = (-1)^k {p_k}^{(\alpha,\beta)}(-x)$, $x \in \mathbb{R}$, $k = 0, 1, \cdots$. Therefore, we may assume in the sequel that $\alpha \geq \beta$. We will assume also that $\alpha \geq \beta \geq -1/2$.

We will often find it convenient to use other normalizations for the Jacobi polynomials. In particular, defining for integer $n \geq 0$ and $x \in (-1,1)$,

$$w_{\alpha,\beta}(x){P_n}^{(\alpha,\beta)}(x) := \frac{(-1)^n}{2^n n!}\frac{d^n}{dx^n} w_{n+\alpha,n+\beta}(x), \tag{2.1}$$

and

$$\kappa_n^{(\alpha,\beta)} := \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}, \tag{2.2}$$

we have [29, Chapter IV]

$${p_n}^{(\alpha,\beta)} = \{\kappa_n^{(\alpha,\beta)}\}^{-1/2} P_n^{(\alpha,\beta)}, \tag{2.3}$$

and

$${P_n}^{(\alpha,\beta)}(1) = \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}. \tag{2.4}$$

For $f \in X^1(\alpha,\beta)$, we may define

$$\hat{f}(k) := \hat{f}(\alpha,\beta;k) := \int_{-1}^1 f(y){p_k}^{(\alpha,\beta)}(y)w_{\alpha,\beta}(y)dy, \qquad k = 0, 1, \cdots,$$

and the Fourier projection

$$\mathbf{s}_n(f,x) := \mathbf{s}_n(\alpha,\beta;f,x) := \sum_{k=0}^{n-1} \hat{f}(k){p_k}^{(\alpha,\beta)}(x), \qquad x \in [-1,1],\ n = 1, 2, \cdots. \tag{2.5}$$

It is well known that if f is analytic on $[-1,1]$ then

$$\limsup_{n\to\infty} \|f-\mathbf{s}_n(f)\|_\infty^{1/n} < 1,$$

so that $\mathbf{s}_n(f)\to f$ exponentially rapidly in the uniform norm. We define the Christoffel–Darboux kernel as in [9] by

$$\begin{aligned} K_n(x,y) &:= K_n(\alpha,\beta;x,y) := \sum_{k=0}^{n-1} p_k{}^{(\alpha,\beta)}(x)p_k{}^{(\alpha,\beta)}(y) \\ &= \frac{\gamma_{n-1}(\alpha,\beta)}{\gamma_n(\alpha,\beta)}\frac{p_n{}^{(\alpha,\beta)}(x)p_{n-1}{}^{(\alpha,\beta)}(y)-p_{n-1}{}^{(\alpha,\beta)}(x)p_n{}^{(\alpha,\beta)}(y)}{x-y}, \end{aligned} \tag{2.6}$$

and observe that

$$\mathbf{s}_n(f,x) = \int_{-1}^1 f(y)K_n(x,y)w_{\alpha,\beta}(y)dy.$$

As usual, if n is not an integer, K_n and $\mathbf{s}_n$ will mean $K_{\lfloor n\rfloor}$ and $\mathbf{s}_{\lfloor n\rfloor}$ respectively.

For $\alpha\ge\beta\ge -1/2$, Koornwinder [20] has proved that there exists a probability measure $\nu^{(\alpha,\beta)}$ on $[0,1]\times[0,\pi]$ such that with

$$Z(x,y;r,\psi) = \frac{1}{2}(1+x)(1+y) + \sqrt{1-x^2}\sqrt{1-y^2}\, r\cos\psi + \frac{1}{2}(1-x)(1-y)\, r^2 - 1, \tag{2.7}$$

we have

$$p_n^{(\alpha,\beta)}(x)p_n^{(\alpha,\beta)}(y) = \int_0^\pi\int_0^1 p_n^{(\alpha,\beta)}(1)p_n^{(\alpha,\beta)}(Z(x,y;r,\psi))\; d\nu^{(\alpha,\beta)}(r,\psi), \quad n=0,1,\cdots. \tag{2.8}$$

We note that the support of the measure $\nu^{(\alpha,\beta)}$ is not necessarily the whole square. For example, in the case when $\alpha=\beta>-1/2$, the measure is equal to $\delta_1\times\nu_\beta$, where (in this paragraph only) δ_1 is the Dirac delta measure at $r=1$ and ν_β is a probability measure on $[0,\pi]$. In the case when $\alpha=\beta=-1/2$, the measure $\nu^{(-1/2,-1/2)}$ is just the average of the Dirac delta measures at the points $(1,0)$ and $(1,\pi)$, and (2.8) reduces to the product formula for cosines. The exact expressions for $\nu^{(\alpha,\beta)}$ are not relevant for this paper. The concrete formulas can be found in [15, p. 308].

An interesting consequence of (2.8) is the following. For almost all $x,y\in[-1,1]$, and $f\in L^1(\alpha,\beta)$, let

$$\mathcal{T}_y f(x) = \int_0^\pi\int_0^1 f(Z(x,y;r,\psi))\; d\nu^{(\alpha,\beta)}(r,\psi). \tag{2.9}$$

Then

$$\widehat{\mathcal{T}_y f}(k) = \hat{f}(k)\frac{p_k{}^{(\alpha,\beta)}(y)}{p_k{}^{(\alpha,\beta)}(1)}. \tag{2.10}$$

If f is a 2π–periodic, continuous function, with trigonometric Fourier coefficients given by $\{b_k\}_{k\in\mathbb{Z}}$, the trigonometric Fourier coefficients of $f(\circ+y)$ are given by $\{e^{iky}b_k\}$. Therefore, it is reasonable to refer to $\mathcal{T}_y$ as a translation operator. The corresponding convolution operator is defined by

$$(f*g)(x) := \int_{-1}^1 f(y)\,\mathcal{T}_y g(x)\; w_{\alpha,\beta}(y)\,dy, \qquad f,g\in L^1(\alpha,\beta).$$

The operator $\mathcal{T}_y$ and the corresponding convolution operator share several interesting properties with their usual analogues for periodic functions.

Proposition 2.1 *For every $1\le p\le\infty$, $f,g\in X^p(\alpha,\beta)$, and $y\in[-1,1]$, we have*

$$\|\mathcal{T}_y f\|_{\alpha,\beta;p} \le \|f\|_{\alpha,\beta;p}, \tag{2.11}$$

$$\lim_{y\to 1-}\|\mathcal{T}_y f - f\|_{\alpha,\beta;p} = 0, \tag{2.12}$$

$$(f*g)(x)=\int_{-1}^{1} f(y)\,\mathcal{T}_y g(x)\;w_{\alpha,\beta}(y)\,dy=\int_{-1}^{1} \mathcal{T}_y f(x)\,g(y)\;w_{\alpha,\beta}(y)\,dy=(g*f)(x), \tag{2.13}$$

$$\|f*g\|_{\alpha,\beta;p}\le \|f\|_{\alpha,\beta;1}\,\|g\|_{\alpha,\beta;p}, \tag{2.14}$$

and

$$\widehat{(f*g)}(k)=\{p_k{}^{(\alpha,\beta)}(1)\}^{-1}\widehat{f}(k)\,\widehat{g}(k). \tag{2.15}$$

Proof. Let $x=\cos\theta$, $y=\cos\varphi$, $r\in[0,1]$, $\psi\in[0,\pi]$. Then

$$\begin{aligned}
Z(x,y;r,\psi) &= \frac{1}{2}(1+x)(1+y)+\frac{1}{2}(1-x)(1-y)+\sqrt{1-x^2}\sqrt{1-y^2}\\
&\quad+\frac{1}{2}(1-x)(1-y)(r^2-1)+\sqrt{1-x^2}\sqrt{1-y^2}(r\cos\psi-1)-1\\
&= xy+\sqrt{1-x^2}\sqrt{1-y^2}+\frac{1}{2}(1-x)(1-y)(r^2-1)+\sqrt{1-x^2}\sqrt{1-y^2}(r\cos\psi-1),
\end{aligned}$$

and we have

$$\begin{aligned}
1-Z(x,y;r,\psi) &= 1-(xy+\sqrt{1-x^2}\sqrt{1-y^2})+\frac{1}{2}(1-x)(1-y)(1-r^2)\\
&\quad+(1-r\cos\psi)\sqrt{1-x^2}\sqrt{1-y^2}\\
&\ge 1-(xy+\sqrt{1-x^2}\sqrt{1-y^2})=1-\cos(\theta-\varphi)\ge 0.
\end{aligned} \tag{2.16}$$

Further, the arithmetic–geometric mean inequality implies that

$$\begin{aligned}
1+Z(x,y;r,\psi) &= \frac{1}{2}(1+x)(1+y)+\frac{1}{2}(1-x)(1-y)\,r^2+\sqrt{1-x^2}\sqrt{1-y^2}\,r\cos\psi\\
&\ge \sqrt{1-x^2}\sqrt{1-y^2}\,r(1+\cos\psi)\ge 0.
\end{aligned}$$

Thus, for every $x,y\in[-1,1]$, $r\in[0,1]$, $\psi\in[0,\pi]$, $Z(x,y;r,\psi)\in[-1,1]$ (cf. [1, p. 10]). It follows from (2.9) that $\mathcal{T}_y$ is a positive operator; i.e., $g(x)\ge 0$ for almost all $x\in[-1,1]$ implies that $\mathcal{T}_y g(x)\ge 0$ for almost all $x\in[-1,1]$. Therefore, for such functions g,

$$\|\mathcal{T}_y g\|_{\alpha,\beta;1}=\int_{-1}^{1}\mathcal{T}_y g(x)w_{\alpha,\beta}(x)dx=\{\kappa_0^{(\alpha,\beta)}\}^{1/2}\widehat{\mathcal{T}_y g}(0)=\{\kappa_0^{(\alpha,\beta)}\}^{1/2}\hat{g}(0)=\int_{-1}^{1}g(x)w_{\alpha,\beta}(x)dx=\|g\|_{\alpha,\beta;1}.$$

Applying this equation with $|f|$ in place of g, we deduce (2.11) in the case when $p=1$ [1, Equation (2.26)]. The case $p=\infty$ is obvious, and the general case follows from the Riesz–Thorin interpolation theorem [2, Theorem 1.1.1]. The equation (2.12) is clear when f is a polynomial, and follows easily in the general case using the Weierstrass approximation theorem and (2.11). The equation (2.15) is an immediate consequence of (2.10) and the relevant definitions. The equation (2.13) follows from (2.15). The estimate (2.14) follows from (2.11) in the case when $p=1$, is obvious when $p=\infty$, and follows from the Riesz–Thorin interpolation theorem in the intermediate cases. □

Dual to the product formula (2.8) is a second product formula:

$$p_n^{(\alpha,\beta)}(x)p_m^{(\alpha,\beta)}(x)=\sum_{k=|n-m|}^{n+m} g^{(\alpha,\beta)}(n,m;k)\frac{p_n{}^{(\alpha,\beta)}(1)p_m{}^{(\alpha,\beta)}(1)}{p_k{}^{(\alpha,\beta)}(1)}\,p_k^{(\alpha,\beta)}(x) \tag{2.17}$$

for all $n,m\in\mathbb{N}_0, x\in[-1,1]$. The coefficients $g^{(\alpha,\beta)}(n,m;k)$ are non-negative for all $k,n,m\in\mathbb{N}_0$ ([1, Theorem 5.1, p. 42]) and, moreover,

$$\sum_{k=|n-m|}^{n+m} g^{(\alpha,\beta)}(n,m;k)=1. \tag{2.18}$$

Thus, we may think of $g^{(\alpha,\beta)}(n,m;\circ)$ as a probability distribution on a subset of $\mathbb{N}_0\times\mathbb{N}_0$. An explicit expression is known for the coefficients $g^{(\alpha,\alpha)}(n,m;k)$ [1, Formula (5.7), p. 39] in the case of ultraspherical polynomials. In particular, if $\alpha = \beta = -1/2$, then $g^{(-1/2,-1/2)}(n,m;k) = 0$ except when $k = |n-m|$ or $k = n+m$, when $g^{(-1/2,-1/2)}(n,m;|n-m|) = g^{(-1/2,-1/2)}(n,m;n+m) = 1/2$, and (2.17) also reduces to the product formula for cosines.

Just as (2.8) can be used via (2.9) to define a generalized convolution of two functions, (2.17) can be used to define a convolution of sequences. If $\mathbf{a} = \{a_k\}_{k=0}^{\infty}$ and $\mathbf{b} = \{b_k\}_{k=0}^{\infty}$, we define formally

$$(\mathbf{a} * \mathbf{b})(k) = \sum_{n,m=0}^{\infty} g^{(\alpha,\beta)}(n,m;k) a_n b_m. \tag{2.19}$$

Analogous to (2.15) and the classical Cauchy formula for the products of power series, we have for $x \in [-1,1]$,

$$\left(\sum_{n=0}^{\infty} a_n {p_n}^{(\alpha,\beta)}(1) {p_n}^{(\alpha,\beta)}(x)\right)\left(\sum_{m=0}^{\infty} b_m {p_m}^{(\alpha,\beta)}(1) {p_m}^{(\alpha,\beta)}(x)\right) = \sum_{k=0}^{\infty} (\mathbf{a} * \mathbf{b})(k) {p_k}^{(\alpha,\beta)}(1) {p_k}^{(\alpha,\beta)}(x). \tag{2.20}$$

We do not wish to complicate our notations further by using different notations for the convolution of sequences and that of functions, since we feel that the context will make it clear which one is intended.

3 Filters and kernels

We will construct kernels localized at 1, and then use the theory of translations in Section 2 to obtain their bivariate version. In the sequel, for integer $n \ge 0$, Π_n denotes the class of all polynomials of degree at most n. We prefer to use the same notation even when n is not an integer; Π_n is then just the class of all polynomials of degree not exceeding the integer part of n. The symbols $c, c_1, \cdots$ will denote generic positive constants depending only on α, β and certain fixed functions to be introduced later.

Let $n \ge 1$ be an integer, $\mathbf{h}_n = \{h_{k,n}\}_{k=0}^{\infty}$. The starting point of the construction of our filters is the kernel

$$\tilde{\Phi}_n(t) := \tilde{\Phi}_n(\alpha,\beta;\mathbf{h}_n,t) := \sum_{k=0}^{4n} h_{k,n} {p_k}^{(\alpha,\beta)}(1) {p_k}^{(\alpha,\beta)}(t). \tag{3.1}$$

We will explain later the choice of the sequence $\mathbf{h}_n$ to ensure different desirable properties of $\tilde{\Phi}_n$.

Theorem 3.1 *Let $\phi : [0,2] \to [0,\infty]$ be a nondecreasing function, continuous in the extended sense. For integer $n \ge 1$, let $\tilde{\Phi}_n$ be as in (3.1), and $S_n \in \Pi_n$ satisfy*

$$S_n(1) = 1, \quad |S_n(t)| \le c_1 \exp(-n\phi(1-t)), \quad t \in [-1,1]. \tag{3.2}$$

Then, with the constant c_1 as in (3.2), we have for $x = \cos\theta$, $y = \cos\varphi \in [-1,1]$,

$$|\mathcal{T}_y(S_n\tilde{\Phi}_n)(x)| \le c_1 \|\tilde{\Phi}_n\|_\infty \exp\left(-n\phi(1-\cos(\theta-\varphi))\right), \tag{3.3}$$

and

$$\sup_{x\in[-1,1]} \int_{-1}^{1} |\mathcal{T}_y(S_n\tilde{\Phi}_n)(x)| w_{\alpha,\beta}(y)dy \le c_1 \int_{-1}^{1} |\tilde{\Phi}_n(t)| w_{\alpha,\beta}(t)dt. \tag{3.4}$$

Moreover, if $h_{k,n} = 1$ for $k = 0, \cdots, 2n$, then for every $P \in \Pi_n$, $x \in [-1,1]$,

$$((S_n\tilde{\Phi}_n) * P)(x) = P(x), \; x \in [-1,1]. \tag{3.5}$$

Remark. In the above theorem, we do not assume any localization on $\tilde{\Phi}_n$; the localization depends entirely on S_n. In view of the results described in [28], it appears that one cannot have the familiar dependence on $n|\theta-\varphi|$. However, in contrast to the main concern in the theory of fast decreasing polynomials, the actual dependence on $\theta-\varphi$ is not important here. On the other hand, it is important that the n-th root of the bound be eventually less than 1 for every fixed $\theta-\varphi$, to ensure spectral approximation of piecewise analytic functions. In practice, one can get the "best of both worlds" by taking a localized kernel $\tilde{\Phi}_n$. For example, if $\tilde{\Phi}_n$ is chosen to satisfy a localization estimate of the form

$$|\tilde{\Phi}_n(\cos\theta)| \le c(Q)\frac{n^{2\alpha+2}}{(1+n\theta)^Q}, \qquad \theta\in[0,\pi], \tag{3.6}$$

then Lemma 3.1 can be used to obtain a sharper version of (3.3) for $x=\cos\theta$, $y=\cos\varphi$:

$$|\mathcal{T}_y(S_n\tilde{\Phi}_n)(x)| \le c(Q)\frac{n^{2\alpha+2}}{(1+n|\theta-\varphi|)^Q}\exp\left(-n\phi(1-\cos(\theta-\varphi))\right). \tag{3.7}$$

A few weeks after we submitted the first version of this paper, we came across a manuscript [16] by Ivanov, Petrushev, and Xu, where the authors prove the existence of a function g such that a kernel $\tilde{\Phi}_n$ obtained by choosing $h_{k,n}$ to be $g(k/n)$ is localized subexponentially. In [27], we have shown by a numerical example that a kernel localized for every value of Q above is not sufficient for the detection of an analytic singularity, that the exponentially localized kernel leads to such a detection even without a significant localization on $\tilde{\Phi}_n$, and demonstrated how a better localized $\tilde{\Phi}_n$ sharpens this detection.

The proof of Theorem 3.1 requires the following lemma, relating the localization of the kernels at 1 with the localization in general.

Lemma 3.1 *Let $F\in L^1(\alpha,\beta)$, $\phi^\diamondsuit:[0,2]\to[0,\infty]$ be a nonincreasing function, and for some constant $A>0$, $|F(t)|\le A\phi^\diamondsuit(1-t)$ for almost all $t\in[-1,1]$. Then for almost all $\theta,\varphi\in[0,\pi]$, $x=\cos\theta$, $y=\cos\varphi$, we have*

$$|\mathcal{T}_yF(x)|\le A\phi^\diamondsuit(1-\cos(\theta-\varphi)).$$

PROOF. In light of (2.9), we deduce using (2.16) and the fact that $\phi^\diamondsuit$ is nonincreasing that

$$\begin{aligned}
|\mathcal{T}_yF(x)| &= \left|\int_0^\pi\int_0^1 F(Z(x,y;r,\psi))\,d\nu^{(\alpha,\beta)}(r,\psi)\right| \\
&\le A\int_0^\pi\int_0^1 \phi^\diamondsuit(1-Z(x,y;r,\psi))d\nu^{(\alpha,\beta)}(r,\psi) \\
&\le A\int_0^\pi\int_0^1 \phi^\diamondsuit(1-\cos(\theta-\varphi))d\nu^{(\alpha,\beta)}(r,\psi) = A\phi^\diamondsuit(1-\cos(\theta-\varphi)).
\end{aligned}$$

□

PROOF OF THEOREM 3.1. Since $|S_n(t)\tilde{\Phi}_n(t)|\le c_1\|\tilde{\Phi}_n\|_\infty\exp(-n\phi(1-t))$ for $|t|\le 1$, the first statement is clear from Lemma 3.1 used with $\exp(-n\phi(t))$ in place of $\phi^\diamondsuit$, and the fact that $\mathcal{T}_y(S_n\tilde{\Phi}_n)(x)$, being a polynomial in x,y, is continuous in x,y. The estimate (3.4) follows from (2.11) and the fact that $|S_n(t)|\le c_1$ for $|t|\le 1$.

If $Q\in\Pi_{2n}$, then $\hat{Q}(k)=0$ for $k>2n$ and $Q(y)=\sum_{k=0}^{2n}\hat{Q}(k){p_k}^{(\alpha,\beta)}(y)$. Hence, if $h_{k,n}=1$ for $k=0,1,\cdots,2n$, then

$$\int_{-1}^1\tilde{\Phi}_n(y)Q(y)w_{\alpha,\beta}(y)dy = Q(1), \qquad Q\in\Pi_{2n}.$$

Let $P\in\Pi_n$. We let $Q(y)=S_n(y)P(y)$, and apply the above equation to obtain

$$\int_{-1}^1 S_n(y)\tilde{\Phi}_n(y)P(y)w_{\alpha,\beta}(y)dy = \int_{-1}^1\tilde{\Phi}_n(y)Q(y)w_{\alpha,\beta}(y)dy = Q(1) = P(1).$$

Next, for $x\in[-1,1]$, we use this equation with $\mathcal{T}_xP$ in place of P to deduce that

$$\begin{aligned}
\int_{-1}^1\mathcal{T}_y(S_n\tilde{\Phi}_n)(x)P(y)w_{\alpha,\beta}(y)dy &= \int_{-1}^1\mathcal{T}_x(S_n\tilde{\Phi}_n)(y)P(y)w_{\alpha,\beta}(y)dy \\
&= \int_{-1}^1 S_n(y)\tilde{\Phi}_n(y)\mathcal{T}_xP(y)w_{\alpha,\beta}(y)dy = \mathcal{T}_xP(1) = P(x).
\end{aligned}$$

This completes the proof of Theorem 3.1. □

While Theorem 3.1 shows the existence of filters which yield exponentially localized kernels, from a practical point of view, we find it necessary to construct an explicitly known filter, albeit at a cost of the great generality in the estimate (3.3). Such a kernel is given by Φ_n^* below. Let the sequence $\mathbf{a}_n = \{a_{k,n}(\alpha,\beta)\}_{k=0}^\infty$ be defined by

$$a_{k,n}(\alpha,\beta) := \begin{cases} \dfrac{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(n+1)\Gamma(n+\beta+1)}{\Gamma(n-k+1)\Gamma(n+k+\alpha+\beta+2)}, & k=0,1,\cdots,n, \\ 0, & \text{if } k\ge n+1, \end{cases} \tag{3.8}$$

and let

$$H_{k,n} := (\mathbf{a}_n * \mathbf{h}_n)(k), \qquad k=0,1,\cdots. \tag{3.9}$$

We define

$$\tilde{\Phi}_n^*(t) := \sum_{k=0}^{5n} H_{k,n} p_k{}^{(\alpha,\beta)}(1) p_k{}^{(\alpha,\beta)}(t), \qquad t\in[-1,1], \tag{3.10}$$

$$\Phi_n^*(x,y) := \mathcal{T}_y\tilde{\Phi}_n^*(x) = \mathcal{T}_x\tilde{\Phi}_n^*(y) = \sum_{k=0}^{5n} H_{k,n} p_k{}^{(\alpha,\beta)}(x) p_k{}^{(\alpha,\beta)}(y), \qquad x,y\in[-1,1]. \tag{3.11}$$

In view of (2.20), we have

$$\tilde{\Phi}_n^*(t) = \tilde{\Phi}_n(t)\ \sum_{k=0}^{n} a_{k,n}(\alpha,\beta) p_k{}^{(\alpha,\beta)}(1) p_k{}^{(\alpha,\beta)}(t), \tag{3.12}$$

where $\tilde{\Phi}_n$ is defined as in (3.1).

The following properties of the kernel Φ_n^* follow directly from Proposition 3.1 below and Theorem 3.1.

Theorem 3.2 *Let $n \ge 1$ be an integer.*
(a) *For $\theta,\varphi\in[0,\pi]$, $x=\cos\theta$, $y=\cos\varphi$,*

$$|\Phi_n^*(x,y)| \le \left(\frac{1+\cos(\theta-\varphi)}{2}\right)^n \|\tilde{\Phi}_n\|_\infty \le \|\tilde{\Phi}_n\|_\infty \exp\left(-\frac{n}{\pi^2}(\theta-\varphi)^2\right). \tag{3.13}$$

(b) *We have*

$$\sup_{x\in[-1,1]}\int_{-1}^1 |\Phi_n^*(x,y)| w_{\alpha,\beta}(y)dy \le \int_{-1}^1 |\tilde{\Phi}_n(t)| w_{\alpha,\beta}(t)dt. \tag{3.14}$$

(c) *If $h_{k,n}=1$ for $k=0,1,\cdots,2n$, then*

$$\int_{-1}^1 \Phi_n^*(x,y)P(y)w_{\alpha,\beta}(y)dy = P(x), \qquad P\in\Pi_n,\ x\in[-1,1]. \tag{3.15}$$

We state an application of this theorem for spectral approximation of functions. If $1\le p\le\infty$ and $f\in L^p(\alpha,\beta)$, we define for $n\ge 0$

$$E_{n,p}(\alpha,\beta;f) := \min_{P\in\Pi_n} \|f-P\|_{\alpha,\beta;p}.$$

Let

$$\sigma_n^*(f,x) := (\tilde{\Phi}_n^* * f)(x) = \int_{-1}^1 \Phi_n^*(x,y)f(y)w_{\alpha,\beta}(y)dy = \sum_k H_{k,n}\hat{f}(k)p_k{}^{(\alpha,\beta)}(x), \qquad x\in[-1,1].$$

The following theorem can be proved using Theorem 3.2 exactly as [27, Theorem 2.1], and its proof will be omitted. The only difference is that the operator σ_n in [27] was defined for more general measures than the Jacobi measures, but naturally, without using any convolution structure, while the operator σ_n^* is defined with a usual filter construction. Several numerical examples are given in [27] to illustrate the theorem with different examples, including one where the target function is infinitely differentiable, but has an analytic singularity on the interval.

Theorem 3.3 (a) *Let* $1 \le p \le \infty$, $\alpha, \beta \ge -1/2$, $f \in L^p(\alpha, \beta)$. *We have* $\sigma_n^*(P) = P$ *for* $P \in \Pi_n$, $\|\sigma_n^*(f)\|_{\alpha,\beta;p} \le c\|f\|_{\alpha,\beta;p}$, *and*

$$E_{5n,p}(\alpha, \beta; f) \le \|f - \sigma_n^*(f)\|_{\alpha,\beta;p} \le c_1 E_{n,p}(\alpha, \beta; f). \tag{3.16}$$

(b) *Let* $f \in C[-1, 1]$, $x_0 \in [-1, 1]$, *and* f *have an analytic continuation to a complex neighborhood of* x_0, *given by* $\{z \in \mathbb{C} : |z - x_0| \le d\}$ *for some* d *with* $0 < d \le 2$. *Then*

$$|f(x) - \sigma_n^*(f, x)| \le c(f, x_0) \exp\left(-c_1(d)n\frac{d^2 \log(e/2)}{e^2 \log(e^2/d)}\right), \qquad x \in [x_0 - d/e, x_0 + d/e] \cap [-1, 1]. \tag{3.17}$$

For the proof of Theorem 3.2, we need the following proposition.

Proposition 3.1 *For integer* $n \ge 1$, $t \in [-1, 1]$,

$$\left(\frac{1+t}{2}\right)^n = \sum_{k=0}^{n} a_{k,n}(\alpha, \beta) {p_k}^{(\alpha,\beta)}(1) {p_k}^{(\alpha,\beta)}(t). \tag{3.18}$$

Hence,

$$\tilde\Phi_n^*(t) = \left(\frac{1+t}{2}\right)^n \tilde\Phi_n(t), \qquad t \in [-1, 1]. \tag{3.19}$$

Proof. The equation (3.18) is the same as [1, Equation (2.30), p. 11], taking the normalization (2.3), (2.2) into account. The equation (3.19) follows from (2.20). □

Proof of Theorem 3.2. In this proof only, let $S_n(t) = ((1 + t)/2)^n$, and $\phi(t) = \log\{(1 - t/2)^{-1}\}$. Then (3.2) is satisfied with $c_1 = 1$. Since $\Phi_n^*(x, y) = \mathcal{T}_y(S_n\tilde\Phi_n)(x)$, (3.14), (3.15), and the first estimate in (3.13) follow from Theorem 3.1. The second inequality in (3.13) is deduced from the estimates

$$\left(\frac{1 + \cos(\theta - \varphi)}{2}\right)^n = (1-\sin^2((\theta-\varphi)/2))^n \le \exp(-n\sin^2((\theta-\varphi)/2)) \le \exp\left(-\frac{n}{\pi^2}(\theta - \varphi)^2\right), \ \theta, \varphi \in [0, \pi].$$

□

Finally, we point out two ways to construct the sequence $\mathbf{h}_n$ so that

$$\sup_{n\ge1} \|\tilde\Phi_n(\alpha, \beta; \mathbf{h}_n)\|_{\alpha,\beta;1} < \infty, \ \|\tilde\Phi_n(\alpha, \beta; \mathbf{h}_n)\|_\infty \le cn^{2\alpha+2}. \tag{3.20}$$

In [22, Lemma 4.6] (cf. also the proof of [22, Theorem 3.1]), we have proved that if $S > \max(\alpha, \beta) + 3/2$ is an integer, and $h : [0, \infty) \to [0, \infty)$ is compactly supported and can be expressed as an S times iterated integral of a compactly supported function of bounded variation, then (3.20) holds with $h_{k,n} = h(k/(4n))$. Since we do not need any localization properties on $\tilde\Phi_n$, we may also use another construction in order to obtain explicit constants in (3.20). Let $\tilde K_n(t) := K_n(1, t)$, where K_n is the Christoffel–Darboux kernel defined in (2.6), and

$$\tilde\Psi_n(t) := \frac{\tilde K_{3n}(t)\tilde K_n(t)}{\tilde K_n(1)}. \tag{3.21}$$

Kernels of this type have been considered and applied in [7, 8].

Necessarily, $\tilde\Psi_n(t) = \tilde\Phi_n(\alpha, \beta; \mathbf{h}_n)$ for a sequence $\mathbf{h}_n$ such that $h_{k,n} = 0$ if $k \ge 4n$. If $P \in \Pi_{2n}$ then $\tilde K_n P \in \Pi_{3n-1}$. Hence, the reproducing property of K_{3n} shows that

$$\int_{-1}^{1} \tilde\Psi_n(t)P(t)w_{\alpha,\beta}(t)dt = P(1).$$

Using this equation with ${p_k}^{(\alpha,\beta)}$, $k = 0, 1, \cdots, 2n$, and observing that none of the quantities ${p_k}^{(\alpha,\beta)}(1)$ is equal to 0, we conclude that $h_{k,n} = 1$, $k = 0, 1, \cdots, 2n$. We remark that for any $x \in [-1, 1]$, we may use the above equation with $\mathcal{T}_x P$ in place of P to conclude that for every $P \in \Pi_{2n}$,

$$\int_{-1}^{1} \mathcal{T}_x\tilde\Psi_n(t)P(t)w_{\alpha,\beta}(t)dt = \int_{-1}^{1} \tilde\Psi_n(t)\mathcal{T}_x P(t)w_{\alpha,\beta}(t)dt = \mathcal{T}_x P(1) = P(x). \tag{3.22}$$

Further, the reproducing property of the Christoffel–Darboux kernel and [29, Formula (4.5.8)] show that for integer $m \geq 1$,

$$\begin{aligned} \int_{-1}^{1} K_m^2(1,t) w_{\alpha,\beta}(t) dt &= K_n(1,1) \\ &= \frac{\Gamma(m+\alpha+\beta+1)\Gamma(m+\alpha+1)}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\alpha+2)\Gamma(m)\Gamma(m+\beta)} = \frac{m^{2\alpha+2}}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\alpha+2)}(1+o(1/m)). \end{aligned}$$

Hence, an application of Schwarz inequality implies that

$$\|\tilde{\Psi}_n\|_{\alpha,\beta;1} \leq \frac{\|\tilde{K}_{3n}\|_{\alpha,\beta;2}\|\tilde{K}_n\|_{\alpha,\beta;2}}{K_n(1,1)} = \left(\frac{K_{3n}(1,1)}{K_n(1,1)}\right)^{1/2} = 3^{\alpha+1}(1+o(1/n)).$$

It is known [29, Formula (4.5.8)] that $K_n(\alpha,\beta;x,1)$ is a multiple of ${P_{n-1}}^{(\alpha+1,\beta)}$. Therefore, in view of [29, Formula (7.32.2)],

$$\|\tilde{\Psi}_n\|_\infty = \tilde{\Psi}_n(1) = K_{3n}(1,1) = \frac{(3n)^{2\alpha+2}}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\alpha+2)}(1+o(1/n)).$$

Thus, (3.20) is satisfied with explicitly defined constants. Explicit bounds for the norm of the kernel were already obtained in [7].

4 Computational considerations

The coefficients $H_{k,n}$ in (3.10) can be computed in the form

$$H_{k,n} = \left({P_k}^{(\alpha,\beta)}(1)\right)^{-1} \int_{-1}^{1} \tilde{\Phi}^*(t) {P_k}^{(\alpha,\beta)}(t) w_{\alpha,\beta}(t) dt$$

using Gauss–Jacobi quadrature formula at the zeros of ${p_{5n}}^{(\alpha,\beta)}$. Algorithms for generating these quadrature formulas are given by Gautschi in [11].

In the case when the coefficients $h_{k,n}$ of $\tilde{\Phi}_n$ are known, one can also use a repeated matrix multiplication that yields all the coefficients in $\mathcal{O}(n)$ operations. We note that the orthonormalized Jacobi polynomials satisfy the recurrence relations

$$\frac{1+t}{2}{p_k}^{(\alpha,\beta)}(t) = \rho_k {p_{k+1}}^{(\alpha,\beta)}(t) + d_k {p_k}^{(\alpha,\beta)}(t) + \rho_{k-1}{p_{k-1}}^{(\alpha,\beta)}(t), \tag{4.1}$$

with the initial terms ${p_{-1}}^{(\alpha,\beta)}(t) = 0$, and

$${p_0}^{(\alpha,\beta)}(t) = \sqrt{\frac{\Gamma(\alpha+\beta+2)}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\beta+1)}}, \tag{4.2}$$

where

$$\rho_0 := \frac{1}{\alpha+\beta+2}\sqrt{\frac{(\alpha+1)(\beta+1)}{\alpha+\beta+3}},\ d_0 := \frac{1}{2} + \frac{\beta-\alpha}{2\alpha+2\beta+3}, \tag{4.3}$$

and for $k = 1, 2, \cdots$,

$$\begin{aligned} \rho_k &:= \sqrt{\frac{(k+1)(k+\alpha+1)(k+\beta+1)(k+\alpha+\beta+1)}{(2k+\alpha+\beta+1)(2k+\alpha+\beta+2)^2(2k+\alpha+\beta+3)}}, \\ d_k &:= \frac{1}{2} + \frac{\beta^2-\alpha^2}{2(2k+\alpha+\beta)(2k+\alpha+\beta+1)}. \end{aligned} \tag{4.4}$$

Let

$$J_n = \begin{pmatrix} d_0 & \rho_0 & 0 & 0 & \cdots & 0 \\ \rho_0 & d_1 & \rho_1 & 0 & \cdots & 0 \\ & \rho_1 & d_2 & \rho_2 & \cdots & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \cdots & \rho_n & d_{n+1} \end{pmatrix}.$$

The numbers $H_{k,n}{p_k}^{(\alpha,\beta)}(1)$, $k = 0, 1, \cdots, 5n$ are given by

$$J^n_{5n-2}(\{h_{k,n}{p_k}^{(\alpha,\beta)}(1)\}_{k=0}^{4n}, \mathbf{0})^T.$$

For the filter $\mathbf{h}_n$, we may choose $h_{k,n} = h(k/(4n))$ where h is the C^∞ function given by

$$h(t) := \begin{cases} 1, & \text{if } 0 \le t \le 1/2, \\ \exp\left(-\dfrac{\exp(2/(1-2t))}{1-t}\right), & \text{if } 1/2 < t < 1, \\ 0, & \text{if } t \ge 1. \end{cases}$$

In Figure 1, we show graphically, the filter $H_{k,16}$, and the kernels $\mathcal{T}_{1/2}\tilde{\Phi}(0,0;\mathbf{h}_{20})$ and $\Phi^*_{16}(0,0;\mathbf{h}_{16},x,1/2)$ (both polynomials of degree at most 79) on the whole interval and the interval $[-1,0]$. It is clear that the kernel $\Phi^*_{16}(0,0;\mathbf{h}_{16},x,1/2)$ is localized much better.

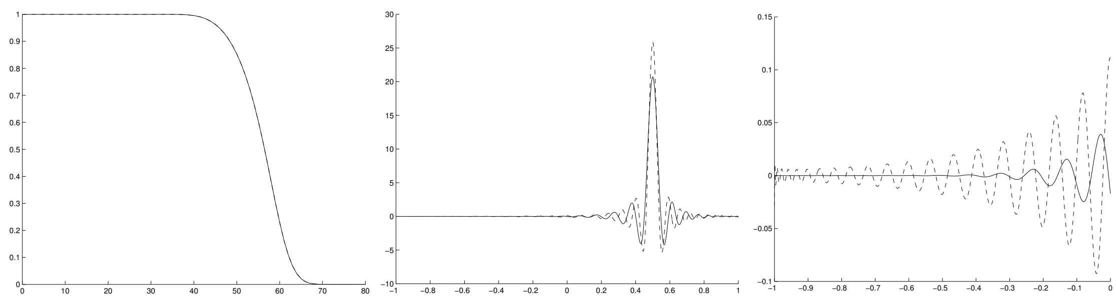

Figure 1: The filter $H_{k,16}$ on the left, the kernels $\mathcal{T}_{1/2}\tilde{\Phi}(0,0;\mathbf{h}_{20})$ (dotted line) and $\Phi^*_{16}(0,0;\mathbf{h}_{16},x,1/2)$ (continuous line) on the interval $[-1,1]$ in the middle, and on the interval $[-1,0]$ on the right. The rightmost picture shows clearly the better localization of the kernel $\Phi^*_{16}(0,0;\mathbf{h}_{16},x,1/2)$.

5 Riesz bases

We recall that if $\mathcal{H}$ is a Hilbert space, a sequence $\{g_k\} \subset \mathcal{H}$ is called a *Riesz basis* for $\mathcal{H}$ if each of the following conditions (a), (b), (c) are satisfied:

(a) The closure of span $\{g_k\}$ is equal to $\mathcal{H}$.

(b) For every sequence $\mathbf{a} = \{a_k\} \subset \mathbb{R}$ with

$$\|\mathbf{a}\|_{\ell^2} := \sqrt{\sum_{k=0}^{\infty} a_k^2} < \infty,$$

the series $\sum a_k g_k$ converges in $\mathcal{H}$,

(c) There exist constants $A, B > 0$ such that for every square summable sequence $\mathbf{a}$,

$$A\|\mathbf{a}\|_{\ell^2} \le \left\| \sum_{k=0}^{\infty} a_k g_k \right\|_{\mathcal{H}} \le B\|\mathbf{a}\|_{\ell^2}. \tag{5.1}$$

The quotient $B/A \geq 1$ describes the quality of the Riesz stability. If $B/A = 1$ the basis is an orthogonal one.

The purpose of this section is to obtain a Riesz basis for $L^2 := L^2(-1/2, -1/2)$, consisting of translates of a sequence of polynomial kernels, such that each element of this basis is localized exponentially near the point used to define the translate. We define the Chebyshev polynomial of degree k by $T_k(\cos\theta) := \cos(k\theta)$, $k = 0, 1, \cdots$. Then the orthonormalized Chebyshev polynomials are given by

$$\mathbb{T}_k(t) := \begin{cases} \pi^{-1/2} T_0(t), & \text{if } k = 0, \\ (\pi/2)^{-1/2} T_k(t), & \text{if } k = 1, 2, \cdots. \end{cases}$$

It is customary to define $\mathbb{T}_{-1}(t) = 0$. Clearly, $\{\mathbb{T}_k\}_{k=0}^{\infty}$ is an orthonormal basis for L^2, but not an exponentially localized one. Following [18, 19], we will obtain an orthogonal decomposition of L^2 into "wavelet spaces": $L^2 = V_0 \oplus \oplus_{j=0}^{\infty} W_{2^j}$. Within each W_n, we will define a Riesz basis, consisting of translates of an exponentially localized kernel Ψ_n. It is easy to check that if f has a formal Chebyshev expansion of the form $f \sim \sum_{k=0}^{\infty} f_k T_k$, then for $y \in [-1, 1]$, $\mathcal{T}_y f$ has the expansion $\sum_{k=0}^{\infty} f_k T_k(y) T_k$.

Obviously, the points for the translations must be chosen very carefully. At a critical point in our construction, we use the fact that all the polynomials $T_{6n+k} + T_{6n-k}$, $(k = 0, \cdots, 6n)$, $T_{12n-k} + T_k$, and $T_{12n+k} + T_k$ $(0 \leq k \leq 12n)$ have T_{6n} as a common factor. For this reason, we have to restrict ourselves to the case $\alpha = \beta = -1/2$.

For the purpose of defining a wavelet-type decomposition, it is convenient to define a kernel having a degree which is a multiple of 8. Also, for the case of Chebyshev polynomials, one can use a piecewise linear function to generate the filter $h_{k,n}$ in (3.1). More than the smoothness, the symmetry of this filter is important for our purpose here. Accordingly, we define

$$a_n(k) := \begin{cases} 0 & \text{if } k < 0, \\ 1/2, & \text{if } k = 0, \\ 1, & \text{if } 1 \leq k \leq 5n, \\ (7n - k)/(2n), & \text{if } 5n + 1 \leq k \leq 7n - 1, \\ 0, & \text{if } k \geq 7n, \end{cases}$$

and define the analogue of the kernel $\tilde{\Phi}_n$ in (3.1) by

$$\tilde{G}_n(t) := \sum_{k=0}^{7n-1} a_n(k) T_k(t).$$

We note that $a_n(0) = 1/2$ rather than 1 to account for the different normalizations. In particular, $\tilde{G}_n * P = P$ for any $P \in \Pi_{5n}$. We define the exponentially localized kernel as in Section 3 by

$$\tilde{G}_n^*(t) := \left(\frac{1+t}{2}\right)^n \tilde{G}_n(t) =: \sum_{k=0}^{8n-1} b_n(k) T_k(t), \qquad t \in [-1, 1]. \tag{5.2}$$

The following lemma summarizes some of the properties of the coefficients $b_n(k)$, which will be needed for our construction of the basis. We adopt the convention that $\binom{2m}{\ell} = 0$ if $\ell \notin [0, 2m]$.

Proposition 5.1 *Let $n, m \geq 1$ be integers, $\mathbf{d} = \{d_k\}_{k=0}^{\infty}$ be a sequence with all $d_k = 0$ for sufficiently large k, $P(t) = \sum_{k=0}^{\infty} d_k T_k(t)$.*
(a) *With*

$$\tilde{d}_{k,m} := \begin{cases} \displaystyle\sum_{\ell=0}^{m} \binom{2m}{m+\ell} d_\ell, & \text{if } k = 0, \\ \displaystyle\binom{2m}{m+k} d_0 + \sum_{\ell=1}^{m} \binom{2m}{m+\ell} (d_{k+\ell} + d_{|\ell-k|}) + \binom{2m}{m} d_k, & \text{if } k \geq 1, \end{cases} \tag{5.3}$$

we have

$$2^m (1+t)^m P(t) = \sum_{k=0}^{\infty} \tilde{d}_{k,m} T_k(t). \tag{5.4}$$

In particular,

$$4^n b_n(k) = \begin{cases} \sum_{\ell=0}^{n} \binom{2n}{n+\ell} a_n(\ell), & \text{if } k=0, \\ \binom{2n}{n+k} a_n(0) + \sum_{\ell=1}^{n} \binom{2n}{n+\ell}(a_n(k+\ell)+a_n(|\ell-k|)) + \binom{2n}{n} a_n(k), & \text{if } k \geq 1, \end{cases} \tag{5.5}$$

$$= \begin{cases} 4^n/2, & \text{if } k=0 \\ 4^n, & \text{if } k=1,\cdots,4n, \\ \sum_{\ell=0}^{2n} \binom{2n}{\ell} a_n(k-n+\ell), & \text{if } k \geq n+1. \end{cases} \tag{5.6}$$

(b) *For all integers k, $0 \leq b_n(k) \leq 1$. For $k \geq 1$, $b_n(k) \geq b_n(k+1)$. For $0 \leq k \leq 2n$, $b_n(6n+k) = 1 - b_n(6n-k)$. In particular, $b_n(6n) = 1/2$ and $b_n(k) \geq 1/2$ for $0 \leq k \leq 6n$.*

Proof. We observe that with $t = \cos\theta$

$$\begin{aligned} (2+2t)^m &= (2+2\cos\theta)^m = e^{-im\theta}(1+e^{i\theta})^{2m} = \binom{2m}{m} + 2\sum_{\ell=1}^{m} \binom{2m}{m+\ell}\cos(\ell\theta) \\ &= \binom{2m}{m} + 2\sum_{\ell=1}^{m} \binom{2m}{m+\ell} T_\ell(t), \end{aligned} \tag{5.7}$$

and that $2T_k(t)T_\ell(t) = T_{k+\ell}(t) + T_{|k-\ell|}(t)$. Using these identities, we calculate that

$$(2+2t)^m P(t) = S_1 + S_2 + S_3, \tag{5.8}$$

where, in this proof only,

$$\begin{aligned} S_1 &:= \sum_{\ell=0}^{m}\sum_{k=0}^{\infty} \binom{2m}{m+\ell} d_k T_{k+\ell}(t), \\ S_2 &:= \sum_{\ell=1}^{m}\sum_{k=0}^{\ell-1} \binom{2m}{m+\ell} d_k T_{\ell-k}(t), \\ S_3 &:= \sum_{\ell=1}^{m}\sum_{k=\ell}^{\infty} \binom{2m}{m+\ell} d_k T_{k-\ell}(t). \end{aligned}$$

With our convention about the combinatorial symbols,

$$\begin{aligned} S_1 &= \sum_{\ell=0}^{m}\sum_{k=\ell}^{\infty} \binom{2m}{m+\ell} d_{k-\ell} T_k(t) = \sum_{\ell=0}^{\infty}\sum_{k=\ell}^{\infty} \binom{2m}{m+\ell} d_{k-\ell} T_k(t) \\ &= \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \binom{2m}{m+\ell} d_{k-\ell} T_k(t) = \binom{2m}{m} d_0 T_0(t) + \sum_{k=1}^{\infty}\sum_{\ell=0}^{k} \binom{2m}{m+\ell} d_{k-\ell} T_k(t). \end{aligned} \tag{5.9}$$

Replacing $\ell - k$ by k in S_2, we obtain

$$\begin{aligned} S_2 &= \sum_{\ell=1}^{m}\sum_{k=1}^{\ell} \binom{2m}{m+\ell} d_{\ell-k} T_k(t) = \sum_{k=1}^{m}\sum_{\ell=k}^{m} \binom{2m}{m+\ell} d_{\ell-k} T_k(t) \\ &= \sum_{k=1}^{m} \binom{2m}{m+k} d_0 T_k(t) + \sum_{k=1}^{m}\sum_{\ell=k+1}^{m} \binom{2m}{m+\ell} d_{\ell-k} T_k(t). \end{aligned} \tag{5.10}$$

Similar changes of variables and interchange of order of summation yield

$$\begin{aligned} S_3 &= \sum_{\ell=1}^{m}\binom{2m}{m+\ell}d_\ell T_0(t)+\sum_{\ell=1}^{m}\sum_{k=\ell+1}^{\infty}\binom{2m}{m+\ell}d_k T_{k-\ell}(t)\\ &= \sum_{\ell=1}^{m}\binom{2m}{m+\ell}d_\ell T_0(t)+\sum_{\ell=1}^{m}\sum_{k=1}^{\infty}\binom{2m}{m+\ell}d_{k+\ell}T_k(t)\\ &= \sum_{\ell=1}^{m}\binom{2m}{m+\ell}d_\ell T_0(t)+\sum_{k=1}^{\infty}\sum_{\ell=1}^{m}\binom{2m}{m+\ell}d_{k+\ell}T_k(t). \end{aligned} \tag{5.11}$$

We substitute the values of S_1, S_2, S_3 from (5.9), (5.10), (5.11) into (5.8) and combine the terms to obtain (5.4) with the coefficients as in (5.3).

The equation (5.5) is the same as (5.3) with $m = n$ and $d_k = a_n(k)$. In (5.6), we prove first the third part, then the first part, and finally, the second part. If $k \ge n+1$, then our convention on the combinatorial symbols implies that

$$\begin{aligned} 4^n b_n(k) &= \sum_{\ell=1}^{n}\binom{2n}{n+\ell}(a_n(k+\ell)+a_n(k-\ell))+\binom{2n}{n}a_n(k)\\ &= \sum_{1\le|\ell|\le n}\binom{2n}{n+\ell}a_n(k+\ell)+\binom{2n}{n}a_n(k)=\sum_{\ell=-n}^{n}\binom{2n}{n+\ell}a_n(k+\ell)\\ &= \sum_{\ell=0}^{2n}\binom{2n}{\ell}a_n(k-n+\ell). \end{aligned}$$

This proves the third part of (5.6).

Next, putting $t = 1$ in (5.7), we see that

$$\binom{2n}{n}+2\sum_{\ell=1}^{n}\binom{2n}{n+\ell}=\sum_{\ell=0}^{2n}\binom{2n}{\ell}=4^n. \tag{5.12}$$

Since $a_n(\ell) = 1$ for $1 \le \ell \le n$, and $a_n(0) = 1/2$, we get

$$4^n b_n(0)=\frac{1}{2}\binom{2n}{n}+\sum_{\ell=1}^{n}\binom{2n}{n+\ell}=\frac{1}{2}4^n.$$

If $1 \le k \le n$, then we note that $a(k+\ell) = 1$ for all ℓ, $0 \le \ell \le n$, $a_n(|\ell-k|) = 1$ if $\ell \ne k$, $0 \le \ell \le n$. So, in view of (5.12),

$$\begin{aligned} 4^n b_n(k) &= \binom{2n}{n+k}a_n(0)+\sum_{\ell=1}^{n}\binom{2n}{n+\ell}(a_n(k+\ell)+a_n(|\ell-k|))+\binom{2n}{n}a_n(k)\\ &= \binom{2n}{n+k}+\sum_{\ell=1}^{n}\binom{2n}{n+\ell}+\sum_{\substack{\ell=1\\ \ell\ne k}}^{n}\binom{2n}{n+\ell}+\binom{2n}{n}\\ &= \binom{2n}{n}+2\sum_{\ell=1}^{n}\binom{2n}{n+\ell}=4^n. \end{aligned}$$

If $n+1 \le k \le 4n$, then for $\ell = 0, \cdots, 2n$, $1 \le k-n+\ell \le 5n$, and $a_n(k-n+\ell) = 1$. Hence, the third part of (5.6) and (5.12) imply that $4^n b_n(k) = 4^n$.

In view of the fact that $0 \le a_n(k) \le 1$ for all k, (5.6) and (5.12) imply that $0 \le b_n(k) \le 1$ for all integer $k \ge 0$. We note that $b_n(k) = b_n(k+1)$ if $1 \le k \le n$. If $k \ge n+1$, then we observe that $a_n(k-n+\ell) \ge a_n(k+1-n+\ell)$ for all $\ell = 0, \cdots, 2n$. Hence, the third formula in (5.6) shows that

$$b_n(k)=4^{-n}\sum_{\ell=0}^{2n}\binom{2n}{\ell}a_n(k-n+\ell)\ge 4^{-n}\sum_{\ell=0}^{2n}\binom{2n}{\ell}a_n(k+1-n+\ell)=b_n(k+1).$$

To prove the last assertion in part (b), we note that that for $-n \le j \le 3n$, $6n+j, 6n-j > n$ and $a_n(6n-j) = 1 - a_n(6n+j)$. Moreover, for $0 \le k \le 2n$, $0 \le \ell \le 2n$, $-n \le k+n-\ell \le 3n$.

So, using the third formula in (5.6) and (5.12), we conclude that for $0 \le k \le 2n$,

$$4^n b_n(6n-k) = \sum_{\ell=0}^{2n} \binom{2n}{\ell} a_n(6n-(k+n-\ell)) = 4^n - \sum_{\ell=0}^{2n} \binom{2n}{\ell} a_n(6n+(k+n-\ell)) = 4^n - 4^n b_n(6n+k).$$

□

To define the wavelet spaces, we first denote the zeros of the Chebyshev polynomial T_{6n} by

$$z_{s,n} = \cos\left(\frac{(2s-1)\pi}{12n}\right), \qquad s = 1, \cdots, 6n.$$

The wavelet spaces will be W_{2^j}, $j = 1, 2, \cdots$, where the space W_n is defined by translates of the kernel $G^*_{2n} - G^*_n$ at these zeros:

$$W_n := \text{span}\left\{\sum_{k=0}^{16n-1} b_{2n}(k) T_k(z_{s,n}) T_k - \sum_{k=0}^{8n-1} b_n(k) T_k(z_{s,n}) T_k =: \Psi_{n,s}\right\}. \tag{5.13}$$

The following proposition gives another description for the spaces W_n, which makes it clear that W_{2n} is orthogonal to W_n, and also that the dimension of each W_n is $6n$. For reasons which will be clear in the statement of Proposition 5.2 below, we write in the remainder of this section,

$$\mathbb{P}_{n,k}(t) = \begin{cases} -(1/2)T_{12n}(t), & \text{if } k = 0, \\ -b_{2n}(12n-k)T_{12n-k}(t) - b_{2n}(12n+k)T_{12n+k}(t), & \text{if } k = 1, \cdots, 4n-1, \\ -T_{8n}(t), & \text{if } k = 4n, \\ -b_n(k)T_{12n-k}(t) + b_n(12n-k)T_k(t), & \text{if } k = 4n+1, \cdots, 6n-1. \end{cases}$$

Proposition 5.2 *Let $n \ge 1$, $1 \le s \le 6n$ be integers. Then for $t \in [-1, 1]$,*

$$\begin{aligned} \Psi_{n,s}(t) &= -\frac{1}{2}T_{12n}(t) - \sum_{k=1}^{4n-1} T_k(z_{s,n}) \left\{b_{2n}(12n-k)T_{12n-k}(t) + b_{2n}(12n+k)T_{12n+k}(t)\right\} \\ &\quad - T_{4n}(z_{s,n})T_{8n}(t) - \sum_{k=1}^{2n-1} T_{6n-k}(z_{s,n}) \left\{b_n(6n-k)T_{6n+k}(t) - b_n(6n+k)T_{6n-k}(t)\right\}, \\ &= \sum_{k=0}^{6n-1} T_k(z_{s,n}) \mathbb{P}_{n,k}(t), \end{aligned} \tag{5.14}$$

and

$$\mathbb{P}_{n,k}(t) = \frac{1}{3n} \sum_{s=1}^{6n} T_k(z_{s,n}) \Psi_{n,s}(t), \qquad k = 0, \cdots, 6n-1. \tag{5.15}$$

In particular, the dimension of W_n is $6n$ and $W_n \perp W_{2n}$.

The main tools in our proof of Proposition 5.2 are Proposition 5.1, the easily verified relations

$$T_{6n-k}(z_{s,n}) = -T_{6n+k}(z_{s,n}),\ T_{12n-k}(z_{s,n}) = T_{12n+k}(z_{s,n}) = -T_k(z_{s,n}), \quad k = 0, \cdots, 6n, \tag{5.16}$$

and the Chebyshev quadrature formula [5, Formula (7.45), Section 7.3]:

$$\int_{-1}^{1} P(t)(1-t^2)^{-1/2} dt = \frac{\pi}{6n} \sum_{s=1}^{6n} P(z_{s,n}), \qquad P \in \Pi_{12n-1}.$$

In particular, the last formula implies that

$$\sum_{s=1}^{6n} T_k(z_{s,n})T_\ell(z_{s,n}) = \begin{cases} 6n, & \text{if } k=\ell=0, \\ 3n, & \text{if } k=\ell,\ k,\ell=1,\cdots,6n-1, \\ 0, & \text{if } k\neq\ell,\ k,\ell=0,\cdots,6n-1. \end{cases} \tag{5.17}$$

Proof of Proposition 5.2. In light of (5.6), we have

$$G_n^*(t) = \frac{1}{2} + \sum_{k=1}^{4n} T_k(t) + \sum_{k=4n+1}^{8n-1} b_n(k)T_k(t).$$

Using this expression also for G_{2n}^*, we conclude using the facts that $b_{2n}(8n)=1$, $T_{8n}(z_{s,n}) = -T_{4n}(z_{s,n})$, that

$$\begin{aligned} \Psi_{n,s}(t) &= b_{2n}(8n)T_{8n}(z_{s,n})T_{8n}(t) + \sum_{k=8n+1}^{16n-1} b_{2n}(k)T_k(z_{s,n})T_k(t) + \sum_{k=4n+1}^{8n-1}(1-b_n(k))T_k(z_{s,n})T_k(t) \\ &= -T_{4n}(z_{s,n})T_{8n}(t) + \sum_{k=8n+1}^{16n-1} b_{2n}(k)T_k(z_{s,n})T_k(t) + \sum_{k=4n+1}^{8n-1}(1-b_n(k))T_k(z_{s,n})T_k(t). \end{aligned} \tag{5.18}$$

Using the facts that $b_{2n}(12n)=1/2$, $T_{12n}(z_{s,n})=-1$, and (5.16), we deduce that

$$\begin{aligned} &\sum_{k=8n+1}^{16n-1} b_{2n}(k)T_k(z_{s,n})T_k(t) = -\frac{1}{2}T_{12n}(t) \\ &\qquad + \sum_{k=1}^{4n-1}\{b_{2n}(12n-k)T_{12n-k}(z_{s,n})T_{12n-k}(t) + b_{2n}(12n+k)T_{12n+k}(z_{s,n})T_{12n+k}(t)\} \\ &= -\frac{1}{2}T_{12n}(t) - \sum_{k=1}^{4n-1} T_k(z_{s,n})\{b_{2n}(12n-k)T_{12n-k}(t) + b_{2n}(12n+k)T_{12n+k}(t)\}. \end{aligned} \tag{5.19}$$

Since $T_{6n}(z_{s,n})=0$ for $s=1,\cdots,6n$, we may use Proposition 5.1(b) and and (5.16) to obtain

$$\begin{aligned} \sum_{k=4n+1}^{8n-1}(1-b_n(k))T_k(z_{s,n})T_k(t) &= \sum_{k=1}^{2n-1}\{(1-b_n(6n-k))T_{6n-k}(z_{s,n})T_{6n-k}(t) \\ &\qquad +(1-b_n(6n+k))T_{6n+k}(z_{s,n})T_{6n+k}(t)\} \\ &= -\sum_{k=1}^{2n-1} T_{6n-k}(z_{s,n})\{b(6n-k)T_{6n+k}(t) - b(6n+k)T_{6n-k}(t)\}. \end{aligned} \tag{5.20}$$

The first equation (5.14) follows from (5.18), (5.19), and (5.20). The second equation in (5.14) is just a change of indexing in the last summation.

Using (5.17), it is not difficult to deduce the equations (5.15) from (5.14). Since the system of functions on the left hand side of (5.15) are clearly orthogonal, the equations (5.15) and (5.14) show that the dimension of W_n is $6n$. Moreover, W_n is the span of the system of functions on the left hand side of (5.15). This leads to the assertion that $W_{2n} \perp W_n$. □

Next, we describe the interpolatory properties and the Riesz bounds for $\{\Psi_{n,s}\}$ as a basis for W_n, and its localization.

Proposition 5.3 *Let $n \geq 1$ be an integer. We have for $s,\ell = 1,\cdots,6n$, and $\theta\in[0,\pi]$,*

$$\Psi_{n,s}(z_{\ell,n}) = \begin{cases} 3n, & \textit{if } k=\ell, \\ 0, & \textit{if } k\neq\ell, \end{cases} \tag{5.21}$$

$$|\Psi_{n,s}(\cos\theta)| \leq 21n\left(\frac{1+\cos(\theta-(2s-1)\pi/(12n))}{2}\right)^n. \tag{5.22}$$

If $d_s \in \mathbb{R}$, $s = 1, \cdots, 6n$, *then*

$$\sqrt{\frac{3n\pi}{4}}\|\{d_s\}\|_{\ell^2} \le \|\sum_{s=1}^{6n} d_s\Psi_{n,s}\|_{-1/2,-1/2;2} \le \sqrt{\frac{3n\pi}{2}}\|\{d_s\}\|_{\ell^2}. \tag{5.23}$$

PROOF. We use (5.16), and the fact that $b_n(6n-k)+b_n(6n+k)=1$ from Proposition 5.1(b), to obtain from (5.14) that

$$\begin{aligned}\Psi_{n,s}(z_{\ell,n}) &= \frac{1}{2}+\sum_{k=1}^{4n-1}T_k(z_{s,n})T_k(z_{\ell,n})\left\{b_{2n}(12n-k)+b_{2n}(12n+k)\right\}\\ &\quad +T_{4n}(z_{s,n})T_{4n}(z_{\ell,n})+\sum_{k=1}^{2n-1}T_{6n-k}(z_{s,n})T_{6n-k}(z_{\ell,n})\left\{b_n(6n-k)+b_n(6n+k)\right\}\\ &= \frac{1}{2}+\sum_{k=1}^{6n-1}T_k(z_{s,n})T_k(z_{\ell,n}).\end{aligned}$$

This leads to (5.21) by using the identity

$$\frac{1}{2}+\sum_{k=1}^{6n-1}T_k(t)T_k(y)=\frac{1}{2}\begin{cases}\dfrac{T_{6n}(t)T_{6n-1}(y)-T_{6n-1}(t)T_{6n}(y)}{t-y}, & \text{if } t\neq y,\\ T'_{6n}(t)T_{6n-1}(t)-T'_{6n-1}(t)T_{6n}(t), & \text{if } t=y.\end{cases}$$

Since $|\tilde{G}_n(t)| \le 7n$, the estimate (5.22) follows Lemma 3.1 as in the proof of Theorem 3.2.

In this proof only, we define the (column) vector valued functions $\mathbf{\Psi} = (\Psi_{n,s})_{s=1}^{6n}$ and $\mathbf{P} = (\mathbb{P}_{n,k})_{k=0}^{6n-1}$. Further, in this proof only, we define the following $6n\times 6n$ matrices: $\mathbf{B} = (T_k(z_{s,n}))_{s=1,k=0}^{6n,6n-1}$, and

$$\mathbf{D} = \sqrt{\pi/2}\,\mathrm{diag}(1/2, (\sqrt{b_{2n}(12n-k)^2+b_{2n}(12n+k)^2})_{k=1}^{4n-1}, 1, (\sqrt{b_n(k)^2+b_n(12n-k)^2})_{k=4n+1}^{6n-1}).$$

The second equation in (5.14) can be expressed in the form $\mathbf{\Psi} = (\mathbf{BD})\mathbf{D}^{-1}\mathbf{P}$. The equations (5.17) take the form

$$\mathbf{B}^T\mathbf{B} = \mathrm{diag}(6n, 3n, \cdots, 3n). \tag{5.24}$$

It is clear that $\mathbf{D}^{-1}\mathbf{P}$ is a vector of orthonormalized polynomials. In this proof only, let $\mathbf{d}$ be the column vector (d_s). Then one can verify by a simple computation taking (5.24) into account that

$$\begin{aligned}\|\sum_s d_s\Psi_{n,s}\|^2_{-1/2,-1/2;2} &= \mathbf{d}^T(\mathbf{BD})^T(\mathbf{BD})\mathbf{d} = \mathbf{d}^T\mathbf{D}\mathbf{B}^T\mathbf{B}\mathbf{D}\mathbf{d}\\ &= (3n\pi/2)\mathbf{d}^T\mathrm{diag}(1/2, (b_{2n}(12n-k)^2+b_{2n}(12n+k)^2)_{k=1}^{4n-1}, 1, (b_n(k)^2+b_n(12n-k)^2)_{k=4n+1}^{6n-1})\mathbf{d}.\end{aligned} \tag{5.25}$$

Since $b_{2n}(12n-k)+b_{2n}(12n+k)=1$, $k=1,\cdots,4n-1$, and $0\le b_{2n}(k)\le 1$, $k=0,\cdots,12n-1$, we see that for $k=1,\cdots,4n-1$,

$$1/2 = (1/2)(b_{2n}(12n-k)+b_{2n}(12n+k)) \le b_{2n}(12n-k)^2+b_{2n}(12n+k)^2 \le b_{2n}(12n-k)+b_{2n}(12n+k) = 1.$$

Similarly, for $k=4n+1,\cdots,6n-1$,

$$1/2 \le b_n(k)^2+b_n(12n-k)^2 \le 1.$$

Consequently, (5.25) leads to (5.23). □

We have now enough preparation to define the basis for $L^2(-1/2,-1/2)$ as in (5.28) below. However, we take a small detour to discuss the basis of W_n dual to $\Psi_{n,s}$.

Proposition 5.4 *Let $n \geq 1$ be an integer. For $t \in [-1,1]$, $\ell = 1, \cdots, 6n$, let*

$$\Psi^D_{n,\ell}(t) := \frac{1}{3n}\sum_{k=0}^{6n-1} T_k(z_{\ell,n})\|\mathbb{P}_{n,k}\|^{-2}_{-1/2,-1/2;2}\mathbb{P}_{n,k}(t). \tag{5.26}$$

Then for $s, \ell = 1, \cdots, 6n$,

$$\int_{-1}^{1} \Psi_{n,s}(t)\Psi^D_{n,\ell}(t)(1-t^2)^{-1/2}dt = \begin{cases} 1, & \text{if } s = \ell, \\ 0, & \text{otherwise.} \end{cases} \tag{5.27}$$

Proof. Let

$$\mathbf{D}_1 = 3n\sqrt{\pi/2}\,\mathrm{diag}(1, (\sqrt{b_{2n}(12n-k)^2 + b_{2n}(12n+k)^2})_{k=1}^{4n-1}, 1, (\sqrt{b_n(k)^2 + b_n(12n-k)^2})_{k=4n+1}^{6n-1})$$

and

$$\mathbf{\Psi}^D := (\Psi^D_{n,s})_{s=1}^{6n} := \mathbf{B}\mathbf{D}_1^{-1}\mathbf{D}^{-1}\mathbf{P},$$

where $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{P}$ are matrices introduced in the proof of Proposition 5.3. The presence of the matrix $\mathbf{B}$ ensures that $\mathbf{\Psi}^D$ is a vector of generalized translates of the function given by the sum of the elements of the vector $\mathbf{D}_1^{-1}\mathbf{D}^{-1}\mathbf{P}$.

The proof of the duality of the bases follows directly from the fact that $\mathbf{D}^{-1}\mathbf{P}$ is a vector of orthonormalized polynomials and that

$$(\mathbf{B}\mathbf{D}_1^{-1})^T\mathbf{B}\mathbf{D} = \mathbf{D}_1^{-1}\mathbf{B}^T\mathbf{B}\mathbf{D}$$

which is the identity matrix of order $6n$.

□

It appears from Figure 2 that the dual basis $\{\Psi^D_{n,\ell}\}$ is also well localized, although less so than the basis $\{\Psi_{n,\ell}\}$ itself. The study of the precise localization properties of the dual basis remains an open question.

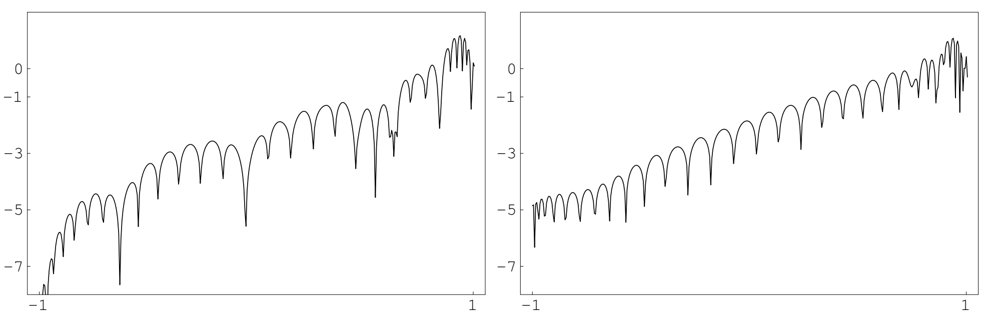

Figure 2: The logplot of $|\Psi_{5,1}|$ on the left, $|\Psi^D_{5,1}|$ on the right.

We now resume our main discussion to define the basis functions by

$$\mathbb{B}_\ell = \begin{cases} \mathbb{T}_\ell & \text{for } \ell \in \{0,1,2,3,4,6\}, \\ \dfrac{7\mathbb{T}_5 + \mathbb{T}_7}{\sqrt{50}} & \text{for } \ell = 5, \\ \dfrac{2^{3/4}}{\sqrt{3\cdot 2^j\pi}}\Psi_{2^j,s} & \text{for } \ell > 6 \text{ and } j, s \text{ uniquely determined by} \\ & \ell = 6\cdot 2^j + s \text{ with } 1 \leq s \leq 6\cdot 2^j,\ j = 0,1,2,\cdots. \end{cases} \tag{5.28}$$

Theorem 5.1 *The functions $\{\mathbb{B}_\ell\}_{\ell=0}^\infty$ are an exponentially localized Riesz basis, consisting of polynomials, for $L^2(-1/2,-1/2)$. For any $\mathbf{d} = \{d_k\}_{k=0}^\infty \in \ell^2$, we have*

$$2^{-1/4}\|\mathbf{d}\|_{\ell^2} \leq \|\sum_{\ell=0}^\infty d_\ell \mathbb{B}_\ell\|_{-1/2,-1/2;2} \leq 2^{1/4}\|\mathbf{d}\|_{\ell^2}. \tag{5.29}$$

Proof. In this proof only, let $V_0 = \text{span}\,\{\mathbb{T}_\ell, \ell = 0, 1, 2, 3, 4, 6, \frac{7\mathbb{T}_5+\mathbb{T}_7}{\sqrt{50}}\}$. Then it is easy to check that $V_0 \perp W_{2^j}$, $j = 0, 1, \cdots$, and the indicated polynomials form an orthonormal basis for V_0. In particular, $L^2(-1/2, -1/2) = V_0 \oplus \oplus_{j=0}^{\infty} W_{2^j}$. The estimates (5.29) follow immediately from (5.23). The fact that the functions are exponentially localized follows from (5.22). □

In the same way as we got the dual basis for $\{\mathbb{B}_\ell\}$ by replacing $\mathbf{\Psi} = \mathbf{BP}$ by $\mathbf{\Psi}^D = \mathbf{BD}_1^{-1}\mathbf{D}^{-1}\mathbf{P}$, we obtain an orthonormal basis of translates by $\mathbf{\Psi}^O = \mathbf{BD}_1^{-1/2}\mathbf{D}^{-1/2}\mathbf{P}$. However, the precise localization properties of this orthonormal basis remains also an open question.

Moreover, we conjecture that the system $\{\mathbb{B}_\ell\}_{\ell=0}^{\infty}$ is actually a Schauder basis for $C([-1, 1])$ (cf. [13, 18]).

References

[1] R. Askey, "Orthogonal Polynomials and Special Functions", Regional Conference Series in Applied Mathematics, **21**, SIAM Philadelphia, 1975.

[2] J. Bergh and J. Löfström, "Interpolation Spaces, an Introduction", Springer Verlag, Berlin, 1976.

[3] W. R. Bloom and H. Heyer, "Harmonic Analysis of Probability Measures on Hypergroups", deGruyter, Berlin, 1994.

[4] R. Archibald and A. Gelb, *A method to reduce the Gibbs ringing artifact in MRI while keeping tissue boundary integrity,* IEEE Trans. Med. Imag. **21** (4) (2002), 305–319.

[5] S. D. Conte and C. de Boor, "Elementary Numerical Analysis", McGraw Hill, Boston, 1980.

[6] K. S. Eckhoff, *Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions,* Math. Comp., **64** (1995), 671–690.

[7] F. Filbir and W. Themistoclakis, *On the construction of de la Vallée Poussin means for orthogonal polynomials using convolution structures,* J. Comput. Anal. Appl., **6** (4), (2004), 297–312.

[8] F. Filbir and W. Themistoclakis, *Polynomial approximation on the sphere using scattered data,* Math. Nachr. **281** (5) (2008), 650–668.

[9] G. Freud, "Orthogonal Polynomials", Académiai Kiadó, Budapest, 1971.

[10] D. Gaier, *Polynomial approximation of piecewise analytic functions,* J. Anal., **4** (1996), 67–79.

[11] W. Gautschi, "Orthogonal Polynomials: Computation and Approximation", Oxford University Press, 2004.

[12] A. Gelb and E. Tadmor, *Detection of edges in spectral data,* Appl. Comput. Harmon. Anal. **7** (1999), no. 1, 101–135.

[13] R. Girgensohn and J. Prestin, *Lebesgue constants for an orthogonal polynomial Schauder basis,* J. Comput. Anal. Appl. **2** (2000), 159-175.

[14] D. Gottlieb, B. Gustafsson, and P. Forssen, *On the direct Fourier method for computed tomography,* IEEE Trans. Med. Imag. **19** (3) (2000), 223–233.

[15] M. E. H. Ismail, "Classical and Quantum Orthogonal Polynomials in One Variable", Cambridge University Press, 2005.

[16] K. G. Ivanov, P. Petrushev, and Y. Xu, *Sub–exponentially localized kernels and frames induced by orthogonal expansions,* arXiv:0809.3421v1 [Math: CA] 19 Sep 2008.

[17] K. G. Ivanov, E. B. Saff, and V. Totik, *Approximation by polynomials with locally geometric rates,* Proc. Amer. Math. Soc., **106** (1989), 153-161.

[18] T. Kilgore, J. Prestin, and K. Selig, *Orthogonal algebraic polynomial Schauder bases of optimal degree,* J. Fourier Anal. Appl., **2** (6) (1996), 597–610.

[19] T. Kilgore, J. Prestin, and K. Selig, *Polynomial wavelets and wavelet packet bases,* Stud. Sci. Math. Hung. **33** (4) (1997), 419-431.

[20] T. Koornwinder, *Jacobi polynomials II. An analytic proof of the product formula,* SIAM J. Math. Anal., **5** (1974), 125–137.

[21] M. Maggioni and H. N. Mhaskar, *Diffusion polynomial frames on metric measure spaces,* Appl. Comput. Harmon. Anal. **24** (3) (2008), 329–353.

[22] H. N. Mhaskar, *Polynomial operators and local smoothness classes on the unit interval,* J. Approx. Theory, **131** (2004), 243-267.

[23] H. N. Mhaskar and D. V. Pai, "Fundamentals of Approximation Theory", CRC Press, 2000.

[24] H. N. Mhaskar and J. Prestin, *On the detection of singularities of a periodic function,* Adv. Comput. Math., **12** (2000), 95–131.

[25] H. N. Mhaskar and J. Prestin, *On local smoothness classes of periodic functions,* J. Fourier Anal. Appl., **11** (3) (2005), 353 - 373.

[26] H. N. Mhaskar and J. Prestin, *Polynomial frames: a fast tour,* in Approximation theory XI: Gatlinburg 2004 (C. K. Chui, M. Neamtu, and L. Schumaker Eds.), Mod. Methods Math., Nashboro Press, Brentwood, TN, 2005, 287–318.

[27] H. N. Mhaskar and J. Prestin, *Polynomial operators for spectral approximation of piecewise analytic functions,* Appl. Comput. Harmon. Anal. **26** (2009), 121–142.

[28] E. B. Saff and V. Totik, "Logarithmic Potentials with External Fields", Springer–Verlag, New York/Berlin, 1997.

[29] G. Szegö, "Orthogonal Polynomials", Amer. Math. Soc. Colloq. Publ. **23**, Amer. Math. Soc., Providence, 1975.

[30] E. Tadmor, *Convergence of spectral methods for nonlinear conservation laws,* SIAM J. Numer. Anal. **26** (1989), 30–44.

[31] E. Tadmor and J. Tanner, *Adaptive filters for piecewise smooth spectral data,* IMA J. Numer. Anal. **25** (4) (2005), 635–647.

[32] J. Tanner, *Optimal filter and mollifier for piecewise smooth spectral data,* Math. Comp. **75** (254) (2006), 767–790.

[33] M. Wei, A. G. Martinez, and A. R. De Pierro, *Detection of edges from spectral data: new results,* Appl. Comput. Harmon. Anal. **22** (2007), 386–393.