Polynomial operators for spectral approximation of piecewise analytic functions

H. N. Mhaskar*
Department of Mathematics, California State University
Los Angeles, California, 90032, USA
email: hmhaska@calstatela.edu

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

Abstract

We propose the construction of a *mixing filter* for the detection of analytic singularities and an *auto-adaptive* spectral approximation of piecewise analytic functions, given either spectral or pseudo-spectral data, without knowing the location of the singularities beforehand. We define a polynomial frame with the following properties. At each point on the interval, the behavior of the coefficients in our frame expansion reflects the regularity of the function at that point. The corresponding approximation operators yield an exponentially decreasing rate of approximation in the vicinity of points of analyticity and a near best approximation on the whole interval. Unlike previously known results on the construction of localized polynomial kernels, we suggest a very simple idea to obtain exponentially localized kernels based on a general system of orthogonal polynomials, for which the Cesàro means of some order are uniformly bounded. The boundedness of these means is known in a number of cases, where no special function properties are known.

1 Introduction

For a 2π–periodic function $f:\mathbb{R}\to\mathbb{R}$, integrable on $[-\pi,\pi]$, the (trigonometric) Fourier coefficients and Fourier projection $S_n(f)$ are defined by

$$\hat{f}(k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\exp(-ikt)dt,\ \ S_n(f,x)=\sum_{|k|\le n}\hat{f}(k)\exp(ikx).$$

It is well known that for a globally smooth function f, $S_n(f)$ approximates f at a spectral rate; for example, if f is analytic on $[-\pi,\pi]$, then $\|f-S_n(f)\|_\infty\le c\rho^n$ for some constant $c=c(f)>0$ and $\rho=\rho(f)\in(0,1)$. This rate of approximation changes drastically if f has any singularity, for example, a jump discontinuity in its first derivative, on $[-\pi,\pi]$.

The problem of detection of singularities of a function from spectral data arises in many important applications, for example, computer tomography [10], nuclear magnetic resonance inversion [2], and conservation laws in differential equations [31]. A closely related problem is to obtain spectral approximation of piecewise smooth or analytic functions. These problems are studied by many authors; some recent references are [33, 32, 34], and references therein. Typically, one finds first the location of singularities

*The research of this author was supported, in part, by grant DMS-0605209 from the National Science Foundation and grant W911NF-04-1-0339 from the U.S. Army Research Office.

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 function $h : [0,\infty) \to [0,\infty)$, and the corresponding *mollifier* is given by

$$\Phi_n(h,x) = \sum_{k\in\mathbb{Z}} h(|k|/n)\exp(ikx), \ \ \sigma_n(h,f,x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\Phi_n(h,x-t)dt.$$

We note that in classical harmonic analysis parlance, Φ_n (respectively, σ_n) is called a summability kernel (respectively, summability operator).

In [23], we have proved that by choosing a suitably smooth, nondecreasing function $h : [0,\infty) \to [0,\infty)$ with $h(t) = 1$ for $0 \le t \le 1/2$ and $h(t) = 0$ for $t \ge 1$, the local maxima of the expressions $|\sigma_{2^n}(h,f,x) - \sigma_{2^{n-1}}(h,f,x)|$ give a very accurate and stable approximation of the location of singularities. This includes not just jump discontinuities in f itself, but also the jump discontinuities in derivatives of f of arbitrarily high order, the maximum order being determined by the smoothness of h. Moreover, we showed in [24] that the polynomials $\sigma_{2^n}(h,f)$ provide a spectral degree of approximation to f at different points of the interval $[-\pi,\pi]$ commensurate with the smoothness of f in the vicinity of these points. Thus, the operators σ_{2^n} are *universal*; i.e., can be constructed only with spectral data, with no a priori knowledge of the location of singularities, and *auto-adaptive*; i.e., their approximation power is asymptotically the best possible given the smoothness of f on different intervals. In particular, they solve both the problems mentioned above in one stroke. Extensions of these ideas in the context of Jacobi polynomials and approximation on the sphere are given in [18, 19], and a survey can be found in [25]. Our methods work both with spectral and pseudo–spectral data.

While these methods are adequate for the detection of jump discontinuities of derivatives of f and an approximation of a smooth f, they are not adequate for a spectral approximation of piecewise analytic functions. In [33], Tanner has described algorithms to obtain an approximation to a piecewise analytic function f yielding approximants which converge exponentially fast on the intervals where f is analytic. He uses filters depending on the point at which the approximation is taking place, and requires, at least in theory, an a priori knowledge of the location of the singularities. It is not clear whether the approximants provide a near best approximation globally.

The problem of approximating piecewise analytic functions has been studied before in approximation theory community by several mathematicians, including Gaier, Ivanov, Saff, and Totik ([8, 28, 11], and references therein). The results in this context are typically stated for approximation of aperiodic functions by algebraic polynomials, but approximation of periodic functions by trigonometric polynomials is often an essential step. For integer $n \ge 0$, let Π_n denote the class of all algebraic polynomials of degree at most n. We find it convenient to extend this notation for real, nonnegative values of n as well. 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 [8], Gaier constructed a sequence of linear operators $\mathcal{G}_n$ such that for each continuous $f : [-1,1] \to \mathbb{R}$, and integer $n \ge 1$, $\mathcal{G}_n(f) \in \Pi_n$, and satisfies the following conditions:

$$\max_{x\in[-1,1]} |f(x) - \mathcal{G}_n(f,x)| \le M(f)e^{-\alpha n} + c_1 E_{n/6,\infty}(f), \tag{1.1}$$

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

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

where $M(f)$ is a positive constant depending only on f, and c_1, c, α are absolute positive constants. Gaier's construction is based on the Fourier-Chebyshev coefficients of f and depends heavily on a resulting contour integral. Necessary and sufficient conditions are given in [28] for a function $\phi : [0,\infty) \to [0,\infty)$ so that there exist polynomials P with $P_n(0) = 1$, and satisfiying $|P_n(x)| \le c_1\exp(-n\phi(x))$, $x \in [-1,1]$, for some positive constant $c_1 = c_1(\phi)$.

In many applications, one may need to use expansions in orthogonal polynomials other than the Chebyshev polynomials [5, 9, 12]. From a computational point of view, it is also desirable to be able to use samples of f rather than the Fourier information.

In this paper, we will solve the twin problems of detection of singularities and localized spectral approximation of piecewise analytic functions using either the coefficients in an expansion with respect to a very general class of orthogonal polynomials, or values of the function at certain points on

$[-1,1]$. In Propositions 2.1 and 5.2, we will construct a Littlewood Paley decomposition of the form $f = \sum_{n=0}^\infty \sum_{k=1}^{c2^n} b_{k,n}(f)\Psi_{k,n}$ where the convergence is uniform on $[-1,1]$, each $\Psi_{k,n} \in \Pi_{2^{n+3}-1}$, and the coefficients $b_{k,n}(f)$ are obtained as a linear combination of either the coefficients of the orthgonal polynomial expansion of f, or values of f at certain points. The coefficients are computed in the form $b_{k,n}(f) = \tau_n(f, y_{k,n})$ for a linear operator τ_n and suitably chosen points $y_{k,n}$. They satisfy the Riesz condition:

$$\sum_{k,n} |b_{k,n}|^2 \sim \int_{-1}^1 |f(t)|^2 d\mu(t),$$

where μ is the measure used to define the orthogonal polynomial system. We will demonstrate in Theorems 2.2 and 5.2 the localization of the coefficients by showing that f is analytic at a point $x \in [-1,1]$ if and only if there is a nondegenerate interval $I \subseteq [-1,1]$ containing x such that

$$\limsup_{n\to\infty} \left\{ \max_{y_{k,n}\in I} |b_{k,n}(f)| \right\}^{1/2^n} < 1.$$

The partial sums $\sum_{n=0}^N \sum_k b_{k,n}(f)\Psi_{k,n}$ are our analogues of the operators $\mathcal{G}_{2^{N+3}}$. We will show in Theorem 2.1, 5.1 that they satisfy an inequality analogous to (1.1), but without the extra term $M(f)e^{-\alpha n}$. The construction of these operators do not require an a priori knowledge of the location of the singularities of the target function, and clearly, these operators are based on a general class of orthogonal polynomials. As shown in [18], the behavior of the coefficients $b_{k,n}(f)$ also characterises the membership of f in different local Besov spaces. We will not elaborate on this aspect in this paper.

In the Fourier space, our construction is described by a *mixing filter*; i.e., a matrix A such that the Fourier coefficients in our approximation to f are given by $A\hat{\mathbf{f}}$, where $\hat{\mathbf{f}}$ is the vector of Fourier coefficients of the target function f. Unlike the classical filter which modifies each frequency separately, a mixing filter takes a linear mixture of all the available data at each frequency. In the physical space, the filter gives mollifiers of the form $\Phi_n^*(x,y)$, which are algebraic polynomials of degree at most $8n$ in x and y, with the property that $|\Phi_n^*(x,y)| \le c_1 \exp(-c_2(x,y)n)$ if $x \ne y$. There are a number of kernels defined in the literature (see [25] for a survey) where the kernels satisfy a bound of the form $c(Q,x,y)/n^Q$ for every integer Q. This rate is not sufficient for our purpose. Moreover, the constructions of such kernels depend heavily on the special function properties of the orthogonal polynomial system in question. In contrast, we require only the existence of a bounded reproducing summability kernel (see Section 5 for precise definition). Freud [7] has proved the existence of such kernels for a very general class of orthogonal polynomials for which no asymptotic expansions are known.

Recently, we have developed in [15] the analogue of our theory of polynomial frames in the context of eigenfunctions of the Laplace–Beltrami operator on an arbitrary smooth manifold. In the case of some important manifolds, these eigenfunctions are closely related to Jacobi polynomials; for example, ultraspherical polynomials in the case of the Euclidean sphere. In Section 2, we introduce our ideas in the context of Jacobi polynomials. The theory is illustrated with a few numerical examples in Section 3. In Section 4, we apply this theory to the case of local approximation of functions on the Euclidean unit sphere. In Section 5, the theory in Section 2 is generalized further to the case of arbitrary systems of orthogonal polynomials, subject to certain technical conditions. The proofs of the new results are given in Section 6.

2 Jacobi polynomials

In this paper, let $\alpha, \beta \ge -1/2$. The Jacobi measure is defined by

$$d\mu^{(\alpha,\beta)}(x) = \begin{cases} (1-x)^\alpha (1+x)^\beta dx, & \text{if } x \in (-1,1), \\ 0, & \text{otherwise.} \end{cases}$$

There exists a unique system of polynomials $\{p_k^{(\alpha,\beta)}\}_{k=0}^\infty$, called orthonormalized Jacobi polynomials, with each $p_k \in \Pi_k$, and having a positive leading coefficient such that

$$\int_{-1}^1 p_k^{(\alpha,\beta)} p_j^{(\alpha,\beta)} d\mu^{(\alpha,\beta)} = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

If $1 \le p \le \infty$, and $A \subseteq [-1,1]$, the space $L^p(\mu^{(\alpha,\beta)};A)$ consists of measurable functions f for which

$$\|f\|_{\mu^{(\alpha,\beta)};p,A} := \begin{cases} \left(\int_A |f|^p d\mu^{(\alpha,\beta)}\right)^{1/p}, & \text{if } 1 \le p < \infty, \\ \operatorname*{ess\,sup}_{t\in A} |f(t)|, & \text{if } p = \infty \end{cases}$$

is finite, with the usual convention that two functions are considered equal if they are equal almost everywhere. The space $X^p(\mu^{(\alpha,\beta)};A)$ denotes $L^p(\mu^{(\alpha,\beta)};A)$ if $1 \le p < \infty$ and the space of bounded uniformly continuous functions on A (equipped with the supremum norm) if $p = \infty$. The mention of the set A will be omitted if $A = [-1,1]$. In particular, $X^\infty(\mu^{(\alpha,\beta)}) = C[-1,1]$, and we will write $\|f\|_\infty$ rather than the more cumbersome notation $\|f\|_{\mu^{(\alpha,\beta)};\infty}$.

We recall that for real $x \ge 0$, Π_x denotes the class of all algebraic polynomials of degree at most x. For $1 \le p \le \infty$ and $f \in L^p(\mu^{(\alpha,\beta)})$, we define the degree of approximation of f from Π_x by

$$E_{x,p}(\alpha,\beta;f) := \min_{P\in\Pi_x} \|f - P\|_{\mu^{(\alpha,\beta)};p}.$$

Of course, $E_{n,\infty}(\alpha,\beta;f) = E_{n,\infty}(f)$ as defined in the introduction. We adopt the following convention regarding constants: the symbols $c, c_1, \cdots$ will denote generic positive constants, dependent only on such fixed parameters in the discussion as p, α, β, etc. Their value may be different at different occurrences, even within the same formula.

It is readily seen that the partial sum of order 2^n of the Jacobi polynomial expansion of a function $f \in L^1(\mu^{(\alpha,\beta)})$ is given by $\int_{-1}^1 f(y)K_n^{(\alpha,\beta)}(\circ,y)d\mu^{(\alpha,\beta)}(y)$, where the Christoffel–Darboux kernel K_n is defined by

$$K_n^{(\alpha,\beta)}(x,y) := \sum_{k=0}^{2^n} p_k^{(\alpha,\beta)}(x)p_k^{(\alpha,\beta)}(y).$$

However, the sequence of these partial sums need not converge to f for every $f \in L^1(\mu^{(\alpha,\beta)})$. To get convergent sums, we need to use a summability method. It is known [30, Theorem 9.1.4] that for every continuous f, the Cesáro means of order $K > \alpha+\beta+1$ of the Jacobi polynomial expansion of f converge uniformly to f. In order to get a near best approximation and exponential localization, we need to introduce an operator based on another related kernel.

Let $K \ge \alpha+\beta+2$ be an integer, $h : [0,\infty) \to \mathbb{R}$ be a function which is a K times iterated integral of a function of bounded variation, $h(x) = 1$ for $0 \le x \le 1/2$, and $h(x) = 0$ for $x > 1$. Then for $x, y \in \mathbb{C}$, $n = 0, 1, \cdots$, we define the kernel

$$\Phi_n(\mu^{(\alpha,\beta)};h,x,y) := \sum_{k=0}^{2n} h\left(\frac{k}{2n}\right) p_k^{(\alpha,\beta)}(x)p_k^{(\alpha,\beta)}(y). \tag{2.1}$$

Using a summation by parts argument or directly as in [18], one can prove that

$$\sup_{n\ge0,x\in[-1,1]} \int_{-1}^1 |\Phi_n(\mu^{(\alpha,\beta)};h,x,y)|d\mu^{(\alpha,\beta)}(y) < \infty.$$

In addition, it is easy to verify that $\int_{-1}^1 \Phi_n(\mu^{(\alpha,\beta)};h,x,y)P(y)d\mu^{(\alpha,\beta)}(y) = P(x)$ for every $P \in \Pi_n$. Therefore, the polynomials $\int_{-1}^1 f(y)\Phi_n(\mu^{(\alpha,\beta)};h,\circ,y)d\mu^{(\alpha,\beta)}(y)$ converge uniformly to f for every continuous f, at a rate comparable to $E_{n,\infty}(\alpha,\beta;f)$. In [18], we have shown that the smoother the h, the better localized the kernels Φ_n are; in particular, if h is infinitely often differentiable, then for every integer Q, $|\Phi_n(\mu^{(\alpha,\beta)};h,x,y)| \le c(Q,x,y)/n^Q$. However, this rate is not enough to detect the possibility of analytic continuation of a function near a point. In order to obtain an exponential rate of decay, we use the following kernel instead:

$$\Phi_n^*(\mu^{(\alpha,\beta)};h,x,y) := \left(\frac{4-(x-y)^2}{4}\right)^n \Phi_{3n}(\mu^{(\alpha,\beta)};h,x,y). \tag{2.2}$$

We will see in Theorem 2.1 that choosing Φ_{3n} above leads to a polynomial reproduction property. The summability operators σ_n^C are defined for $f \in L^1(\mu^{(\alpha,\beta)})$, $x \in \mathbb{C}$, $n = 0, 1, \cdots$, by

$$\sigma_n^C(\alpha,\beta;h,f,x) := \int_{-1}^{1} f(y)\Phi_n^*(\mu^{(\alpha,\beta)};h,x,y)d\mu^{(\alpha,\beta)}(y). \tag{2.3}$$

We note that $\sigma_n^C(\alpha,\beta;h,f) \in \Pi_{8n}$. Since $\Phi_n^*(\mu^{(\alpha,\beta)};h,x,y)$ is a symmetric polynomial of degree $8n$ in each of its variables, one has the representation

$$\Phi_n^*(\mu^{(\alpha,\beta)};h,x,y) = \sum_{k=0}^{8n}\sum_{j=0}^{8n} a_{n;k,j}^{(\alpha,\beta)}(h) p_k^{(\alpha,\beta)}(x) p_j^{(\alpha,\beta)}(y),$$

where, for each integer $n \geq 0$, $(a_{n;k,j}^{(\alpha,\beta)}(h))$ is a symmetric matrix. Defining the Jacobi coefficients of $f \in L^1(\mu^{(\alpha,\beta)})$ by

$$\hat{f}(\alpha,\beta;j) = \int_{-1}^{1} f p_j^{(\alpha,\beta)} d\mu^{(\alpha,\beta)}, \qquad j = 0, 1, \cdots,$$

it follows that

$$\sigma_n^C(\alpha,\beta;h,f) = \sum_{k=0}^{8n}\left(\sum_{j=0}^{8n} a_{n;k,j}^{(\alpha,\beta)}(h)\hat{f}(\alpha,\beta;j)\right) p_k^{(\alpha,\beta)}.$$

Thus, the operators σ_n^C can be computed using finitely many Jacobi coefficients of f.

From a computational point of view, we would like to define discrete versions of these operators, which are obtained using Gauss quadrature formulas. For $m \geq 1$, let $x_{k,m}$, $k = 1, \cdots, m$, be the zeros of $p_m^{(\alpha,\beta)}$, and

$$\lambda_{k,m} := \left(\sum_{j=1}^{m-1}\left(p_j^{(\alpha,\beta)}(x_{k,m})\right)^2\right)^{-1}$$

be the corresponding Cotes numbers. We define the discretized versions of the operators by

$$\sigma_n^D(\alpha,\beta;h,f,x) := \sum_{k=1}^{8n+1} \lambda_{k,8n+1} f(x_{k,8n+1}) \Phi_n^*(\mu^{(\alpha,\beta)};h,x,x_{k,8n+1}). \tag{2.4}$$

The following theorem is our generalization of the result of Gaier in the context of Jacobi polynomials.

Theorem 2.1 *Let $1 \leq p \leq \infty$, $\alpha, \beta \geq -1/2$, $f \in L^p(\mu^{(\alpha,\beta)})$. For integer $n \geq 0$, let $\sigma_n(f)$ denote either $\sigma_n^C(\alpha,\beta;h,f)$ or $\sigma_n^D(\alpha,\beta;h,f)$.*
(a) *We have $\sigma_n(P) = P$ for $P \in \Pi_n$, $\|\sigma_n(f)\|_{\mu^{(\alpha,\beta)};p} \leq c\|f\|_{\mu^{(\alpha,\beta)};p}$, and*

$$E_{8n,p}(\alpha,\beta;f) \leq \|f - \sigma_n(f)\|_{\mu^{(\alpha,\beta)};p} \leq c_1 E_{n,p}(\alpha,\beta;f). \tag{2.5}$$

(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 $\{z \in \mathbb{C} : |z - x_0| \leq d\}$ for some d with $0 < d \leq 2$. Then*

$$|f(x) - \sigma_n(f,x)| \leq c(f,x_0)\exp\left(-n\frac{d^2\log(e/2)}{e^2\log(e^2/d)}\right), \qquad x \in [x_0 - d/e, x_0 + d/e] \cap [-1,1]. \tag{2.6}$$

We note a few interesting features of this theorem. First, we are able to drop the extra term $M(f)e^{-\alpha n}$ in (1.1) at the expense of a higher estimate $c_1 E_{n/8,\infty}(f)$ in place of $c_1 E_{n/6,\infty}(f)$. Second, our construction can be based either on the coefficients in general Jacobi polynomial expansions, or based on values of the function. Finally, we think that our proofs are simpler than those given by Gaier in [8].

Next, we describe a Littlewood–Paley expansion of functions in $X^p(\mu^{(\alpha,\beta)})$, where the analyticity of the target function at a point can be completely characterised using certain coefficients of this expansion.

Towards this end, we define the continuous and discrete frame operators by

$$\begin{aligned}
\tau_n^C(\alpha,\beta;h,f,x) &:= \begin{cases} \sigma_1^C(\alpha,\beta;h,f,x), & \text{if } n=0,\\ \sigma_{2^n}^C(\alpha,\beta;h,f,x)-\sigma_{2^{n-1}}^C(\alpha,\beta;h,f,x), & \text{if } n=1,2,\cdots,\end{cases}\\
\tau_n^D(\alpha,\beta;h,f,x) &:= \begin{cases} \sigma_1^D(\alpha,\beta;h,f,x), & \text{if } n=0,\\ \sigma_{2^n}^D(\alpha,\beta;h,f,x)-\sigma_{2^{n-1}}^D(\alpha,\beta;h,f,x), & \text{if } n=1,2,\cdots.\end{cases}
\end{aligned} \tag{2.7}$$

We note that $\tau_n^C(\alpha,\beta;h,f), \tau_n^D(\alpha,\beta;h,f) \in \Pi_{2^{n+3}}$. The next proposition demonstrates the use of these operators in obtaining a Littlewood–Paley decomposition of functions in $X^p(\mu^{(\alpha,\beta)})$, $1\le p\le\infty$.

Proposition 2.1 *Let $1\le p\le\infty$, $\alpha,\beta\ge -1/2$, $f\in X^p(\mu^{(\alpha,\beta)})$. If $N_n \ge 2^{n+3}+1$ are integers, one has the Littlewood–Paley decomposition*

$$\begin{aligned}
f &= \sum_{n=0}^{\infty} \tau_n^C(\alpha,\beta;h,f)\\
&= \sum_{n=0}^{\infty}\sum_{k=1}^{N_n} \lambda_{k,N_n}\tau_n^C(\alpha,\beta;h,f,x_{k,N_n})\left\{K_{n+3}^{(\alpha,\beta)}(\circ,x_{k,N_n}) - K_{n-1}^{(\alpha,\beta)}(\circ,x_{k,N_n})\right\},
\end{aligned} \tag{2.8}$$

with convergence in the sense of $X^p(\mu^{(\alpha,\beta)})$. Moreover, we have for every $f\in L^2(\mu^{(\alpha,\beta)})$,

$$\begin{aligned}
c_1\int_{-1}^{1}|f(t)|^2(1-t)^\alpha(1+t)^\beta dt &\le \sum_{n=0}^{\infty}\sum_{k=1}^{N_n}\lambda_{k,N_n}|\tau_n^C(\alpha,\beta;h,f,x_{k,N_n})|^2\\
&\le c_2\int_{-1}^{1}|f(t)|^2(1-t)^\alpha(1+t)^\beta dt.
\end{aligned}$$

If $f\in C[-1,1]$, then $f=\sum_{n=0}^\infty \tau_n^D(\alpha,\beta;h,f)$, with convergence being uniform.

The following theorem demonstrates that, unlike the Fourier–Jacobi coefficients $\hat f(\alpha,\beta;j)$, the behavior of the coefficients $\tau_n^C(\alpha,\beta;h,f,x_{k,N_n})$ in the Littlewood–Paley expansion (2.8) for points x_{k,N_n} in a neighborhood of a point x_0 reflects the (analytic) regularity of the function f at x_0.

Theorem 2.2 *Let $\alpha,\beta\ge -1/2$, $x_0\in[-1,1]$ and $f\in C[-1,1]$. For integer $n\ge 0$, let $\tau_n(f)$ denote either $\tau_n^C(\alpha,\beta;h,f)$ or $\tau_n^D(\alpha,\beta;h,f)$.*
(a) *The function f has an extension as an analytic function in a complex neighborhood of x_0 if and only if there exists a nondegenerate interval $I\subseteq[-1,1]$ containing x_0 such that*

$$\limsup_{n\to\infty}\max_{x\in I}|\tau_n(f,x)|^{1/2^n} < 1. \tag{2.9}$$

(b) *The function f has an extension as an analytic function in a complex neighborhood of x_0 if and only if there exists a nondegenerate interval $I\subseteq[-1,1]$ containing x_0 such that*

$$\limsup_{n\to\infty}\left\{\max_{x_{k,2^{n+6}+1}\in I}|\tau_n(f,x_{k,2^{n+6}+1})|\right\}^{1/2^n} < 1. \tag{2.10}$$

We note that choosing $N_n = 2^{n+6}+1$ in Proposition 2.1, and using τ_n^C in place of τ_n, (2.10) gives a characterization of regular points of f in terms of the coefficients in the Littlewood–Paley expansion (2.8).

3 Numerical examples

In this section, we illustrate the construction of our localized kernels and their approximation properties using some numerical examples. We wish to point out several aspects of our theory: (1) the property of near best approximation on the whole interval, (2) the effect of smoothness of the function h on the ability of the operators in local approximation, (3) the utility of using the exponentially localized kernels rather than C^∞ kernels.

In our examples below, we use the Chebyshev polynomials; i.e., the polynomials T_n defined by $T_n(\cos\theta) = \cos n\theta$, $\theta \in [0,\pi]$, $n = 0, 1, \cdots$. The polynomials

$$p_n^T = \begin{cases} 1, & \text{if } n = 0, \\ \sqrt{2}T_n, & \text{if } n = 1, 2, \cdots, \end{cases}$$

are orthonormalized with respect to the measure

$$d\mu^T(x) = \frac{d\mu^{(-1/2,-1/2)}(x)}{\pi} = \frac{dx}{\pi(1-x^2)^{1/2}}, \qquad x \in (-1,1).$$

For the sake of brevity, we write $\Phi_n^T(h,x,y)$ for $\Phi_n(\mu^{(-1/2,-1/2)}; h, x, y)$ and $\Phi_n^{*T}(h,x,y)$ in place of $\Phi_n^*(\mu^{(-1/2,-1/2)}; h, x, y)$. The corresponding operators are

$$\begin{aligned} V_n^T(h,f,x) &:= \frac{1}{8n+1}\sum_{k=1}^{8n+1} f\left(\cos\frac{(2k-1)\pi}{16n+2}\right)\Phi_{4n}^T\left(h, x, \cos\frac{(2k-1)\pi}{16n+2}\right) \\ \sigma_n^T(h,f,x) &:= \frac{1}{8n+1}\sum_{k=1}^{8n+1} f\left(\cos\frac{(2k-1)\pi}{16n+2}\right)\Phi_n^{*T}\left(h, x, \cos\frac{(2k-1)\pi}{16n+2}\right). \end{aligned}$$

Clearly, both V_n^T and σ_n^T yield polynomials in Π_{8n}.

For the function h, we will consider two functions. Both h_1 and h_∞ below are equal to 1 on $[0,1/2]$, and equal to 0 on $[1,\infty)$. For $t \in (1/2,1)$ they are defined by

$$\begin{aligned} h_1(t) &:= 2 - t, \\ h_\infty(t) &:= \exp\left(-\frac{\exp(2/(1-2t))}{1-t}\right). \end{aligned}$$

The function h_1 is clearly piecewise linear, but not differentiable, and the function h_∞ is a C^∞ function. The operator $V_n^T(h_1)$ is a discretization of the classical de la Vallée Poussin operator. It is known that this operator provides a near best approximation globally in the sense that for any continuous function $f : [-1,1] \to \mathbb{R}$,

$$E_{n,\infty}(f) \le \|f - V_n^T(h_1,f)\|_\infty \le cE_{n/2,\infty}(f).$$

We demonstrate this global approximation property of the kernels $V_n^T(h_1)$ and $\sigma_n^T(h_1)$ in the case of two functions, the first of which is

$$f_a(x) := |x - 1/4|, \qquad x \in [-1,1].$$

To define the second function, we recall that the cardinal B-spline of order 4 is the function defined by (cf. [4, Formula (4.1.12), p. 84])

$$M_4(x) = \frac{1}{6}\left\{x_+^3 - 4(x-1)_+^3 + 6(x-2)_+^3 - 4(x-3)_+^3 + (x-4)_+^3\right\}$$

where $a_+ = \max(a,0)$. We define $f_b(x) = M_4(2x+2)$. Thus, f_b is analytic on $(-1,1)$, except at $\pm 1/2, 0$, where it is twice continuously differentiable. It is well known that $E_{n,\infty}(f_a) \sim n^{-1}$, while $E_{n,\infty}(f_b) \sim n^{-3}$.

In this section only, let $\mathcal{C}$ denote the set of 10,000 equidistant points on $[-1,1]$. We estimate the supremum norm of continuous functions f by $\max_{t\in\mathcal{C}} |f(t)|$. In particular, we write

$$\epsilon_n(f,V) := \max_{x\in\mathcal{C}} |f(x) - V_n^T(h_1,f,x)|, \quad \epsilon_n(f,\sigma) := \max_{x\in\mathcal{C}} |f(x) - \sigma_n^T(h_1,f,x)|,$$

and

$$\delta_n(f,V) := \log_2 \frac{\epsilon_n(f,V)}{\epsilon_{2n}(f,V)}, \quad \delta_n(f,\sigma) := \log_2 \frac{\epsilon_n(f,V)}{\epsilon_{2n}(f,V)}.$$

n	$\epsilon_n(f_a,\sigma)$	$\epsilon_n(f_a,V)$	$\epsilon_n(f_b,\sigma)$	$\epsilon_n(f_b,V)$
8	$1.8065*10^{-2}$	$1.3609*10^{-2}$	$1.3838*10^{-4}$	$4.7691*10^{-5}$
16	$9.8889*10^{-3}$	$8.0887*10^{-3}$	$1.684*10^{-5}$	$6.0351*10^{-6}$
32	$4.8372*10^{-3}$	$3.8924*10^{-3}$	$2.0823*10^{-6}$	$7.5158*10^{-7}$
64	$2.3075*10^{-3}$	$1.7856*10^{-3}$	$2.5918*10^{-7}$	$9.3814*10^{-8}$

Table 1: Maximum absolute errors.

n	$\delta_n(f_a,\sigma)$	$\delta_n(f_a,V)$	$\delta_n(f_b,\sigma)$	$\delta_n(f_b,V)$
8	0.8693	0.7506	3.0387	2.9823
16	1.0316	1.0552	3.0156	3.0054
32	1.0678	1.1243	3.0062	3.0021

Table 2: The smoothness index as predicted by δ_n's.

Table 1 shows the decay of errors $\epsilon_n(f,V)$ and $\epsilon_n(f,\sigma)$ for different values of n.

In light of the direct theorems of approximation theory, the quantities $\delta_n(f_a,V)$ and $\delta_n(f_a,\sigma)$ should be close to 1, and the corresponding quantities for f_b should be close to 3. Table 2 confirms this fact (cf. Theorem 2.2(a).)

It is clear from Table 1 that the maximum error is less with the de la Vallée Poussin operators than the exponentially localized operators, although their rate of decrease in both cases is ommensurate with the theoretical degree of approximation.

Next, we illustrate the better localization of the kernels σ_n^T than that given by V_n^T. In all the subfigures of Figure 1 below, the value $(-k)$ on the y axis corresponds to the value 10^{-k} for the errors plotted in the figures. We note that only $8n+1$ values of the function are used in the computation of the transforms V_n^T and σ_n^T. In particular, in the top left figure in Figure 1, an absolute error of less than 10^{-20} is obtained away from the singularity, using only 513 samples of the function f_a.

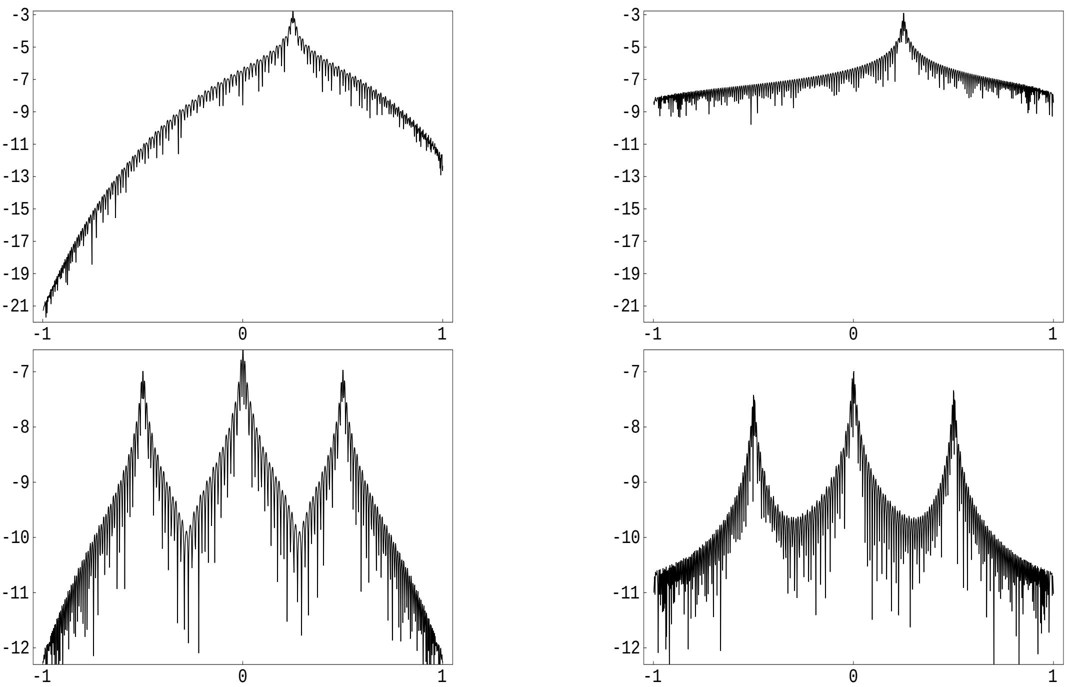

Figure 1: Clockwise, the graphs represent $\log_{10}|f_a-\sigma_{64}^T(h_1,f_a)|$, $\log_{10}|f_a-V_{64}^T(h_1,f_a)|$, $\log_{10}|f_b-V_{64}^T(h_1,f_b)|$, and $\log_{10}|f_b-\sigma_{64}^T(h_1,f_b)|$.

Next, we wish to illustrate the advantage of the operator $\sigma_n^T(h_1)$ over the operator $V_n^T(h_\infty)$ in the detection of an analytic singularity. We consider the function

$$f_c(x) = \chi(x) + \sum_{k=0}^{\infty} \frac{\exp(-\sqrt{k})}{k+1} \cos(k\pi/3) T_k(x),$$

where, in this section only, $\chi(x) = 1$ if $x \in [-1, -1/\sqrt{2}]$ and $\chi(x) = 0$ otherwise. Using Pringsheim's theorem [16, Theorem 17.13], it is possible to show that f_c is infinitely often differentiable in $(-1, 1)$, except for the jump discontinuity at $-1/\sqrt{2}$, but cannot be extended as an analytic function in any neighborhood of 1/2. Figure 2 shows that the operator $\sigma_n^T(h_1)$ can detect this singularity also in the presence of the much stronger jump discontinuity, whereas the operator $V_n^T(h_\infty)$ does not perform so well. Moreover, the approximation at 1/2 is better with $\sigma_n^T(h_1)$ than $V_n^T(h_\infty)$. We note that we have used only 129 evaluations of f_c in computing the various transforms used in Figure 2.

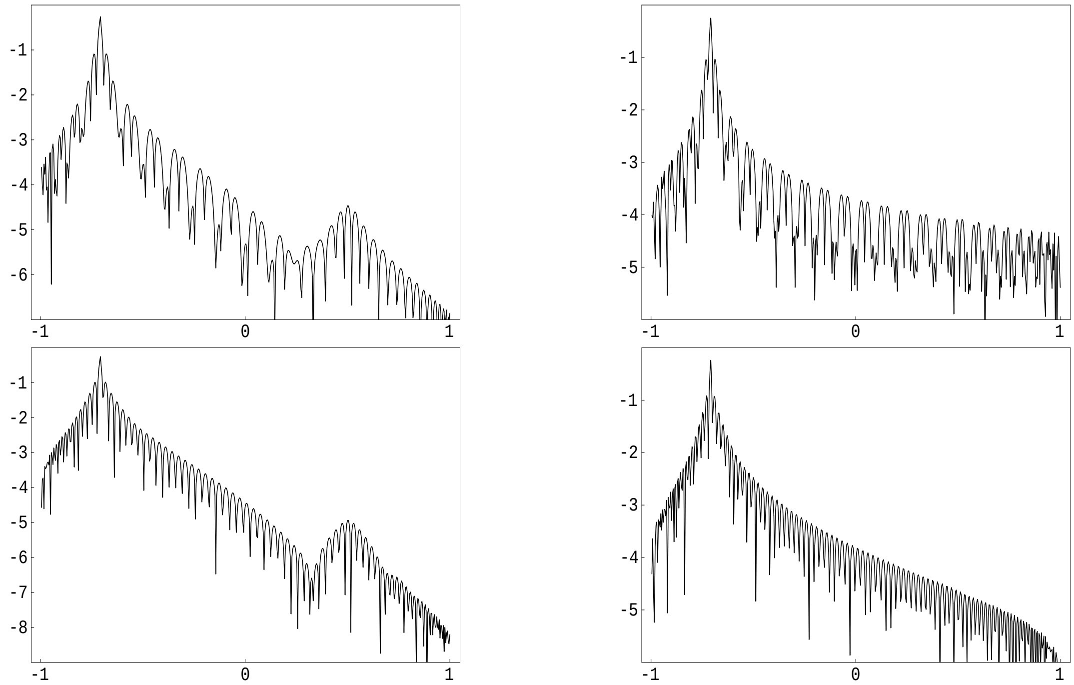

Figure 2: Clockwise, the graphs represent $\log_{10}|f_c - \sigma_{16}^T(h_1, f_c)|$, $\log_{10}|f_c - V_{16}^T(h_1, f_c)|$,$\log_{10}|f_c - V_{16}^T(h_\infty, f_c)|$, $\log_{10}|f_c - \sigma_{16}^T(h_\infty, f_c)|$.

4 Approximation on the sphere

The constructions described for the unit interval can be adapted easily for the unit sphere of a Euclidean space. Several numerical experiments illustrating the superior local approximation properties of our operators are given in [13]. Although the operators there are not exponentially localized, we do not feel that additional experiments in this direction will add anything fundamentally new to this subject. Therefore, we restrict our discussion to a sketch of the adaptations of the theory, and indicate the differences.

Let $q \geq 1$ be an integer,

$$\mathbb{S}^q := \{(x_1, \cdots, x_{q+1}) \in \mathbb{R}^{q+1} \ : \ \sum_{j=1}^{q+1} x_j^2 = 1\}.$$

A spherical cap, centered at $\mathbf{x}_0 \in \mathbb{S}^q$ and radius α is defined by

$$\mathbb{S}^q_\alpha(\mathbf{x}_0) := \{\mathbf{x} \in \mathbb{S}^q \ : \ \mathbf{x}\cdot\mathbf{x}_0 \ge \cos\alpha\}.$$

We note that for any $\mathbf{x}_0 \in \mathbb{S}^q$, $\mathbb{S}^q_\pi(\mathbf{x}_0) = \mathbb{S}^q$. The surface area (volume element) measure on $\mathbb{S}^q$ will be denoted by μ_q^*, and we write $\mu_q^*(\mathbb{S}^q) =: \omega_q$. The spaces $X^p(\mathbb{S}^q)$ and $C(\mathbb{S}^q)$ on the sphere are defined analogously to the case of the interval.

A spherical polynomial of degree m is the restriction to $\mathbb{S}^q$ of a polynomial in $q+1$ real variables with total degree m. For $x \ge 0$, the class of all spherical polynomials of degree at most x will be denoted by Π^q_x. For integer $\ell \ge 0$, the class of all homogeneous, harmonic, spherical polynomials of degree ℓ will be denoted by $\mathbb{H}^q_\ell$, and its dimension by d^q_ℓ. For each integer $\ell \ge 0$, let $\{Y_{\ell,k} \ : \ k = 1,\cdots,d^q_\ell\}$ be a μ_q^*-orthonormalized basis for $\mathbb{H}^q_\ell$. It is known (cf. [29, 26]) that for any integer $n \ge 0$, $\{Y_{\ell,k} \ : \ \ell = 0,\cdots,n, \ k = 1,\cdots,d^q_\ell\}$ is an orthonormal basis for Π^q_n. The connection with the theory of orthogonal polynomials on $[-1,1]$ is the following addition formula (cf. [26], where the notation is different):

$$\sum_{k=1}^{d^q_\ell} Y_{\ell,k}(\mathbf{x})Y_{\ell,k}(\mathbf{y}) = \omega_{q-1}^{-1} p_\ell^{(q/2-1,q/2-1)}(1)p_\ell^{(q/2-1,q/2-1)}(\mathbf{x}\cdot\mathbf{y}), \qquad \ell = 0,1,\cdots.$$

It is proved in [21], (see also [6]) that if $n \ge 1$ is an integer, $\mathcal{C}_n$ is a finite set of points on $\mathbb{S}^q$ such that

$$\max_{\mathbf{x}\in\mathbb{S}^q}\min_{\mathbf{y}\in\mathcal{C}_n} \ \mathrm{dist}\ (\mathbf{x},\mathbf{y}) \le c/n,$$

for a judiciously chosen constant $c > 0$, then there exist nonnegative weights $w_\mathbf{x}$, $\mathbf{x} \in \mathcal{C}_n$ such that

$$\sum_{\mathbf{x}\in\mathcal{C}_n} w_\mathbf{x} P_1(\mathbf{x})P_2(\mathbf{x}) = \int_{\mathbb{S}^q} P_1P_2 d\mu_q^*, \qquad P_1, P_2 \in \Pi^q_{8n}. \tag{4.1}$$

We will say that $\mathcal{C}_n$ admits a positive quadrature formula of order $16n$.

It is possible to define analytic functions on $\mathbb{S}^q$ using a parametrization. A more immediate, parameter-free definition that comes to mind is to think of functions on $\mathbb{S}^q$ as functions of $q+1$ real variables, and expect an analytic continuation to a subset of $\mathbb{C}^{q+1}$, or at least, a real analytic continuation to a subset of $\mathbb{R}^{q+1}$. Since $\mathbb{S}^q$ is not a set of uniqueness of functions analytic on $\mathbb{C}^{q+1}$, it is not possible to define the extent of such analytic continuation using spectral data on the sphere. Nevertheless, one can define an analogue from the approximation theory point of view as follows. Let $\mathbf{x}_0 \in \mathbb{S}^q$, $\mathcal{K}$ be a spherical cap centered at $\mathbf{x}_0$ and $f \ : \ \mathcal{K} \to \mathbb{C}$. We write

$$E_n(\mathcal{K}, f) := \inf_{P\in\Pi^q_n} \|f - P\|_{\mu_q^*;\infty,\mathcal{K}}.$$

The class $\mathcal{A}_q(\mathbf{x}_0)$ is defined to be the class of all functions $f \in C(\mathbb{S}^q)$ such that for some spherical cap $\mathcal{K}$ centered at $\mathbf{x}_0$,

$$\limsup_{n\to\infty} E_n(\mathcal{K}, f)^{1/n} < 1.$$

Let $h : [0,\infty) \to \mathbb{R}$ be a q times iterated integral of a function of bounded variation, $h(x) = 1$ if $x \in [0, 1/2]$, and $h(x) = 0$ if $x \in [1,\infty)$. The role of the kernels $K_n^{(\alpha,\beta)}$ and $\Phi_n(\mu^{(\alpha,\beta)}; h, x, y)$ is played respectively by

$$K_n^S(\mathbf{x}\cdot\mathbf{y}) := \omega_{q-1}^{-1}\sum_{\ell=0}^{2^n} p_\ell^{(q/2-1,q/2-1)}(1)p_\ell^{(q/2-1,q/2-1)}(\mathbf{x}\cdot\mathbf{y}) = \omega_{q-1}^{-1}K_n^{(q/2-1,q/2-1)}(1,\mathbf{x}\cdot\mathbf{y})$$

and

$$\Phi_n^S(q; h, \mathbf{x},\mathbf{y}) := \omega_{q-1}^{-1}\sum_{\ell=0}^{2n} h(\ell/(2n))p_\ell^{(q/2-1,q/2-1)}(1)p_\ell^{(q/2-1,q/2-1)}(\mathbf{x}\cdot\mathbf{y}) = \Phi_n(\mu^{(q/2-1,q/2-1)}; h, 1, \mathbf{x}\cdot\mathbf{y}).$$

For $n=0,1,\cdots$, $\mathbf{x},\mathbf{y}\in\mathbb{S}^q$, let

$$\Phi^*_n(h,\mathbf{x},\mathbf{y})=\left(\frac{1+\mathbf{x}\cdot\mathbf{y}}{2}\right)^n\Phi^S_{3n}(q;h,\mathbf{x},\mathbf{y}).$$

Let $\{\mathcal{C}_n\}$ be a sequence of finite subsets of $\mathbb{S}^q$, with each $\mathcal{C}_n$ admitting a positive quadrature of order $16n$. We define the analogues of the operators σ^C_n, σ^D_n, τ^C_n, and τ^D_n by

$$\begin{aligned}
\sigma^{C,S}_n(h,f,\mathbf{x}) &:= \int_{\mathbb{S}^q}\Phi^*_n(h,\mathbf{x},\mathbf{y})f(\mathbf{y})d\mu^*_q(\mathbf{y}),\\
\sigma^{D,S}_n(h,f,\mathbf{x}) &:= \sum_{\mathbf{y}\in\mathcal{C}_n} w_{\mathbf{y}}\Phi^*_n(h,\mathbf{x},\mathbf{y})f(\mathbf{y}),\\
\tau^{C,S}_n(h,f,\mathbf{x}) &:= \begin{cases}\sigma^{C,S}_1(h,f,\mathbf{x}), & \text{if } n=0,\\ \sigma^{C,S}_{2^n}(h,f,\mathbf{x})-\sigma^{C,S}_{2^{n-1}}(h,f,\mathbf{x}), & \text{if } n=1,2,\cdots,\end{cases}\\
\tau^{D,S}_n(h,f,\mathbf{x}) &:= \begin{cases}\sigma^{D,S}_1(h,f,\mathbf{x}), & \text{if } n=0,\\ \sigma^{D,S}_{2^n}(h,f,\mathbf{x})-\sigma^{D,S}_{2^{n-1}}(h,f,\mathbf{x}), & \text{if } n=1,2,\cdots.\end{cases}
\end{aligned}$$

The following proposition is the analogue of Proposition 2.1, and can be proved in exactly the same way.

Proposition 4.1 *Let $1\le p\le\infty$, $\{\mathcal{C}_n\}$ be a sequence of finite subsets of $\mathbb{S}^q$, with each $\mathcal{C}_n$ admitting a positive quadrature of order at least $16n$. Then for $f\in X^p(\mathbb{S}^q)$,*

$$f=\sum_{n=0}^\infty\tau^{C,S}_n(h,f)=\sum_{n=0}^\infty\sum_{\mathbf{y}\in\mathcal{C}_n}w_{\mathbf{y}}\tau^{C,S}_n(h,f,\mathbf{y})\left\{K^S_{n+3}(\circ\cdot\mathbf{y})-K^S_{n-1}(\circ\cdot\mathbf{y})\right\}, \tag{4.2}$$

with the series converging in the sense of $X^p(\mathbb{S}^q)$. If $f\in C(\mathbb{S}^q)$, then also $f=\sum_{n=0}^\infty\tau^{D,S}_n(h,f)$ with the series converging uniformly. Further, if $f\in L^2(\mathbb{S}^q)$,

$$c_1\|f\|^2_{\mu^*_q;2,\mathbb{S}^q}\le\sum_{n=0}^\infty\|\tau^{C,S}_n(h,f)\|^2_{\mu^*_q;2,\mathbb{S}^q}=\sum_{n=0}^\infty\sum_{\mathbf{y}\in\mathcal{C}_n}w_{\mathbf{y}}|\tau^{C,S}_n(h,f,\mathbf{y})|^2\le c_2\|f\|^2_{\mu^*_q;2,\mathbb{S}^q}.$$

The analogue of Theorem 2.2 and 2.1 is the following, slightly weaker statement.

Theorem 4.1 *Let $\{\mathcal{C}_n\}$ be a sequence of finite subsets of $\mathbb{S}^q$, with each $\mathcal{C}_n$ admitting a positive quadrature of order at least $16n$, $f\in C(\mathbb{S}^q)$, and $\mathbf{x}_0\in\mathbb{S}^q$. Let σ_n denote either $\sigma^{C,S}_n$ or $\sigma^{D,S}_n$ and similarly for τ_n.*
(a) *For integer $n\ge1$, we have*

$$\|f-\sigma_n(h,f)\|_{\mu^*_q;\infty,\mathbb{S}^q}\le cE_n(\mathbb{S}^q,f). \tag{4.3}$$

If $f\in\mathcal{A}_q(\mathbf{x}_0)$, then there exists a nondegenerate spherical cap $\mathcal{K}\subseteq\mathbb{S}^q$ with center at $\mathbf{x}_0$ and $\rho\in(0,1)$ (depending upon $\mathbf{x}_0$ and f) such that

$$\|f-\sigma_n(h,f)\|_{\mu^*_q;\infty,\mathcal{K}}\le c_1(f,\mathbf{x}_0)\rho^n. \tag{4.4}$$

(b) *The function $f\in\mathcal{A}_q(\mathbf{x}_0)$ if and only if there exists a non-degenerate spherical cap $\mathcal{K}\subseteq\mathbb{S}^q$ with center at $\mathbf{x}_0$ such that*

$$\limsup_{n\to\infty}\|\tau_n(h,f)\|^{1/2^n}_{\mu^*_q;\infty,\mathcal{K}}<1. \tag{4.5}$$

5 The general case

In this section, we state our main results in a very general form. Thus, instead of the Jacobi measure, we will consider an arbitrary measure, supported on an arbitrary compact subset of $[-1,1]$. Instead of achieving the discretization of the summability and frame operators using Gauss quadrature formula, we will formulate our "discretization" using general functionals. In order to state our results, we need certain notions from measure theory and potential theory. For the convenience of the reader, we review the measure theoretic notions in Section 5.1; the ideas from potential theory are reviewed in Section 5.2. The generalizations of the operators and new results in Section 2 are given in Section 5.3.

5.1 Measures

We observe that if $N \geq 1$ is an integer, $\{x_k\}_{k=1}^N$, $\{w_k\}_{k=1}^N$ are real numbers, a sum of the form $\sum_{k=1}^N w_k f(x_k)$ can be expressed as a Stieltjes integral $\int f d\nu$, where ν is the measure that associates the mass w_k with each point x_k. The total variation measure in this case is given by $|\nu|(B) = \sum_{x_k \in B} |w_k|$, $B \subset \mathbb{R}$. We prefer to use the integral notation rather than the more explicit sum notation for a number of reasons. First, the precise locations of the points x_k, the values of w_k, and sometimes, even the value of N do not play a significant role in our theory. The use of the integral notation avoids the need to prescribe these quantities explicitly, and develop additional notation for these. Second, we wish our theory to be applicable to all L^p spaces. If $p < \infty$, point evaluations are not well defined for every f in the space, and we have to use some other local measurements, for example, averages over small subintervals around certain points. Again, the details of exactly what these points and the corresponding subintervals are, and even the nature of the local measurements do not play any significant role in our theory. The integral notation allows us to treat both the case of continuous functions and elements of L^p in a unified manner.

Let ν be a (possibly signed) measure on $\mathbb{R}$ that is either positive and finite, or has a bounded variation on $\mathbb{R}$, $|\nu|$ denote ν if ν is a positive measure, and its total variation measure if it is a signed measure. We recall that the support of ν, denoted by supp (ν) is the set of all $x \in \mathbb{R}$ such that $|\nu|(I) > 0$ for every interval I containing x. If $A \subseteq \mathbb{R}$ is $|\nu|$-measurable, $|\nu|(A) > 0$, and $f : A \to \mathbb{R}$ is $|\nu|$-measurable, we write

$$\|f\|_{\nu;p,A} := \begin{cases} \left\{ \int_A |f|^p d|\nu| \right\}^{1/p}, & \text{if } 1 \leq p < \infty, \\ |\nu| - \text{ess sup}_{t \in A} |f(t)|, & \text{if } p = \infty. \end{cases}$$

The class of measurable functions f for which $\|f\|_{\nu;p,A} < \infty$ is denoted by $L^p(\nu; A)$, with the standard convention that two functions are considered equal if they are equal $|\nu|$-almost everywhere on A. The class of all uniformly continuous, bounded functions on A (equipped with the norm of $L^\infty(\nu)$) will be denoted by $C(\nu; A)$. The class $X^p(\nu; A)$ will denote $L^p(\nu; A)$ if $1 \leq p < \infty$ and $C(\nu; A)$ if $p = \infty$.

In the sequel, we will assume that μ is a fixed, finite, positive, Borel measure with supp (μ) being an infinite subset of $[-1, 1]$. The mention of the set A will be omitted if $A = \text{supp }(\mu)$. Thus, for example, we will write $X^p(\mu) = X^p(\mu; \text{supp }(\mu))$ and $C(\mu) = C(\mu; \text{supp }(\mu))$.

We now formulate certain assumptions on our measures.

Definition 5.1 *A sequence $\{\nu_n\}$ will be called an **M–Z (Marcinkiewicz–Zygmund) sequence** if each of the following conditions is satisfied, with constants independent of n.*

1. *Each ν_n is a Borel, finite, positive or signed measure having bounded variation on $[-1, 1]$.*
2. $$\|T\|_{\nu_n;p} \leq c\|T\|_{\mu;p}, \qquad T \in \Pi_{16n},\ p = 1, \infty. \tag{5.1}$$
3. $$\int T_1 T_2 d\nu_n = \int T_1 T_2 d\mu, \qquad T_1, T_2 \in \Pi_{8n}. \tag{5.2}$$

*If $1 \leq p \leq \infty$, a sequence $\{\nu_n\}$ will be called **p-compatible** (with μ) if each μ–measurable function is also ν_n–measurable for each n, and $\|f\|_{\nu_n;p} \leq c\|f\|_{\mu;p}$ for every $f \in X^p(\mu)$.*

In the case of Jacobi polynomials, it is proved by Lubinsky, Máté, and Nevai [14, Theorem 5] that the measures ν_n that associate the mass $\lambda_{k,8n+1}$ with each $x_{k,8n+1}$, $k = 1, \cdots, 8n+1$, form an ∞–compatible M–Z sequence. In general, it is natural to construct measures to satisfy (5.2) using Gauss quadrature formulas based on the zeros of a sufficiently high degree orthogonal polynomial p_N. However, if supp (μ) is not an interval, then the zeros of the corresponding orthogonal polynomials might not be all in supp (μ), in which case, such a measure would not be ∞–compatible. We will prove the following proposition to demonstrate the existence of ∞–compatible, M–Z sequences of measures supported on finite subsets of supp (μ).

Proposition 5.1 *Let $\mu(\{x\}) = 0$ for every $x \in [-1, 1]$. Then there exists an ∞–compatible M–Z sequence $\{\nu_n\}$ of measures such that each of the sets* supp (ν_n) *is a finite subset of* supp (μ).

5.2 Potential theory ideas

In this section, we review briefly certain ideas from potential theory, based on the discussion in [28, Chapter 2.4]. The *logarithmic energy* of a positive measure ν on $\mathbb{C}$ is defined by

$$\mathcal{E}(\nu) := \int\int \log\left(\frac{1}{|x-y|}\right) d\nu(x)d\nu(y),$$

whenever the integral is well defined. For example, we recall (cf. [28, Chapter I, Example 3.5]) that

$$\int\int \log\left(\frac{2}{|x-t|}\right) d\mu^{(-1/2,-1/2)}(x)d\mu^{(-1/2,-1/2)}(t) = \pi^2 \log 4.$$

Since

$$\max_{x\in[-1,1]}(1-x)^a(1+x)^b = \left(\frac{2}{a+b}\right)^{a+b} a^a b^b, \qquad a,b \ge 0,$$

it follows that if $\alpha, \beta \ge -1/2$ and $\alpha+\beta+1 > 0$, then

$$\begin{aligned}
&\int\int \log\left(\frac{2}{|x-t|}\right) d\mu^{(\alpha,\beta)}(x)d\mu^{(\alpha,\beta)}(t)\\
\le\ & \left(\frac{2}{\alpha+\beta+1}\right)^{\alpha+\beta+1} (\alpha+1/2)^{\alpha+1/2}(\beta+1/2)^{\beta+1/2}\\
&\quad \times \int\int \log\left(\frac{2}{|x-t|}\right) d\mu^{(-1/2,-1/2)}(x)d\mu^{(-1/2,-1/2)}(t)\\
=\ & \pi^2(\log 4)\left(\frac{2}{\alpha+\beta+1}\right)^{\alpha+\beta+1} (\alpha+1/2)^{\alpha+1/2}(\beta+1/2)^{\beta+1/2}.
\end{aligned}$$

Thus, $\mu^{(\alpha,\beta)}$ has finite logarithmic energy.

If $A \subseteq \mathbb{C}$ is a compact set, the capacity of A, cap (A) is defined by

$$\log(1/\text{cap }(A)) = \inf \mathcal{E}(\nu),$$

where the infimum is taken over all unit, positive, Borel measures ν, with supp $(\nu) \subseteq A$. If $A \subset \mathbb{C}$ is compact, and cap $(A) > 0$, the infimum above is attained by a measure, called equilibrium measure of A, denoted by μ_A. We write

$$G_A(z) := \int \log \||x-z|d\mu_A(x) + \log(1/\text{cap }(A)). \tag{5.3}$$

For example, if $A = [a,b]$, one has cap $([a,b]) = (b-a)/4$, and

$$G_{[a,b]}(z) = \log\frac{|2z-a-b+2\sqrt{(z-a)(z-b)}|}{b-a}, \qquad z \in \mathbb{C}\setminus[a,b].$$

A point on the outer boundary of A (i.e., the boundary of the unbounded component of $\overline{\mathbb{C}}\setminus A$) is called *regular* if G_A is continuous at z, or equivalently, $G_A(z) = 0$. The set A called regular if each of the points on its outer boundary is regular. It is clear that every interval $[a,b]$ is a regular set. Examples of other regular and nonregular points and sets are given in [28].

We end this subsection by recalling the well known Bernstein–Walsh inequality (cf. [28, Estimate (2.4), p. 153]).

Lemma 5.1 *Let $A \subseteq [-1,1]$ be a regular set, $m \ge 0$ be an integer, $P \in \Pi_m$. Then for any $z \in \mathbb{C}$,*

$$|P(z)| \le \exp(mG_A(z)) \max_{x\in A}|P(x)|. \tag{5.4}$$

In particular, for any $x_0 \in \mathbb{R}$, $\ell > 0$, $L \ge \ell$,

$$\max_{x\in[x_0-L,x_0+L]}|P(x)| \le \left(\frac{2L}{\ell}\right)^m \max_{x\in[x_0-\ell,x_0+\ell]}|P(x)|. \tag{5.5}$$

5.3 Polynomial operators

It is well known [7, Chapter 1] that there exists a unique system of polynomials $p_k(x) = \gamma_k x^k + \cdots$, $\gamma_k > 0$, $k = 0, 1, 2, \cdots$, such that for $k, j = 0, 1, \cdots$,

$$\int p_k p_j d\mu = \begin{cases} 1, & \text{if } k = j, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover (cf. [7, Chapter 2]), any function $f \in L^1(\mu)$ is uniquely determined by the sequence of its coefficients

$$\hat{f}(k) := \int f p_k d\mu, \qquad k = 0, 1, \cdots. \tag{5.6}$$

Next, we define the kernels. For $x, y \in \mathbb{C}$, we will write

$$K_n(x, y) := \sum_{k=0}^{2^n} p_k(x) p_k(y), \qquad n = 0, 1, 2, \cdots, \tag{5.7}$$

and $K_{-1}(x, y) = 0$. For an integer $n \geq 0$, a function $\Phi_n : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ will be called a *reproducing summability kernel (of order n)*, if each of the following four conditions is satisfied. For each $x, y \in \mathbb{C}$, $\Phi_n(x, y) = \Phi_n(y, x)$, $\Phi_n(x, \circ) \in \Pi_{2n}$,

$$\int \Phi_n(x, y) P(y) d\mu(y) = P(x), \qquad P \in \Pi_n, \tag{5.8}$$

and

$$\sup_{x \in \text{supp }(\mu)} \int |\Phi_n(x, y)| d\mu(y) \leq c, \tag{5.9}$$

where c is a constant independent of n, depending at most on μ and the whole sequence $\{\Phi_n\}$. We assume in the sequel that there exists a sequence $\{\Phi_n\}$ of reproducing summability kernels.

In [18], we have proved that the kernels $\Phi_n(\mu^{(\alpha,\beta)}, h, x, y)$ defined in (2.1) are reproducing summability kernels if $\alpha, \beta \geq -1/2$. In [7, Section IV.3], Freud has shown the strong $(C, 1)$ summability of a very general class of orthogonal polynomials. If h is an integral of a function of bounded variation, $h(x) = 1$ for $0 \leq x \leq 1/2$, and $h(x) = 0$ for $x > 1$, it can be shown using a summation by parts argument that a kernel similar to the one defined in (2.1) with these orthogonal polynomials satisfies all the properties mentioned above. In the case of the Jacobi measure, we were able to use the special function properties of Jacobi polynomials to obtain localization estimates on the kernels $\Phi_n^{(\alpha,\beta)}$ (cf. [18]). These techniques cannot be used to obtain localization estimates on the kernels in general, for example, for the orthogonal polynomial systems discussed by Freud. Nevertheless, our simple construction below allows one to construct exponentially localized kernels based only on the summability estimates. In turn, the localization allows one to use the ideas in [18] to obtain a characterization of local Besov spaces on the interval also in the case of these more general systems of orthogonal polynomials.

For $x, y \in \mathbb{C}$, and $n = 0, 1, \cdots$, let

$$\Phi_n^*(x, y) = \left(\frac{4 - (x - y)^2}{4} \right)^n \Phi_{3n}(x, y).$$

If ν is a Borel, finite, positive or signed measure (with bounded variation), and $f \in L^1(\nu)$, we define the operators

$$\sigma_n(\nu; f, x) := \int \Phi_n^*(x, y) f(y) d\nu(y), \qquad x \in \mathbb{C},\ n = 0, 1, \cdots.$$

With $\nu = \mu^{(\alpha,\beta)}$ and $\Phi_n(\mu^{(\alpha,\beta)}, h, x, y)$ in place of $\Phi_n(x, y)$, $\sigma_n(\nu; f)$ reduces to $\sigma_n^C(\alpha, \beta; h, f)$. We observe that Φ_n^* being a symmetric polynomial in x and y, has an expansion of the form

$$\Phi_n^*(x, y) = \sum_{j=0}^{8n} \sum_{k=0}^{8n} a_{n;j,k} p_j(x) p_k(y),$$

where $(a_{n;j,k})$ is a symmetric matrix. Therefore, taking ν to be the measure μ, we see that

$$\sigma_n(\mu; f) = \sum_{j=0}^{8n} \left(\sum_{k=0}^{8n} a_{n;j,k} \hat{f}(k) \right) p_j$$

is a polynomial with coefficients given as a finite linear combination of the coefficients $\{\hat{f}(k)\}_{k=0}^{8n}$. The more general definition allows us to compute these operators using, for example, values of f.

Our generalization of Theorem 2.1 is the following.

Theorem 5.1 *Let $1 \le p \le \infty$ and $\{\nu_n\}$ be a p-compatible M–Z sequence of measures.*
(a) *We have $\sigma_n(\nu_n; P) = P$ for all $P \in \Pi_n$, and*

$$\|\sigma_n(\nu_n; f)\|_{\mu;p} \le c\|f\|_{\nu_n;p}, \qquad f \in L^p(\nu_n),\ 1 \le p \le \infty. \tag{5.10}$$

Further, for each $f \in X^p(\mu)$,

$$\|f - \sigma_n(\nu_n; f)\|_{\mu;p} \le cE_{n,p}(f). \tag{5.11}$$

(b) *Let $\mathcal{E}(\mu) < \infty$, supp (μ) be a regular set, $f \in C(\mu)$, $\{\nu_n\}$ be an ∞-compatible M–Z sequence of measures, $x_0 \in$ supp (μ), $0 < d \le 2$, and f have an analytic continuation to a complex neighborhood of $\{z \in \mathbb{C} : |z - x_0| \le d\}$ of x_0. Then*

$$|f(x) - \sigma_n(\nu_n; f, x)| \le c(f, x_0) \exp\left(-n\frac{d^2 \log(e/2)}{e^2 \log(e^2/d)}\right), \quad x \in [x_0 - d/e, x_0 + d/e] \cap \text{supp } (\mu). \tag{5.12}$$

If $\{\nu_n\}$ is a sequence of finite positive or signed Borel measures having a bounded variation on $[-1, 1]$, we define for $z \in \mathbb{C}$

$$\tau_n(\nu_{2^n}; f, z) = \begin{cases} \sigma_1(\nu_1; f, z), & \text{if } n = 0, \\ \sigma_{2^n}(\nu_{2^n}; f, z) - \sigma_{2^{n-1}}(\nu_{2^{n-1}}; f, z), & \text{if } n \ge 1. \end{cases} \tag{5.13}$$

Clearly, the operator τ_n depends upon two measures: ν_{2^n} and $\nu_{2^{n-1}}$. Although we have to mention the measure to distinguish between the general case and the continuous case, when each $\nu_{2^n} = \mu$, we prefer to keep the notation simpler rather than using the more cumbersome notation $\tau_n(\nu_{2^n}, \nu_{2^{n-1}}; f, x)$. In the Jacobi case, we choose $\Phi_n(\mu^{(\alpha,\beta)}; h, x, y)$ in place of $\Phi_n(x, y)$. Choosing each ν_n to be $\mu^{(\alpha,\beta)}$, we obtain $\tau_n(\mu^{(\alpha,\beta)}; f, x) = \tau_n^C(\alpha, \beta; h, f, x)$. We obtain $\tau_n^D(\alpha, \beta; h, f, x)$ by choosing ν_{2^n} to be the measure that associates the mass $\lambda_{k,2^{n+3}+1}$ with each $x_{k,2^{n+3}+1}$, $k = 1, \cdots, 2^{n+3} + 1$.

The following proposition, generalizing Proposition 2.1, shows a representation of any function in $X^p(\mu)$, $1 \le p \le \infty$, in terms of the operators and kernels introduced so far. The theorem uses two sequences of measures. The sequence $\{\nu_n\}$ is determined by the kind of information we have regarding the target function f. Thus, if one starts with the coefficients $\{\hat{f}(k)\}$, then each of the measures ν_n is equal to μ. On the other hand, if a set of values of the form $\{f(x_{k,N_n})\}$ are available at a system of points, we should choose ν_n to be an M–Z measure supported at the points $\{x_{k,N_n} : k = 1, \cdots, N_n\}$, if such a measure can be found. The choice of the sequence $\{\mu_n\}$ is required only to satisfy (5.2), and may be used judiciously to obtain a parsimoneous representation, or a representation with other desirable properties depending upon the application.

Proposition 5.2 *Let $1 \le p \le \infty$, $\{\nu_n\}$ be a p-compatible M–Z sequence, $\{\mu_n\}$ be a sequence of measures satisfying (5.2), and $f \in X^p(\mu)$. We have*

$$f = \sum_{n=0}^{\infty} \tau_n(\nu_{2^n}; f), \tag{5.14}$$

where the convergence of the series is in the norm of $X^p(\mu)$. In the case when each $\nu_{2^n} = \mu$, we have further

$$f = \sum_{n=0}^{\infty} \int \tau_n(\mu; f, y) \{K_{n+3}(\circ, y) - K_{n-1}(\circ, y)\} d\mu_{2^n}(y). \tag{5.15}$$

Moreover, for $f \in L^2(\mu)$, we have the frame property:

$$c_1\|f\|_{\mu;2}^2 \le \sum_{n=0}^{\infty} \|\tau_n(\mu;f)\|_{\mu;2}^2 = \sum_{n=0}^{\infty} \|\tau_n(\mu;f)\|_{\mu_{2^n};2}^2 \le c_2\|f\|_{\mu;2}^2. \tag{5.16}$$

Next, we describe the generalization of Theorem 2.2.

Theorem 5.2 *Let $\mathcal{E}(\mu) < \infty$, supp (μ) be a regular set, $x_0 \in$ supp (μ), $f \in C(\mu)$, and $\{\nu_n\}$ be an ∞-compatible M–Z sequence of measures.*
(a) *The function f has an analytic continuation to a complex neighborhood of x_0 if and only if there exists a non-degenerate interval I with $x_0 \in I$ such that*

$$\limsup_{n\to\infty} \|\tau_n(\nu_{2^n};f)\|_{\mu;\infty,I}^{1/2^n} < 1. \tag{5.17}$$

(b) *The function f has an analytic continuation to a complex neighborhood of x_0 if and only if there exists a non-degenerate interval I with $x_0 \in I$ such that*

$$\limsup_{n\to\infty} \|\tau_n(\nu_{2^n};f)\|_{\nu_{2^{n+3}};\infty,I}^{1/2^n} < 1. \tag{5.18}$$

We note that $\mathcal{E}(\mu) \ge \log(1/2)(\mu([-1,1]))^2$. The condition $\mathcal{E}(\mu) < \infty$ implies, in particular, that $\mu(\{x\}) = 0$ for each $x \in [-1,1]$. Thus, in view of Proposition 5.1, the measures $\{\nu_n\}$ as required in Theorem 5.2 always exist. We observe again that the operators $\tau_n(\nu_{2^n};f)$ are defined using global information about f; the coefficients $\{\hat{f}(k)\}$ in the case when each ν_{2^n} is equal to μ. Nevertheless, the exponential localization of these operators enables us to obtain the characterization of local analyticity of the function. Similar characterizations of local Besov spaces can also be obtained, using the ideas in [18].

6 Proofs

Since the results in Section 5 clearly generalize those in Section 2, and the results in Section 4 depend upon those in Section 2, we first prove the results in Section 5, then explain their application to obtain the results in Section 2, and finally, indicate the proofs of the results in Section 4.
Proof of Proposition 5.1. This proof follows the ideas in [21]. Without loss of generality, we assume that $\mu([-1,1]) = 1$. Let $n \ge 1$ be an integer. Since Π_n is a finite dimensional space, there exists a constant, to be denoted in this proof only by B_n such that

$$\int_{-1}^{1} |P'(t)|dt \le B_n \int_{-1}^{1} |P(t)|d\mu(t), \qquad P \in \Pi_n. \tag{6.1}$$

Our assumption that $\mu(\{x\}) = 0$ for each $x \in [-1,1]$ implies that the function $x \mapsto \mu([-1,x))$ is a continuous, nondecreasing function on $[-1,1]$, with the range of this function being $[0,1]$. Therefore, there exist intervals I_k with mutually disjoint interiors such that $[-1,1] = \cup I_k$, and $\mu(I_k) \le 1/(4B_n)$ for each I_k. In this proof only, let $\mathcal{I}$ be the set of integers k such that $I_k \cap$ supp (μ) is not empty, and we choose a point $x_k \in I_k \cap$ supp (μ) for each $k \in \mathcal{I}$. In view of (6.1), we have for any $P \in \Pi_n$,

$$\begin{aligned}
\left|\|P\|_{\mu;1} - \sum_{k\in\mathcal{I}} \mu(I_k)|P(x_k)|\right| &= \left|\sum_{k\in\mathcal{I}} \left(\int_{I_k} |P(t)|d\mu(t) - \int_{I_k} |P(x_k)|d\mu(t)\right)\right| \\
&\le \sum_{k\in\mathcal{I}} \int_{I_k} |P(t) - P(x_k)|d\mu(t) \le \sum_{k\in\mathcal{I}} \int_{I_k}\int_{I_k} |P'(u)|dud\mu(t) \\
&\le \frac{1}{4B_n}\int_{-1}^{1} |P'(u)|du \le (1/4)\|P\|_{\mu;1}.
\end{aligned}$$

Therefore,

$$(3/4)\|P\|_{\mu;1} \le \sum_{k\in\mathcal{I}} \mu(I_k)|P(x_k)| \le (5/4)\|P\|_{\mu;1}. \tag{6.2}$$

In this proof only, let N be the number of elements in $\mathcal{I}$, T denote the linear operator defined by $T(P) = (P(x_k))_{k\in\mathcal{I}}$, and V be the range of T. The estimates (6.2) imply that T is invertible on V. In this proof only, let x^* denote the linear functional defined on V by $x^*(T(P)) = \int Pd\mu$. We equip $\mathbb{R}^N$ by the norm $(r_k)_{k\in\mathcal{I}} \mapsto \sum_{k\in\mathcal{I}} \mu(I_k)|r_k|$. The estimates (6.2) imply that the dual norm of x^* with respect to this norm is bounded from above by 4/3. The Hahn–Banach theorem, together with the characterization of the dual of $\mathbb{R}^N$, implies the existence of $w_k \in \mathbb{R}$, $k \in \mathcal{I}$, such that the functional $(r_k)_{k\in\mathcal{I}} \mapsto \sum_{k\in\mathcal{I}} w_k r_k$ extends x^*, and has the dual norm bounded from above by 4/3. In this proof only, let μ_n be the measure that associates the mass w_k with x_k, $k \in \mathcal{I}$. It is easy to check that the total variation measure of μ_n associates the mass $|w_k|$ with each x_k, $k \in \mathcal{I}$. The statement that the functional $(r_k)_{k\in\mathcal{I}} \mapsto \sum_{k\in\mathcal{I}} w_k r_k$ extends x^* means that

$$\int Pd\mu_n = \int Pd\mu, \qquad P \in \Pi_n. \tag{6.3}$$

The statement about the dual norm means that $|w_k| \le (4/3)\mu(I_k)$ for $k \in \mathcal{I}$. Therefore, the estimates (6.2) imply that

$$\|P\|_{\mu_n;1} = \sum_{k\in\mathcal{I}} |w_k||P(x_k)| \le (4/3) \sum_{k\in\mathcal{I}} \mu(I_k)|P(x_k)| \le (5/3)\|P\|_{\mu;1}, \qquad P \in \Pi_n.$$

We note that each μ_n is supported on a finite subset of supp (μ). Therefore, each μ_n is trivially ∞–compatible. Setting $\nu_n = \mu_{16n}$, $n = 1, 2, \cdots$, we have thus shown that the sequence $\{\nu_n\}$ is an ∞-compatible M–Z sequence of measures, with each ν_n supported on a finite subset of supp (μ). □

Proof of Theorem 5.1(a). Let $P \in \Pi_n$, $x \in [-1, 1]$, and $Q_x \in \Pi_{3n}$ be defined by

$$Q_x(y) = P(y)\left(\frac{4 - (x-y)^2}{4}\right)^n, \qquad y \in \mathbb{R}.$$

Consequently, (5.2) and (5.8) imply that

$$\begin{aligned} \sigma_n(\nu_n; P, x) &= \int P(y)\left(\frac{4 - (x-y)^2}{4}\right)^n \Phi_{3n}(x, y)d\nu_n(y) \\ &= \int Q_x(y)\Phi_{3n}(x, y)d\nu_n(y) = \int Q_x(y)\Phi_{3n}(x, y)d\mu(y) \\ &= Q_x(x) = P(x). \end{aligned}$$

Since $|4 - (x-y)^2| \le 4$ for all $x, y \in [-1, 1]$, the conditions (5.1) and (5.9) imply that

$$\sup_{n\ge 0} \sup_{x\in \text{supp }(\mu)} \int |\Phi_n^*(x, y)|d|\nu_n|(y) \le \sup_{n\ge 0} \sup_{x\in \text{supp }(\mu)} \int |\Phi_{3n}(x, y)|d\mu(y) \le c. \tag{6.4}$$

Therefore, for any $f \in L^\infty(\nu_n)$ and $x \in$ supp (μ), we have

$$|\sigma_n(\nu_n; f, x)| \le \int |\Phi_n^*(x, y)||f(y)|d|\nu_n|(y) \le c\|f\|_{\nu_n;\infty}. \tag{6.5}$$

Thus, (5.10) is satisfied if $p = \infty$. If $f \in L^1(\nu_n)$ and $g \in L^\infty(\mu)$, we verify using Fubini's theorem that

$$\int \sigma_n(\nu_n; f, x)g(x)d\mu(x) = \int f(y)\sigma_n(\mu; g, y)d\nu_n(y).$$

Since $\sigma_n(\mu; g) \in \Pi_{8n}$, the condition (5.1) implies that $\|\sigma_n(\mu; g)\|_{\nu_n;\infty} \le c\|\sigma_n(\mu; g)\|_{\mu;\infty}$. Therefore, using (6.5) with μ in place of ν_n, we obtain that for every $f \in L^1(\nu_n)$ and $g \in L^\infty(\mu)$,

$$\begin{aligned} \left|\int \sigma_n(\nu_n; f, x)g(x)d\mu(x)\right| &\le \int |f(y)\sigma_n(\mu; g, y)|d|\nu_n|(y) \\ &\le \|\sigma_n(\mu; g)\|_{\nu_n;\infty}\|f\|_{\nu_n;1} \le c\|\sigma_n(\mu; g)\|_{\mu;\infty}\|f\|_{\nu_n;1} \\ &\le c\|g\|_{\mu;\infty}\|f\|_{\nu_n;1}. \end{aligned}$$

Therefore, the Hahn-Banach theorem implies that $\|\sigma_n(\nu_n; f)\|_{\mu;1} \le c\|f\|_{\nu_n;1}$ for every $f \in L^1(\nu_n)$. Thus, we have proved (5.10) for $p = 1, \infty$. An application of the Riesz–Thorin interpolation theorem [3, Theorem 1.1.1] now yields (5.10) for $1 < p < \infty$.

Consequently, for any $P \in \Pi_n$,

$$E_{8n,p}(f) \le \|f - \sigma_n(\nu_n; f)\|_{\mu;p} = \|f - P - \sigma_n(\nu_n; f - P)\|_{\mu;p} \le c\|f - P\|_{\mu;p}.$$

Since P is arbitrary, this implies (5.11). □

In order to prove Theorem 5.1(b), we need two lemmas. First, we recall a well known fact from the theory of approximation of analytic functions [27, Chapter IX, Section 3]. For the sake of completion, we will sketch a proof.

Lemma 6.1 *Let $x_0 \in \mathbb{R}$, $d > 0$, and f be analytic on the complex neighborhood $\{z \in \mathbb{C} : |z - x_0| \le d\}$. Then for every $\ell \in (0, d)$ and integer $n \ge 1$, there exists a polynomial $P \in \Pi_{n-1}$ such that*

$$|f(x) - P(x)| \le c(f, x_0, d/\ell)(\ell/d)^n, \qquad x \in [x_0 - \ell, x_0 + \ell]. \tag{6.6}$$

PROOF. With an appropriate affine transform, we may assume that $\ell = 1$, $x_0 = 0$, denote in this proof only, $\delta = d/\ell > 1$, and assume that f is analytic on the disc $|z| \le \delta$. Let P be the partial sum of the power series expansion of f around 0 of degree $n - 1$. Then for $x \in [-1, 1]$,

$$f(x) - P(x) = \frac{x^n}{2\pi i} \oint_{|\xi|=\delta} \frac{f(\xi)d\xi}{\xi^n(\xi - x)}.$$

Since $|x| \le 1$ and $|\xi - x| \ge |\xi| - |x| \ge \delta - 1$ for $x \in [-1, 1]$, $|\xi| = \delta$, we deduce that

$$|f(x) - P(x)| \le \frac{c(f, \delta)}{2\pi\delta^n(\delta - 1)} \oint_{|\xi|=\delta} |d\xi| = c(f, \delta)\delta^{-n}.$$

□

The next lemma helps us to estimate the norms of $\sigma_n(\nu_n; f)$ on small intervals in terms of the norms of f on slightly larger intervals.

Lemma 6.2 *Let $f \in C(\mu)$, $x_0 \in \text{supp}\ (\mu)$, $\ell \in (0, 2)$, $I = [x_0 - \ell, x_0 + \ell] \cap \text{supp}\ (\mu)$, $J = [x_0 - 2\ell, x_0 + 2\ell] \cap [-1, 1]$, and $\{\nu_n\}$ be an M–Z sequence of measures. Then for every integer $n \ge 0$ and $x \in I$,*

$$|\sigma_n(\nu_n; f, x)| \le c\|f\|_{\nu_n;\infty,J} + c_1 \left(\frac{4 - \ell^2}{4}\right)^n \|f\|_{\nu_n;\infty}. \tag{6.7}$$

PROOF. Let $x \in I$. If $y \in [-1, 1] \setminus J$, then

$$\frac{4 - (x - y)^2}{4} \le \frac{4 - \ell^2}{4},$$

and hence,

$$|\Phi_n^*(x, y)| \le \left(\frac{4 - \ell^2}{4}\right)^n |\Phi_{3n}(x, y)|.$$

Therefore, (5.1) and (5.9) imply that

$$\begin{aligned} \left|\int_{[-1,1]\setminus J} \Phi_n^*(x, y)f(y)d\nu_n(y)\right| &\le \left(\frac{4 - \ell^2}{4}\right)^n \|f\|_{\nu_n;\infty} \int |\Phi_{3n}(x, y)|d|\nu_n|(y) \\ &\le c_1 \left(\frac{4 - \ell^2}{4}\right)^n \|f\|_{\nu_n;\infty}. \end{aligned}$$

The estimate (6.4) implies that

$$\left|\int_J \Phi_n^*(x, y)f(y)d\nu_n(y)\right| \le \|f\|_{\nu_n;\infty,J} \int |\Phi_n^*(x, y)|d|\nu_n|(y) \le c\|f\|_{\nu_n;\infty,J}.$$

Together with the definition of $\sigma_n(\nu_n; f, x)$, we are thus led to (6.7). $\square$

PROOF OF THEOREM 5.1(b). In this proof, constants denoted by $c, c_1, \cdots$ may depend upon x_0 and f. In this proof only, let

$$m = n\frac{d^2}{e^2\log(e^2/d)}.$$

Since $d \le 2$ and $e^2\exp(-4/e^2) > 4.3 \ge d$, it is not difficult to see that $m \le n$. In view of Lemma 6.1, we find a polynomial $P \in \Pi_m = \Pi_{\lfloor m\rfloor}$, such that

$$|f(x) - P(x)| \le c(e/2)^{-m}, \qquad x \in [x_0 - 2d/e, x_0 + 2d/e]. \tag{6.8}$$

Hence, $|P(x)| \le c_1$ for $x \in [x_0 - 2d/e, x_0 + 2d/e]$, and (5.5) (used with 2 in place of L and $2d/e$ in place of ℓ) implies that $|P(x)| \le c_2(2e/d)^m$ for $x \in [x_0 - 2, x_0 + 2] \supseteq [-1, 1]$. Consequently,

$$|f(x) - P(x)| \le c_3(2e/d)^m, \qquad x \in \text{supp }(\mu). \tag{6.9}$$

Let $J := [x_0 - 2d/e, x_0 + 2d/e] \cap [-1, 1]$. Since the measures ν_n are ∞–compatible, the estimates (6.7) (with d/e in place of ℓ), (6.8), and (6.9) imply that for $x \in [x_0 - d/e, x_0 + d/e] \cap \text{supp }(\mu) =: I$ and integer $n \ge 1$,

$$\begin{aligned}
|\sigma_n(\nu_n; f - P, x)| &\le c_4\left\{\|f - P\|_{\nu_n;\infty,J} + \left(\frac{4 - d^2/e^2}{4}\right)^n \|f - P\|_{\nu_n;\infty}\right\}\\
&\le c_5\left\{(e/2)^{-m} + (2e/d)^m\left(\frac{4 - d^2/e^2}{4}\right)^n\right\}\\
&\le c_6\{(e/2)^{-m} + (2e/d)^m\exp(-nd^2/e^2)\}.
\end{aligned} \tag{6.10}$$

Since $m \le n$, $\sigma_n(\nu_n; P) = P$. Using (6.8) and (6.10) we conclude that for $x \in I$,

$$\begin{aligned}
|f(x) - \sigma_n(\nu_n; f, x)| &= |f(x) - P(x) - \sigma_n(\nu_n; f - P, x)|\\
&\le c_7\{(e/2)^{-m} + (2e/d)^m\exp(-nd^2/e^2)\}.
\end{aligned}$$

In view of our choice of m,

$$|f(x) - \sigma_n(\nu_n; f, x)| \le c_8\exp\left(-n\frac{d^2\log(e/2)}{e^2\log(e^2/d)}\right), \qquad x \in I.$$

$\square$

PROOF OF PROPOSITION 5.2. We note that $f \in X^p(\mu)$ implies that $E_{n,p}(f) \to 0$ as $n \to \infty$. The equation (5.14) follows from (5.13) and (5.11). Since $\tau_n(\mu; f) \in \Pi_{2^{n+3}}$, we verify easily that

$$\tau_n(\mu; f, x) = \int \tau_n(\mu; f, y)K_{n+3}(x, y)d\mu(y). \tag{6.11}$$

Further, since $K_{n-1}(x, y)$, as a function of y, is in $\Pi_{2^{n-1}}$, $\tau_n(\mu; K_{n-1}(x, \circ)) = 0$. Therefore,

$$\int \tau_n(\mu; f, y)K_{n-1}(x, y)d\mu(y) = \int f(z)\tau_n(\mu; K_{n-1}(x, \circ), z)d\mu(z) = 0,$$

and (6.11) may be rewritten in the form

$$\tau_n(\mu; f, x) = \int \tau_n(\mu; f, y)\left(K_{n+3}(x, y) - K_{n-1}(x, y)\right)d\mu(y).$$

Since μ_{2^n} satisfies the quadrature formula (5.2), this implies (5.15).

In the remainder of this proof only, let

$$\mathbb{P}_m(f, x) := \int f(y)\left(K_m(x, y) - K_{m-1}(x, y)\right)d\mu(y), \qquad x \in \mathbb{R},\ m = 0, 1, \cdots.$$

We note that the Parseval identity implies that

$$f = \sum_{m=0}^{\infty} \mathbb{P}_m(f), \qquad \|f\|_{\mu;2}^2 = \sum_{m=0}^{\infty} \|\mathbb{P}_m(f)\|_{\mu;2}^2,$$

where the convergence of the first series is in the sense of $L^2(\mu)$.

Next, we observe that $\mathbb{P}_m(P) = 0$ if $P \in \Pi_{2^{m-1}}$. If $n \ge 0$ is an integer, and $n+4 \le m$, then $\tau_n(\mu; f) \in \Pi_{2^{n+3}} \subseteq \Pi_{2^{m-1}}$, and $\mathbb{P}_m(\tau_n(\mu; f)) = 0$. Similarly, if $n-1 \ge m$, then for each $x \in \mathbb{R}$, $K_m(x,\circ) - K_{m-1}(x,\circ) \in \Pi_{2^m} \subseteq \Pi_{2^{n-1}}$, and hence, for each $x, t \in \mathbb{R}$,

$$\int (\Phi^*_{2^n}(y,t) - \Phi^*_{2^{n-1}}(y,t))\,(K_m(x,y) - K_{m-1}(x,y))\,d\mu(y) = 0.$$

Therefore, if $n-1 \ge m$, then

$$\begin{aligned}
&\mathbb{P}_m(\tau_n(\mu; f)) \\
&= \int \left(\int f(t)\,(\Phi^*_{2^n}(y,t) - \Phi^*_{2^{n-1}}(y,t))\,d\mu(t) \right) (K_m(x,y) - K_{m-1}(x,y))\,d\mu(y) \\
&= \int f(t) \int (\Phi^*_{2^n}(y,t) - \Phi^*_{2^{n-1}}(y,t))\,(K_m(x,y) - K_{m-1}(x,y))\,d\mu(y)d\mu(t) = 0.
\end{aligned}$$

Hence, (5.14) implies that for any integer $m \ge 0$,

$$\begin{aligned}
\|\mathbb{P}_m(f)\|_{\mu,2}^2 &= \left\| \sum_{n=0}^{\infty} \mathbb{P}_m(\tau_n(\mu; f)) \right\|_{\mu;2}^2 = \left\| \sum_{n=\max(0,m-3)}^{m} \mathbb{P}_m(\tau_n(\mu; f)) \right\|_{\mu;2}^2 \\
&\le \left(\sum_{n=\max(0,m-3)}^{m} \|\mathbb{P}_m(\tau_n(\mu; f))\|_{\mu;2} \right)^2 \\
&\le 4 \sum_{n=\max(0,m-3)}^{m} \|\mathbb{P}_m(\tau_n(\mu; f))\|_{\mu;2}^2 \\
&\le 4 \sum_{n=\max(0,m-3)}^{m} \|\tau_n(\mu; f)\|_{\mu;2}^2.
\end{aligned}$$

This implies

$$\|f\|_{\mu;2}^2 = \sum_{m=0}^{\infty} \|\mathbb{P}_m(f)\|_{\mu,2}^2 \le 4 \sum_{m=0}^{\infty} \sum_{n=\max(0,m-3)}^{m} \|\tau_n(\mu; f)\|_{\mu;2}^2 \le 16 \sum_{n=0}^{\infty} \|\tau_n(\mu; f)\|_{\mu;2}^2.$$

The proof of the second inequality in (5.16) is similar. Thus, arguing as before, we see that $\tau_n(\mu; \mathbb{P}_m(f)) = 0$ except when $n \le m \le n+3$. So, for any integer $n \ge 0$,

$$\begin{aligned}
\|\tau_n(\mu; f)\|_{\mu;2}^2 &= \left\| \sum_{m=0}^{\infty} \tau_n(\mu; \mathbb{P}_m(f)) \right\|_{\mu;2}^2 = \left\| \sum_{m=n}^{n+3} \tau_n(\mu; \mathbb{P}_m(f)) \right\|_{\mu;2}^2 \\
&\le \left(\sum_{m=n}^{n+3} \|\tau_n(\mu; \mathbb{P}_m(f))\|_{\mu;2} \right)^2 \le 4 \sum_{m=n}^{n+3} \|\tau_n(\mu; \mathbb{P}_m(f))\|_{\mu;2}^2 \\
&\le c \sum_{m=n}^{n+3} \|\mathbb{P}_m(f)\|_{\mu;2}^2.
\end{aligned}$$

Consequently,

$$\sum_{n=0}^{\infty} \|\tau_n(\mu; f)\|_{\mu;2}^2 \le c \sum_{m=0}^{\infty} \|\mathbb{P}_m(f)\|_{\mu;2}^2 = c\|f\|_{\mu;2}^2.$$

□

Proof of Theorem 5.2. We will prove part (a). The equivalence of (a) and (b) is a simple consequence of Lemma 6.2 and the fact that $\sigma_{2^{n+3}}(\nu_{2^{n+3}}; P) = P$ for all $P \in \Pi_{2^{n+3}}$. Let $2 > \ell > 0$, (5.17) be satisfied for $J = [x_0 - 2\ell, x_0 + 2\ell] \cap [-1, 1]$ in place of I, and $0 < \rho_1 < 1$ and integer N be chosen so that for all integer $n \geq N$,

$$\|\tau_n(\nu_{2^n}; f)\|_{\mu;\infty,J} < \rho_1^{2^n}.$$

Since $\tau_n(\nu_{2^n}; f) \in \Pi_{2^{n+3}}$ and $\|\tau_n(\nu_{2^n}; f)\|_{\mu;\infty} \leq c\|f\|_{\mu,\infty}$, we see from Lemma 6.2 applied with μ in place of ν_{2^n} that for every $x \in I := [x_0 - \ell, x_0 + \ell] \cap \text{supp }(\mu)$,

$$|\tau_n(\nu_{2^n}; f, x)| = |\sigma_{2^{n+3}}(\mu; \tau_n(\nu_{2^n}; f), x)| \leq c\rho_1^{2^n} + c_1(1 - \ell^2/4)^{2^n}\|f\|_{\mu,\infty}. \tag{6.12}$$

Since $x_0 \in \text{supp }(\mu)$, $\mu(I) > 0$, and hence, it is easy to see that the restriction of μ to I has a finite logarithmic energy. Therefore, cap $(I) > 0$. In view of a result of Wiener (cf. [28, Theorem 1.1, Appendix A]), I is a regular set. Letting $\rho := \max(\rho_1, (1 - \ell^2/4))$, we obtain from (5.4) and (6.12) that for every $z \in \mathbb{C}$, and integer $n \geq N$,

$$|\tau_n(\nu_{2^n}; f, z)| \leq c(f, x_0, \ell) \left(\rho \exp(G_I(z))\right)^{2^n}.$$

We observe that $0 < \rho < 1$. Therefore, the series $\sum_{n=0}^\infty \tau_n(\nu_{2^n}; f, z)$ converges uniformly and absolutely on compact subsets of the region $\{z \in \mathbb{C} \ : \ |G_I(z)| < \log(1/\rho)\}$. Since I is regular, this is an open neighborhood of I. In view of (5.14), the sum of this series is an analytic function that coincides with f on I.

The converse assertion follows immediately from Theorem 5.1(b) and the relevant definitions. □

It is proved by Lubinsky, Máté, and Nevai [14, Theorem 5] that the measures ν_n that associate the mass $\lambda_{k,2^{n+3}+1}$ with each $x_{k,2^{n+3}+1}$, $k = 1, \cdots, 2^{n+3}+1$, form an ∞–compatible M–Z sequence. Further, it is proved in [18] that the kernel $\Phi_n(\mu^{(\alpha,\beta)}; h, x, y)$ is a reproducing kernel of order n. Theorem 2.1 (respectively, Theorem 2.2) follows from Theorem 5.1 (respectively, Theorem 5.2) since $\mu^{(\alpha,\beta)}$ has finite logarithmic energy and $[-1, 1]$ is a regular set. Proposition 2.1 follows similarly from Proposition 5.2.

We now turn our attention to the proofs of the results in Section 4. The ideas are the same; we only sketch the proofs when the technical details are essentially different. First, we recall from [20, Theorem 3.3] that if each $\mathcal{C}_n$ admits a positive quadrature formula of order $16n$, and $w_{\mathbf{x}}$, $\mathbf{x} \in \mathcal{C}_n$, are nonnegative weights satisfying (4.1), then

$$\sum_{\mathbf{x}\in\mathcal{C}_n} w_{\mathbf{x}}|P(\mathbf{x})| \leq c\int_{\mathbb{S}^q} |P(\xi)|d\mu_q^*(\xi), \qquad P \in \Pi_{16n}^q.$$

Thus, the sequence of measures associating the weight $w_{\mathbf{x}}$ with $\mathbf{x} \in \mathcal{C}_n$ is the spherical analogue of a sequence of M–Z measures.

We further recall an analogue of the Bernstein–Walsh inequality from [17, Estimate (22)]: For $m = 0, 1, \cdots$, $0 < \alpha < \beta \leq \pi$,

$$\max_{\mathbf{x}\in\mathbb{S}_\beta^q} |P(\mathbf{x})| \leq \left(\frac{\pi(2\beta - \alpha)}{\alpha}\right)^{2m} \max_{\mathbf{x}\in\mathbb{S}_\alpha^q} |P(\mathbf{x})|, \qquad P \in \Pi_m^q. \tag{6.13}$$

In [17], it was assumed that $\beta < \pi$. However, the same estimate holds clearly for $\beta = \pi$ because of continuity.

Proposition 4.1 is proved exactly as Proposition 5.2. There are no new ideas involved, and we omit the proof.

Proof of Theorem 4.1. The proof of (4.3) is similar to that of (5.11). The estimate (4.4) is proved analogously to Theorem 5.1(b) with the following differences. The role of Lemma 5.1 is played by (6.13). The analogue of Lemma 6.2 can be proved in exactly the same way, using the fact that $(1 + \mathbf{x} \cdot \mathbf{y})^n < 2^n$ if $\mathbf{x} \neq \mathbf{y}$. The definition of the class $\mathcal{A}_q(\mathbf{x}_0)$ is the substitute for Lemma 6.1. This is the reason why the estimate (4.4) is evidently weaker than the estimate (5.12). Except for these technical differences, the proof of (4.4) follows that of (5.12) in exactly the same way. The proof of part (b) is similar to that of Theorem 5.2, with no new ideas. □

Acknowledgement.
We are grateful to Charles Chui and the referees for their many useful comments towards a substantial improvement in the first draft of this paper.

References

[1] V. Andrievskii and H.–P. Blatt, "Discrepancy of signed measures and polynomial approximation", Springer Verlag, New York/Berlin, 2002.

[2] 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** (2002), 305–319.

[3] J. Bergh and J. Löfström, "Interpolation spaces, an introduction", Springer Verlag, Berlin, 1976.

[4] C. K. Chui, "An introduction to wavelets", Academic Press, Boston, 1992.

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

[6] F. Filbir and W. Themistoclakis, *Polynomial approximation on the sphere using scattered data,* Accepted for publication in Math. Nachrichten.

[7] G. Freud, "Orthogonal Polynomials", Académiai Kiado, Budapest, 1971.

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

[9] A. Gelb and E. Tadmor, *Detection of edges in spectral data,* Appl. Comput. Harmon. Anal. **7** (1999), 101–135.

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

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

[12] Q. T. Le Gia and H. N. Mhaskar, *Polynomial operators and local approximation of solutions of pseudo-differential equations on the sphere,* Numer. Math. **103** (2006), 299–322.

[13] Q. T. Le Gia and H. N. Mhaskar, *Localized linear polynomial operators and quadrature formulas on the sphere,* http://www.calstatela.edu/faculty/hmhaska/postscript/quadconst4.pdf.

[14] D.S. Lubinsky, A. Máté and P. Nevai, *Quadrature sums involving pth powers of polynomials,* SIAM J. Math. Anal. **18** (1987), 531–544.

[15] M. Maggioni and H. N. Mhaskar, *Diffusion polynomial frames on metric measure spaces,* Accepted for publication in Appl. Comput. Harmon. Anal.

[16] A. I. Markushevich, "Theory of functions of a complex variable, Vol. 1", Prentice Hall, New Jersey, 1965.

[17] H. N. Mhaskar, *Local quadrature formulas on the sphere, II,* in "Advances in Constructive Approximation" (M. Neamtu and E. B. Saff Eds), Nashboro Press, Nashville, 2004, 333–344.

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

[19] H. N. Mhaskar, *On the representation of smooth functions on the sphere using finitely many bits,* Appl. Comput. Harmon. Anal. **18** (2005), 215–233.

[20] H. N. Mhaskar, *Weighted quadrature formulas and approximation by zonal function networks on the sphere,* J. Complexity **22** (2006), 348–370.

[21] H. N. Mhaskar, F. J. Narcowich and J. D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature,* Math. Comput. **70** (2001), 1113–1130. (Corrigendum: Math. Comput., **71** (2001), 453–454.)

[22] H. N. Mhaskar and D. V. Pai, "Fundmentals of approximation theory", CRC Press, 2000.

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

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

[25] 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.), Nashboro Press, Brentwood, 2005, 287–318.

[26] C. Müller, "Spherical harmonics", Lecture Notes in Mathematics, Vol. 17, Springer Verlag, Berlin, 1966.

[27] I. P. Natanson, "Constructive function theory, I", Frederick Ungar Publ., New York, 1964.

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

[29] E. M. Stein and G. Weiss, "Fourier analysis on Euclidean spaces", Princeton University Press, Princeton, New Jersey, 1971.

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

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

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

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

[34] 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.