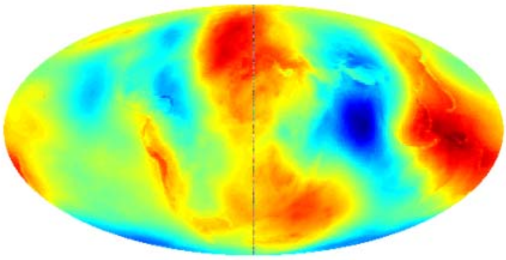
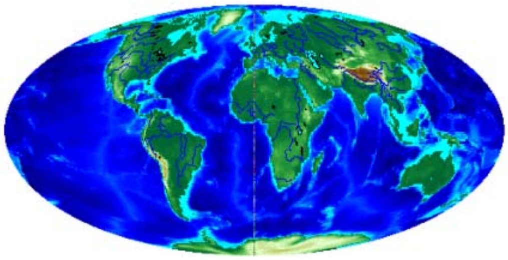




Siegen Preprints on Geomathematics

Optimally Localized Approximate Identities on the 2-Sphere

V. Michel

Geomathematics Group
Department of Mathematics
University of Siegen
Germany

www.geomathematics-siegen.de

3

Optimally Localized Approximate Identities on the 2-Sphere

Volker Michel
Geomathematics Group
Department of Mathematics
University of Siegen
Emmy-Noether-Campus
Walter-Flex-Str. 3
57068 Siegen
Germany
Email: michel@mathematik.uni-siegen.de

November 26, 2010

Abstract

We introduce a method to construct approximate identities on the 2–sphere which have an optimal localization. This approach can be used to accelerate the calculations of approximations on the 2–sphere essentially with a comparably small increase of the error. The localization measure in the optimization problem includes a weight function which can be chosen under some constraints. For each choice of weight function, existence and uniqueness of the optimal kernel are proved as well as the generation of an approximate identity in the bandlimited case. Moreover, the optimally localizing approximate identity for a certain weight function is calculated and numerically tested.

Key Words: approximate identity, optimal localization, sphere, fast approximation, convolution, existence, uniqueness, convergence, Legendre polynomials.

AMS(2000) Classification: 41A35, 41A55, 42C25, 65D15, 86-08

1 Introduction

Approximate identities yield one popular way of constructing approximations on the 2–sphere Ω (see, for example, [1, 2, 3]). This method is based on the calculation of spherical convolutions

$$(K_J * F)(\xi) := \int_\Omega K_J(\xi \cdot \eta) F(\eta)\, \mathrm{d}\omega(\eta), \quad \xi \in \Omega,$$

where $F \in \mathrm{L}^2(\Omega)$ is the function that has to be approximated and $K_J \in \mathrm{L}^2[-1,1]$ is a given kernel or, in other words, radial basis function. Some known criteria tell us when a sequence of such kernels $(K_J)_{J\in\mathbb{N}_0}$ establishes an approximate identity, which means that

$$\lim_{J\to\infty} \|K_J * F - F\|_{\mathrm{L}^2(\Omega)} = 0 \text{ for all } F \in \mathrm{L}^2(\Omega).$$

From the numerical point of view an approximate identity is connected to the implementation of a spherical quadrature rule in the sense that the right–hand side of

$$(K_J * F)(\xi) = \int_\Omega K_J(\xi \cdot \eta) F(\eta)\, \mathrm{d}\omega(\eta) \approx \sum_{k=1}^{N} w_k K_J(\xi \cdot \eta_k) F(\eta_k)$$

is regarded as an approximation to F at the single point $\xi \in \Omega$. Plotting the approximation on the sphere, i.e. at a point grid $\{\xi_j\}_{j=1,\ldots,M}$, consequently requires M spherical numerical integrations which is very expensive.

Earlier works already developed accelerations of several aspects of calculations on the sphere, see, for example, [4, 5]. We will address here another aspect, namely there is a perspective to accelerate the computations of the convolutions since the frequently used kernels K_J show some localization behavior, i.e. the function $\Omega \ni \eta \mapsto K_J(\xi \cdot \eta)$ has its maximum at $\eta = \xi$ and is almost zero if η is essentially distant to ξ. Therefore, each quadrature could be reduced to the neighborhood of the corresponding ξ instead of always taking the whole integration grid $\{\eta_k\}_{k=1,\ldots,N}$.

The localization of such radial basis functions varies from kernel to kernel. Several authors have studied the construction of optimally localizing basis functions, in particular on the sphere, see e.g. [6, 7, 8, 9, 10, 11, 12, 13]. Some of these works are based on an uncertainty principle on the 2–sphere (see [14] for an uncertainty principle). We will show in this paper that it is possible to construct bandlimited approximate identities with optimal localization based on a modified localization measure. Existence and uniqueness results are proved for the bandlimited as well as the non–bandlimited case. Moreover, a convergence proof, i.e. the evidence for an approximate identity, is given in the bandlimited case which is the relevant one for a numerical realization. Afterwards, the results of numerical tests are analyzed. We observe that the new approximate identity yields small approximation errors and shows at the same time a low sensitivity to the reduction of the integration grid. The new kernels represent a trade–off between two popular kernels (see, for example, [1]):

- the Shannon scaling function, which is an optimum in the sense that it yields a minimal approximation error $\|K_J * F - F\|_{\mathrm{L}^2(\Omega)}$ under some constraints on bandlimitation, but has a high sensitivity to the grid reduction,

- the cp scaling function (where “cp” refers to the used cubic polynomial symbol), which has a good localization and, thus, a low sensitivity to the omission of points but yields by itself a much higher approximation error.

Hence, the new kernels should be considered as an alternative in future calculations.

2 Nomenclature

By $\mathbb{R}$ and $\mathbb{N}$ we denote the set of all real numbers and all positive integers, respectively, such that $\mathbb{N}_0$ represents the set of all non–negative numbers. The standard Euclidean inner product and its induced norm in $\mathbb{R}^3$ are denoted by $x \cdot y$ and $|x|$, $x, y \in \mathbb{R}^3$, respectively. Moreover, the unit sphere in $\mathbb{R}^3$ is denoted by $\Omega := \{\xi \in \mathbb{R}^3 | \, |\xi| = 1\}$.
The system $\{P_n\}_{n\in\mathbb{N}_0}$ that is determined uniquely by the requirements

1. Every P_n is a univariate polynomial of degree n,
2. $\int_{-1}^{1} P_n(t)P_m(t)\,\mathrm{d}t = 0$ for all $n, m \in \mathbb{N}_0$ with $n \neq m$,
3. $P_n(1) = 1$ for each degree $n \in \mathbb{N}_0$,

is called the sequence of Legendre polynomials. Their norm is given by

$$\|P_n\|_{\mathrm{L}^2[-1,1]} = \sqrt{\frac{2}{2n+1}}\,.$$

The properties of these functions are well investigated in the literature of orthogonal polynomials, see e.g. [15].
Recall that the set $\mathrm{B}(D)$ of all bounded functions on a domain D is a Banach space if it is equipped with the norm $\|f\|_\infty := \sup_{x\in D} |f(x)|$, $f \in \mathrm{B}(D)$.

3 Approximate Identities on the 2–sphere

In this section we will briefly summarize known results on approximate identities on the 2–sphere. For further details we refer, for example, to [1, 2, 3, 16]. We consider here kernels that only depend on the spherical distance of the two arguments on the unit sphere Ω, i.e. functions $K \in \mathrm{L}^2[-1,1]$ of the form

$$\Omega^2 \ni (\xi, \eta) \mapsto K(\xi \cdot \eta).$$

Such kernels admit a representation in Legendre polynomials $\{P_n\}_{n\in\mathbb{N}_0}$ by

$$K(t) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, K^\wedge(n) P_n(t), \quad t \in [-1,1],$$

where the series converges in $\mathrm{L}^2[-1,1]$–sense and $K^\wedge(n)$, $n \in \mathbb{N}_0$, which will be called here the n–th Legendre coefficient of K, can be calculated by

$$K^\wedge(n) = 2\pi \int_{-1}^{1} K(t) P_n(t)\,\mathrm{d}t.$$

The tool for establishing approximate identities out of radial basis functions such as K is the spherical convolution.

Definition 3.1 *The spherical convolution of a kernel* $K \in \mathrm{L}^2[-1,1]$ *and a function* $F \in \mathrm{L}^2(\Omega)$ *is defined by*

$$(K * F)(\xi) := \int_\Omega K(\xi \cdot \eta) F(\eta) \, \mathrm{d}\omega(\eta), \quad \xi \in \Omega.$$

It is easy to verify by means of the Hölder inequality that $K * F \in \mathrm{L}^2(\Omega)$.

Definition 3.2 *A sequence* $(K_J)_{J \in \mathbb{N}_0}$ *of kernels in* $\mathrm{L}^2[-1,1]$ *is called an approximate identity if*

$$\lim_{J \to \infty} \|K_J * F - F\|_{\mathrm{L}^2(\Omega)} = 0,$$

for all $F \in \mathrm{L}^2(\Omega)$.

We prove the following variation of known criteria for approximate identities (see also [2, 3, 17]).

Theorem 3.3 *Let* $(K_J)_{J \in \mathbb{N}_0} \subset \mathrm{L}^2[-1,1]$ *be a given sequence such that there exists* $M \in \mathbb{R}^+$ *satisfying* $|K_J^\wedge(n)| \leq M$ *for all* $n, J \in \mathbb{N}_0$. *Then* $(K_J)_J$ *is an approximate identity if and only if*

$$\lim_{J \to \infty} K_J^\wedge(n) = 1 \text{ for all } n \in \mathbb{N}_0. \tag{1}$$

Proof. Let $F \in \mathrm{L}^2(\Omega)$ be arbitrary. It can be represented in terms of an orthonormal basis $\{Y_{n,j}\}_{n \in \mathbb{N}_0;\, j=1,\ldots,2n+1}$ of spherical harmonics (for further details see, e.g., [1, 18]) as

$$F = \sum_{n=0}^{\infty} \sum_{j=1}^{2n+1} \langle F, Y_{n,j} \rangle_{\mathrm{L}^2(\Omega)} \, Y_{n,j},$$

which yields

$$K_J * F = \sum_{n=0}^{\infty} \sum_{j=1}^{2n+1} K_J^\wedge(n) \, \langle F, Y_{n,j} \rangle_{\mathrm{L}^2(\Omega)} \, Y_{n,j},$$

where each series converges in $\mathrm{L}^2(\Omega)$–sense. Hence, the Parseval identity implies

$$\|K_J * F - F\|^2_{\mathrm{L}^2(\Omega)} = \sum_{n=0}^{\infty} \sum_{j=1}^{2n+1} \left(K_J^\wedge(n) - 1\right)^2 \langle F, Y_{n,j} \rangle^2_{\mathrm{L}^2(\Omega)} \, . \tag{2}$$

On the one hand, if $(K_J)_J$ is an approximate identity then choosing $F = Y_{m,k}$ ($m \in \mathbb{N}_0$, $k \in \{1, ..., 2m+1\}$ arbitrary but fixed) yields

$$0 = \lim_{J \to \infty} \|K_J * F - F\|^2_{\mathrm{L}^2(\Omega)} = \lim_{J \to \infty} \left(K_J^\wedge(m) - 1\right)^2 .$$

On the other hand, if (1) holds then we use the fact that

$$\begin{aligned} \left(K_J^\wedge(n) - 1\right)^2 \langle F, Y_{n,j} \rangle^2_{\mathrm{L}^2(\Omega)} &\leq (M+1)^2 \langle F, Y_{n,j} \rangle^2_{\mathrm{L}^2(\Omega)} \, , \\ \sum_{n=0}^{\infty} \sum_{j=1}^{2n+1} (M+1)^2 \langle F, Y_{n,j} \rangle^2_{\mathrm{L}^2(\Omega)} &= (M+1)^2 \|F\|^2_{\mathrm{L}^2(\Omega)} \end{aligned}$$

yields the uniform convergence of the series in (2) with respect to $J \in \mathbb{N}_0$ due to the Weierstraß criterion. This implies

$$\begin{aligned}\lim_{J\to\infty} \|K_J * F - F\|^2_{\mathrm{L}^2(\Omega)} &= \sum_{n=0}^{\infty}\sum_{j=1}^{2n+1} \lim_{J\to\infty} \left(K_J^\wedge(n) - 1\right)^2 \langle F, Y_{n,j}\rangle^2_{\mathrm{L}^2(\Omega)} \\ &= 0.\end{aligned}$$

■

4 Optimal Localization

4.1 Quantifying Localization

The question is how to measure localization on a sphere. We present here an alternative approach to those introduced in [6, 7, 8, 9, 13]. We choose a function $w : [-1,1] \to \mathbb{R}$, which is bounded and monotonically decreasing and satisfies $w(t) > 0$ for $t < 1$ and $w(1) = 0$, for example $w(t) = \sqrt{1-t}$. The non–localization of $K \in \mathrm{L}^2[-1,1]$ is then measured by

$$\mathcal{L}(K) := \int_{-1}^{1} w(t)(K(t))^2 \,\mathrm{d}t = \frac{1}{2\pi}\int_\Omega w(\xi\cdot\eta)(K(\xi\cdot\eta))^2 \,\mathrm{d}\omega(\eta),$$

$\xi \in \Omega$ arbitrary. The idea is that $K * F$ as an approximation to F can numerically be calculated by

$$(K * F)(\xi) = \int_\Omega K(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta) \approx \sum_{k=1}^{N} w_k K\left(\xi\cdot\eta_k\right) F\left(\eta_k\right)$$

by an appropriate quadrature method such as in [1, Chapter 7], [19], and [20]. If η_k is sufficiently far away from ξ then $K(\xi\cdot\eta_k)$ should be almost zero and, thus, negligible. The minimization of $\mathcal{L}$ should produce a kernel K with such a property. If $K(\xi\cdot\eta)$ were large for η far away from ξ, i.e. for $\xi\cdot\eta$ being much smaller than 1, then a large value $w(\xi\cdot\eta)$ would punish such a structure. The advantages are a faster calculation with an acceptable inaccuracy and an easier treatment of spherical data sets that only regionally change in time: The convolution only has to be recalculated in a slightly extended version of the area of change.

However, numerical experiments show that optimal localization alone does not yield an approximate identity in general. We, therefore, extend the functional $\mathcal{L}$ in the following way.

Definition 4.1 *A function $w : [-1,1] \to \mathbb{R}$ is called an admissible weight function if it is bounded and monotonically decreasing with $w(t) > 0$ for $t < 1$ and $w(1) = 0$. Moreover, let $\mu = (\mu_n) \subset \mathbb{R}_0^+$ with $\sum_{n=0}^{\infty}\mu_n < +\infty$, i.e. $\mu \in \mathrm{l}^1$, be a given sequence. The functional $\mathcal{F}_{w,\mu} : \mathrm{L}^2[-1,1] \to \mathbb{R}$ is then defined by*

$$\mathcal{F}_{w,\mu}(K) := \int_{-1}^{1} w(t)(K(t))^2\,\mathrm{d}t + \sum_{n=0}^{\infty}\mu_n\left(1 - K^\wedge(n)\right)^2, \quad K \in \mathrm{L}^2[-1,1].$$

In addition, let the operator $M_\mu : \mathbb{R}^\infty \to \mathbb{R}^\infty$ be defined by

$$M_\mu : (a_n) \mapsto (\mu_n a_n)\,,$$

where $\mathbb{R}^\infty := \{(a_n)_{n\in\mathbb{N}_0} | a_n \in \mathbb{R}\}$ is the set of all real sequences.

Note that we will discuss restrictions of M_μ to subdomains later.
Whereas the first part of the functional punishes a bad localization, the second one punishes a strong deviation from an approximate identity (cf. Theorem 3.3). Note that the sequence $(K^\wedge(n))_{n\in\mathbb{N}_0}$ of the Legendre coefficients of a function $K \in \mathrm{L}^2[-1,1]$ necessarily has to tend to 0, such that a series $\sum_{n=0}^{\infty}(1-K^\wedge(n))^2$ would always be infinite. This is compensated by the requirement that $\mu \in \mathrm{l}^1$.
The coefficients μ_n can control a balance between both objectives. The larger μ_n, the stronger the focus on the property of an approximate identity.

4.2 Bandlimited Kernels

Now we are able to prove the following existence and uniqueness result.

Theorem 4.2 *Let w be an admissible weight function, $N \in \mathbb{N}$, and $\mu \in \mathrm{l}^1$ with $\mu_n = 0$ for $n > N$. Then we have: Within the set $\mathrm{Pol}_N[-1,1]$ of all polynomials of degree $\leq N$ on $\mathbb{R}$, restricted to $[-1,1]$, the functional $\mathcal{F}_{w,\mu}$ has a unique global minimizer*

$$K^* = \sum_{n=0}^{N} (K^*)^\wedge(n)\, \frac{2n+1}{4\pi}\, P_n, \tag{3}$$

where the coefficient vector $k^* := ((K^*)^\wedge(n))_{n=0,\ldots,N} \in \mathbb{R}^{N+1}$ is the solution of the positive definite system of linear equations

$$(p + M_\mu)\, k^* = \mu, \tag{4}$$

where the entries $p_{n,m}$ $(n, m = 0, ..., N)$ of the matrix $p \in \mathbb{R}^{(N+1)\times(N+1)}$ are given by

$$p := \left(\frac{(2n+1)(2m+1)}{16\pi^2} \int_{-1}^{1} w(t) P_n(t) P_m(t)\,\mathrm{d}t \right)_{n,m=0,\ldots,N}.$$

Note that in (4) the projection of μ to $\mathbb{R}^{N+1}$ (i.e. $(\mu_0, ..., \mu_N)$) and the restriction of the mapping M_μ to $\mathbb{R}^{N+1} \to \mathbb{R}^{N+1}$ (represented by an $(N+1)\times(N+1)$–matrix) are used.

Proof. Since the Legendre polynomials are orthogonal in $\mathrm{L}^2[-1,1]$ and, thus, $\{P_n\}_{n=0,\ldots,N}$ is a basis of $\mathrm{Pol}_N[-1,1]$ we can write for $K \in \mathrm{Pol}_N[-1,1]$

$$\begin{aligned}\mathcal{F}_{w,\mu}(K) &= \langle wK, K\rangle_{\mathrm{L}^2[-1,1]} + \sum_{n=0}^{N} \mu_n \left(1 - K^\wedge(n)\right)^2 \\ &= \sum_{n=0}^{N}\sum_{m=0}^{N} \frac{2n+1}{4\pi}\,\frac{2m+1}{4\pi}\, K^\wedge(n)K^\wedge(m)\,\langle wP_n, P_m\rangle_{\mathrm{L}^2[-1,1]} + \sum_{n=0}^{N} \mu_n \left(1 - K^\wedge(n)\right)^2 \\ &= \sum_{n=0}^{N}\sum_{m=0}^{N} K^\wedge(n)K^\wedge(m) p_{n,m} + \sum_{n=0}^{N} \mu_n \left(1 - K^\wedge(n)\right)^2 \\ &= k^{\mathrm{T}} p k + (\mathbf{1} - k)^{\mathrm{T}} M_\mu (\mathbf{1} - k),\end{aligned}$$

where we used the abbreviations $k := (K^\wedge(n))_{n=0,\ldots,N} \in \mathbb{R}^{N+1}$ and $\mathbf{1} := (1, ..., 1)^{\mathrm{T}} \in \mathbb{R}^{N+1}$. Necessary for a minimum of $\mathcal{F}_{w,\mu}$ is, thus,

$$2pk + 2M_\mu(\mathbf{1} - k) \cdot (-1) = 0,$$

which is true if and only if

$$(p + M_\mu)\, k = \mu. \tag{5}$$

Since the weighted L^2–inner product

$$\langle f, g\rangle_{\mathrm{L}^2{}_w[-1,1]} := \int_{-1}^{1} f(t)g(t)w(t)\,dt$$

is still an inner product as $w \neq 0$ on $[-1, 1[$, the matrix p is a Gram matrix with the entries

$$p_{n,m} = \left\langle \frac{2n+1}{4\pi}\, P_n,\ \frac{2m+1}{4\pi}\, P_m \right\rangle_{\mathrm{L}^2{}_w[-1,1]}.$$

Due to the linear independence of the Legendre polynomials, p is positive definite and, thus, so is $p + M_\mu$, since $\mu_n \geq 0$ for all n. Hence, (5) has one and only one solution k^*. Since the Hessian $2p + 2M_\mu$ is positive definite, the solution k^* yields the only minimum. ■

It is clear that the values of μ may not be too small if one wants to obtain an approximate identity. A more detailed result yields the following theorem.

Theorem 4.3 *Let w be an admissible weight function and let $(\mu^{(J)})_{J\in\mathbb{N}_0} \subset \mathrm{l}^1$ be a sequence of non–negative finite sequences $\mu^{(J)} \in \mathrm{l}^1$ with $\mu_n^{(J)} = 0 \Leftrightarrow n > N_J$, where $\lim_{J\to\infty} N_J = \infty$. For each $J \in \mathbb{N}_0$, let K_J be the unique minimizer of $\mathcal{F}_{w,\mu^{(J)}}$ in $\mathrm{Pol}_{N_J}[-1,1]$ given by Theorem 4.2. If $(\mu^{(J)})_J$ satisfies*

$$\lim_{J\to\infty} \frac{\mu_n^{(J)}}{(N_J+1)^2} = \infty \tag{6}$$

for all $n \in \mathbb{N}_0$ and

$$\gamma := \sup_{n\in\mathbb{N}_0}\ \sup_{\substack{J\in\mathbb{N}_0\\ \text{with } n\leq N_J}} \frac{(N_J+1)^2}{\mu_n^{(J)}} < \infty, \tag{7}$$

then $\{K_J\}_{J\in\mathbb{N}_0}$ is an approximate identity.

Proof. We observe that choosing the so–called Shannon kernel $\Phi_J = \sum_{n=0}^{N_J} \frac{2n+1}{4\pi} P_n$ yields

$$\begin{aligned}
\mathcal{F}_{w,\mu^{(J)}}(\Phi_J) &= \int_{-1}^{1} w(t) \left(\sum_{n=0}^{N_J} \frac{2n+1}{4\pi} P_n(t) \right)^2 \mathrm{d}t \\
&\leq \|w\|_\infty \int_{-1}^{1} \left(\sum_{n=0}^{N_J} \frac{2n+1}{4\pi} P_n(t) \right)^2 \mathrm{d}t \\
&= \|w\|_\infty \sum_{n=0}^{N_J} \left(\frac{2n+1}{4\pi} \right)^2 \|P_n\|^2_{\mathrm{L}^2[-1,1]} \\
&= \|w\|_\infty \sum_{n=0}^{N_J} \frac{2n+1}{8\pi^2} \\
&= \|w\|_\infty \frac{(N_J+1)^2}{8\pi^2} \quad .
\end{aligned}$$

Assume that $K_J^\wedge(\tilde{n})$ does not converge to 1 as $J \to \infty$ for one $\tilde{n} \in \mathbb{N}_0$, i.e. there exists $\varepsilon > 0$ such that for all J_0 there is $J(J_0) \geq J_0$ with

$$\left| K^\wedge_{J(J_0)}(\tilde{n}) - 1 \right| > \varepsilon.$$

Due to condition (6) we find J_1 such that for all $J \geq J_1$ we have

$$\frac{\mu_{\tilde{n}}^{(J)}}{(N_J+1)^2} \geq \frac{\|w\|_\infty}{8\pi^2\varepsilon^2}.$$

Without loss of generality we can assume that $\tilde{n} \leq N_{J_1}$ since $\lim_{J\to\infty} N_J = \infty$. Hence, for $J = J(J_1)$ we obtain

$$\mathcal{F}_{w,\mu^{(J)}}(K_J) > \mu_{\tilde{n}}^{(J)} \varepsilon^2 \geq \frac{\|w\|_\infty}{8\pi^2} (N_J+1)^2 \geq \mathcal{F}_{w,\mu^{(J)}}(\Phi_J),$$

which contradicts the optimality of K_J. Thus, $\lim_{J\to\infty} K_J^\wedge(n) = 1$ for all $n \in \mathbb{N}_0$.
Moreover, from the considerations above we know that for $J \in \mathbb{N}_0$ and all $n \leq N_J$ we have

$$\mu_n^{(J)} \left(1 - K_J^\wedge(n)\right)^2 \leq \mathcal{F}_{w,\mu^{(J)}}(K_J) \leq \|w\|_\infty \frac{(N_J+1)^2}{8\pi^2}$$

and, hence,

$$\left(1 - K_J^\wedge(n)\right)^2 \leq \frac{\|w\|_\infty}{8\pi^2} \cdot \frac{(N_J+1)^2}{\mu_n^{(J)}} \quad .$$

Taking into account the bandlimitedness of the kernels and using

$$\left| K_J^\wedge(n) \right| = \left| -1 + 1 - K_J^\wedge(n) \right| \leq 1 + \left| 1 - K_J^\wedge(n) \right|$$

we obtain

$$\left|K_J^\wedge(n)\right| \le 1 + \frac{\sqrt{\|w\|_\infty}}{2\sqrt{2}\,\pi}\sqrt{\gamma}$$

for all $n \in \mathbb{N}_0$ and all $J \in \mathbb{N}_0$. Hence, Theorem 3.3 yields the desired result. ■

Note that (6) does not imply (7), as the following counterexample shows: If $N_J = J$ and $\mu_n^{(J)} = (J+1)^3 n^{-2}$ for all $n \le J$, then

$$\frac{\mu_n^{(J)}}{\left(N_J+1\right)^2} = \frac{J+1}{n^2} \longrightarrow \infty \text{ as } J \to \infty,$$

but

$$\sup_{n\in\mathbb{N}_0} \sup_{\substack{J\in\mathbb{N}_0\\ \text{with } n\le N_J}} \frac{\left(N_J+1\right)^2}{\mu_n^{(J)}} = \sup_{n\in\mathbb{N}_0} \sup_{\substack{J\in\mathbb{N}_0\\ \text{with } n\le N_J}} \frac{n^2}{J+1} = \sup_{n\in\mathbb{N}_0} \frac{n^2}{n+1} = \infty.$$

4.3 Non–bandlimited Kernels

In the following, we will investigate the treatment of non–bandlimited kernels where the previous matrices become operators between separable normed spaces. We will now assume that $(\mu_n) \subset \mathbb{R}^+$.

Definition 4.4 *For $\alpha \in \mathbb{R}$ the space $\mathrm{l}(\alpha)$ is defined as the space of all sequences (a_n) with*

$$\|(a_n)\|_{\mathrm{l}(\alpha)} := \sum_{n=0}^{\infty} (2n+1)^\alpha \,|a_n| < +\infty.$$

In analogy to the classical l^p spaces, $1 \le p \le \infty$, one can show that all $\mathrm{l}(\alpha)$ are Banach spaces. Obviously, $\mathrm{l}(\alpha) \subset \mathrm{l}(\beta)$ for $\alpha > \beta$ and $\mathrm{l}(\alpha) \subset \mathrm{l}^1 \subset \mathrm{l}^p \subset \mathrm{l}^\infty$ for $\alpha \ge 0$ and $1 \le p \le \infty$.

Definition 4.5 *Let w be an admissible weight function. We define the operator $A : \mathrm{l}(-1/2) \to \mathrm{l}(-1/2-\varepsilon)$, for $\varepsilon > 0$ fixed, by*

$$A\,(a_n) := \left(\sum_{m=0}^{\infty} a_m \left\langle P_n, P_m \right\rangle_{\mathrm{L}^2{}_w[-1,1]}\right)_{n\in\mathbb{N}_0}.$$

The image of A is indeed a subspace of $\mathrm{l}(-1/2-\varepsilon)$ since

$$\begin{aligned} \left|\left\langle P_n, P_m \right\rangle_{\mathrm{L}^2{}_w[-1,1]}\right| &\le \|w\|_\infty \,\|P_n\|_{\mathrm{L}^2[-1,1]} \,\|P_m\|_{\mathrm{L}^2[-1,1]} \\ &= \|w\|_\infty \,\frac{2}{\sqrt{(2n+1)(2m+1)}} \end{aligned}$$

due to the Hölder inequality. Consequently,

$$\begin{aligned} \left|\sum_{m=0}^{\infty} a_m \left\langle P_n, P_m \right\rangle_{\mathrm{L}^2{}_w[-1,1]}\right| &\le \left(\sum_{m=0}^{\infty} |a_m| \,\frac{1}{\sqrt{2m+1}}\right) \frac{2}{\sqrt{2n+1}} \,\|w\|_\infty \\ &= \|(a_m)\|_{\mathrm{l}(-1/2)} \,\frac{2}{\sqrt{2n+1}} \,\|w\|_\infty \end{aligned}$$

and, hence,

$$\|A\,(a_n)\|_{\mathrm{l}(-1/2-\varepsilon)} \leq \sum_{n=0}^{\infty}(2n+1)^{-\frac{1}{2}-\varepsilon}\,\frac{2}{\sqrt{2n+1}}\,\|(a_m)\|_{\mathrm{l}(-1/2)}\,\|w\|_\infty < +\infty .$$

Note that the last two inequalities imply that the sequence obtained by $A(a_n)$ even converges to 0 and that the linear operator A is bounded.
Furthermore, A may be applied to $(\frac{2m+1}{4\pi}\,a_m)$, if $(a_m) \in \mathrm{l}(1/2)$, since $\|(a_m)\|_{\mathrm{l}(1/2)} = \|((2m+1)a_m)\|_{\mathrm{l}(-1/2)}$. Hence, we obtain for a sequence $(a_m) \in \mathrm{l}(1/2)$ the estimate

$$\begin{aligned}\left|\left\langle\left(\frac{2n+1}{4\pi}\,a_n\right), A\left(\frac{2m+1}{4\pi}\,a_m\right)\right\rangle_{\mathrm{l}^2}\right| &= \left|\sum_{n=0}^{\infty}\frac{2n+1}{4\pi}\,a_n\sum_{m=0}^{\infty}\frac{2m+1}{4\pi}\,a_m\,\langle P_n, P_m\rangle_{\mathrm{L}^2{}_w[-1,1]}\right| \\ &\leq \sum_{n=0}^{\infty}\frac{2n+1}{4\pi}\,|a_n|\,\frac{2}{\sqrt{2n+1}}\,\|w\|_\infty\,\|(a_m)\|_{\mathrm{l}(1/2)}\,\frac{1}{4\pi} \\ &= \|(a_m)\|^2_{\mathrm{l}(1/2)}\,\frac{1}{8\pi^2}\,\|w\|_\infty < +\infty .\end{aligned}$$

Note that the operator

$$\begin{aligned} T : \mathrm{l}(\alpha) &\rightarrow \mathrm{l}(\alpha-1), \\ (a_n) &\mapsto \left(\frac{2n+1}{4\pi}\,a_n\right)\end{aligned}$$

is bijective. Moreover, for $a = (a_n) \in \mathrm{l}(1/2)$, we have due to the Beppo Levi theorem and the Hölder inequality that

$$\begin{aligned}\left\|\sum_{n=N}^{\infty}a_n\,\frac{2n+1}{4\pi}\,P_n\right\|^2_{\mathrm{L}^2{}_w[-1,1]} &\leq \sum_{n=N}^{\infty}\sum_{m=N}^{\infty}|a_n a_m|\,\frac{(2n+1)(2m+1)}{16\pi^2}\,\|w\|_\infty\,\frac{2}{\sqrt{(2n+1)(2m+1)}} \\ &= \|w\|_\infty\left(\sum_{n=N}^{\infty}|a_n|\,\frac{2n+1}{4\pi}\sqrt{\frac{2}{2n+1}}\right)^2 \longrightarrow 0 \quad \text{as } N\rightarrow\infty,\end{aligned}$$

i.e. the Legendre series $\sum_{n=0}^{\infty}a_n\,\frac{2n+1}{4\pi}\,P_n$ is strongly convergent in $\mathrm{L}^2{}_w[-1,1]$ sense and for analogous reasons also in ordinary $\mathrm{L}^2[-1,1]$ sense. Hence, the functional that has to be minimized may be written as

$$\mathcal{F}_{w,\mu}(K) = \langle Tk, A(Tk)\rangle_{\mathrm{l}^2} + \langle(\mathbf{1}-k), M_\mu(\mathbf{1}-k)\rangle_{\mathrm{l}^2}\,, \tag{8}$$

with $k := (K^\wedge(n))_{n\in\mathbb{N}_0} \in \mathrm{l}(1/2)$ and $\mathbf{1} := (1)_{n\in\mathbb{N}_0}$. The last expression in (8) is always finite, since

$$\sum_{n=0}^{\infty}\mu_n\left(1-K^\wedge(n)\right)^2 \leq \max_{m\in\mathbb{N}_0}\left(1+\left|K^\wedge(m)\right|\right)^2\sum_{n=0}^{\infty}\mu_n < +\infty$$

and since every $k \in \mathrm{l}(1/2)$ is bounded. Note that the l^2 inner product only serves here as a formal abbreviation since, in general, $A(Tk)$ may not be assumed to be in l^2 whereas the

occurring series nevertheless converges for that particular constellation.
We regard (8) as a functional depending on the sequence $k \in \mathrm{l}(1/2)$. For calculating its Fréchet derivative we consider

$$\begin{aligned}
&\langle T(k+h), AT(k+h)\rangle_{\mathrm{l}^2} - \langle Tk, ATk\rangle_{\mathrm{l}^2} \\
&\quad = \langle T(k+h), AT(k+h)\rangle_{\mathrm{l}^2} - \langle Tk, AT(k+h)\rangle_{\mathrm{l}^2} + \langle Tk, AT(k+h)\rangle_{\mathrm{l}^2} - \langle Tk, ATk\rangle_{\mathrm{l}^2} \\
&\quad = \langle Th, AT(k+h)\rangle_{\mathrm{l}^2} + \langle Tk, ATh\rangle_{\mathrm{l}^2}
\end{aligned}$$

for $h, k \in \mathrm{l}(1/2)$ such that, consequently,

$$\begin{aligned}
|\langle T(k+h), AT(k+h)\rangle_{\mathrm{l}^2} - \langle Tk, ATk\rangle_{\mathrm{l}^2} - \langle Th, ATk\rangle_{\mathrm{l}^2} - \langle Tk, ATh\rangle_{\mathrm{l}^2}|\, &\frac{1}{\|h\|_{\mathrm{l}(1/2)}} \\
&= |\langle Th, ATh\rangle_{\mathrm{l}^2}|\, \frac{1}{\|h\|_{\mathrm{l}(1/2)}} \\
&\leq \|h\|_{\mathrm{l}(1/2)}\, \frac{1}{8\pi^2}\, \|w\|_\infty .
\end{aligned}$$

The mapping

$$\begin{aligned}
\mathrm{l}(1/2) \ni h \quad \mapsto \quad & \langle Th, ATk\rangle_{\mathrm{l}^2} + \langle Tk, ATh\rangle_{\mathrm{l}^2} \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, h_n \sum_{m=0}^{\infty} \frac{2m+1}{4\pi}\, K^\wedge(m) \langle P_n, P_m\rangle_{\mathrm{L}^2{}_w[-1,1]} \\
&\quad + \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, K^\wedge(n) \sum_{m=0}^{\infty} \frac{2m+1}{4\pi}\, h_m \langle P_n, P_m\rangle_{\mathrm{L}^2{}_w[-1,1]} \\
&= 2\langle Th, ATk\rangle_{\mathrm{l}^2}
\end{aligned}$$

consequently, represents the Fréchet derivative of the first part of the functional in $k \in \mathrm{l}(1/2)$. This mapping is indeed bounded, since

$$|\langle Th, ATk\rangle_{\mathrm{l}^2}| \leq \|h\|_{\mathrm{l}(1/2)} \|k\|_{\mathrm{l}(1/2)}\, \frac{1}{8\pi^2}\, \|w\|_\infty .$$

For the second part one can easily derive that

$$\begin{aligned}
\frac{1}{\|h\|_{\mathrm{l}(1/2)}} &\left| \sum_{n=0}^{\infty} \mu_n \left(1 - K^\wedge(n) - h_n\right)^2 - \sum_{n=0}^{\infty} \mu_n \left(1 - K^\wedge(n)\right)^2 - 2\sum_{n=0}^{\infty} \mu_n h_n \left(K^\wedge(n) - 1\right) \right| \\
&\leq \frac{1}{\|h\|_{\mathrm{l}(1/2)}} \sum_{n=0}^{\infty} \mu_n h_n^2 \\
&\leq \frac{1}{\|h\|_{\mathrm{l}(1/2)}}\, \|h\|_\infty\, \|\mu\|_\infty \sum_{n=0}^{\infty} |h_n| \\
&\leq 2\|\mu\|_\infty \|h\|_{\mathrm{l}(1/2)}
\end{aligned}$$

since $\sup_{m\in\mathbb{N}_0} |h_m| \leq 2\|h\|_{\mathrm{l}^1} \leq 2\|h\|_{\mathrm{l}(1/2)}$. Note that the linear mapping $\mathrm{l}(1/2) \ni h \mapsto \langle M_\mu k - \mu, h\rangle_{\mathrm{l}^2}$ is bounded since

$$\begin{aligned}\left|\langle M_\mu k - \mu, h\rangle_{\mathrm{l}^2}\right| &= \left|\sum_{n=0}^{\infty}\mu_n h_n \left(K^\wedge(n) - 1\right)\right| \\ &\leq \|\mu\|_\infty \|k - \mathbf{1}\|_\infty \sum_{n=0}^{\infty} |h_n| \\ &\leq \|\mu\|_\infty \left(\|k\|_\infty + 1\right) \|h\|_{\mathrm{l}(1/2)}.\end{aligned}$$

The Fréchet derivative of the whole functional at $k \in \mathrm{l}(1/2)$ is now

$$\begin{aligned}\mathrm{l}(1/2) \ni h &\mapsto 2\langle Th, ATk\rangle_{\mathrm{l}^2} + 2\left\langle M_\mu k - \mu, h\right\rangle_{\mathrm{l}^2} \\ &= 2\left\langle h, TATk + M_\mu k - \mu\right\rangle_{\mathrm{l}^2}.\end{aligned}$$

A necessary condition for a minimum is, thus,

$$(TAT + M_\mu)k = \mu.$$

The second Fréchet derivative is obviously constant with respect to k and given by

$$(\mathrm{l}(1/2))^2 \ni (h_1, h_2) \mapsto 2\left\langle h_1, TATh_2 + M_\mu h_2\right\rangle_{\mathrm{l}^2}.$$

Due to the properties of μ, the operator M_μ is strictly positive, i.e. $\langle M_\mu a, a\rangle_{\mathrm{l}^2} > 0$ for all $a \in \mathrm{l}(1/2) \setminus \{0\}$. Moreover, if $a \in \mathrm{l}(1/2)$, then

$$\langle a, TATa\rangle_{\mathrm{l}^2} = \sum_{n=0}^{\infty} a_n \frac{2n+1}{4\pi} \sum_{m=0}^{\infty} \frac{2m+1}{4\pi}\, a_m \langle P_n, P_m\rangle_{\mathrm{L}^2{}_w[-1,1]}.$$

Due to the strong convergence of the Legendre series we are allowed to conclude that

$$\begin{aligned}\langle a, TATa\rangle_{\mathrm{l}^2} &= \left\langle \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, a_n P_n, \sum_{m=0}^{\infty} \frac{2m+1}{4\pi}\, a_m P_m \right\rangle_{\mathrm{L}^2{}_w[-1,1]} \\ &= \left\| \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, a_n P_n \right\|^2_{\mathrm{L}^2{}_w[-1,1]}.\end{aligned}$$

Due to the linear independence of the Legendre polynomials this expression vanishes if and only if $a = 0$. Consequently, $TAT + M_\mu$ is strictly positive and invertible on its image. Provided that $\mu \in (TAT + M_\mu)(\mathrm{l}(1/2))$, we have a unique solution

$$k = (TAT + M_\mu)^{-1}\, \mu.$$

Since $\langle h, TATh + M_\mu h\rangle_{\mathrm{l}^2} > 0$ for all $h \in \mathrm{l}(1/2) \setminus \{0\}$ we have a minimum of the functional. We have, thus, proved the following theorem.

Theorem 4.6 *Let the operators T, A, and M_μ be defined as introduced above where $\mathrm{l}^1 \ni \mu = (\mu_n) \subset \mathbb{R}^+$. If $\mu \in (TAT + M_\mu)(\mathrm{l}(1/2))$ then $\mathcal{F}_{w,\mu}$ has a unique global minimizer in the set of all $K \in \mathrm{L}^2[-1,1]$ with $k = (K^\wedge(n))_{n\in\mathbb{N}_0} \in \mathrm{l}(1/2)$. This is given by*

$$k = (TAT + M_\mu)^{-1}\, \mu.$$

5 Numerical Results

5.1 Determination of Optimal Kernels

Based on the results derived above optimal bandlimited kernels of different degrees were computed by solving the corresponding system of linear equations given by (4). The weight function $w(t) = \sqrt{1-t}$ was chosen based on preliminary numerical tests with various weight functions. Note that the behavior of the root function already allows a high penalty at small distances. The corresponding integrals $\langle P_m, P_n \rangle_{\mathrm{L^2}_w[-1,1]}$ were calculated numerically via the composite Newton–Cotes formula of order 4 with 1,000,001 grid points (see Appendix A). It should be noted that there also exists the following explicit representation in terms of hypergeometric functions (see [21], p. 447).

$$\int_{-1}^{1} \sqrt{1-t} P_m(t) P_n(t)\,\mathrm{d}t = \frac{\left(-\frac{1}{2}\right)_n}{\left(\frac{3}{2}\right)_{n+1}} 2^{3/2} {}_4F_3\left(-m, m+1, \frac{3}{2}, \frac{3}{2}; 1, n+\frac{5}{2}, \frac{3}{2}-n; 1\right).$$

For degrees $N_J = 2^J$, $J \geq 4$, the spectral weights were defined as

$$\mu_n^{(J)} := \begin{cases} 10^{-4}(\log J)\,(N_J+1)^2 & , n \leq N_J \\ 0 & , n > N_J \end{cases} = \begin{cases} 10^{-4}(\log(\log_2 N_J))\,(N_J+1)^2 & , n \leq N_J \\ 0 & , n > N_J \end{cases},$$

where $\log = \log_e$ represents the natural logarithm. Note that the right-hand representation of $\mu^{(J)}$ also allows the use of arbitrary, i.e. not necessarily dyadic, degrees $N \in \mathbb{N}$. The choice of this sequence is based on preliminary numerical tests and on the idea that the condition of an approximate identity should not be weighted much stronger than required by Theorem 4.3 to keep a good localization. In Figure 1, we represent the obtained kernels for certain selected degrees. Note that the functions show a very good localization relative to their degree and simultaneously tend to Legendre coefficients 1 with increasing scale J. The conditions of the obtained matrices are shown in Table 1. All calculations presented here were done with MATLAB.

degree	condition
16	11.9348
32	6.8289
64	3.8448
128	2.3625
256	1.6514

Table 1: Condition of the Matrices p for selected total degrees

5.2 Application to a Test Example

We used the determined optimal kernels for spherical convolutions with the function $f(x) = \sin(34x_1)\cos(59x_2)\sin(12x_3)$, since it contains short, medium, and long wavelength parts. The

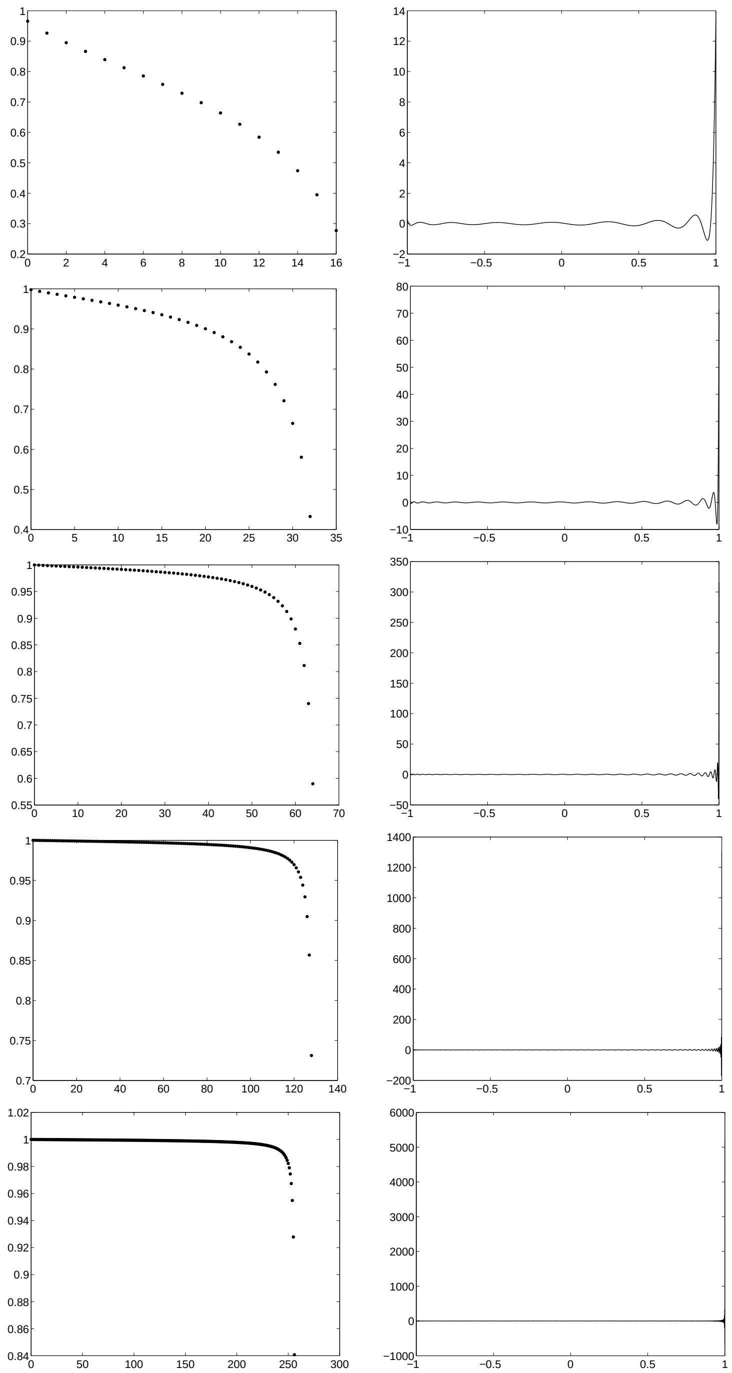


Figure 1: Coefficients (left column) and plots (right column) of the calculated optimal kernels for degrees $N_J = 2^J$ with $J = 4$ (top row)$, ..., 8$ (bottom row)

results are compared with those obtained for convolutions with the Shannon kernels Φ_J^{Sh} and the cp–kernels Φ_J^{cp} whose coefficients are given by a cubic polynomial:

$$\left(\Phi_J^{\mathrm{Sh}}\right)^\wedge(n) = \begin{cases} 1, & n \leq N_J \\ 0, & n > N_J \end{cases}, \quad \left(\Phi_J^{\mathrm{cp}}\right)^\wedge(n) = \begin{cases} \left(1-\frac{n}{N_J}\right)^2\left(1+2\,\frac{n}{N_J}\right), & n \leq N_J \\ 0, & n > N_J \end{cases}.$$

Note that here $N_J = 2^J$. The Shannon kernel is a kind of bandlimited optimum as approximate identity since its coefficients are 1 up to the maximal degree, whereas its localization behavior is known to be bad. On the other hand, the cp–kernel is very popular for approximations on the sphere due to its nice behavior in space.

We investigate two kinds of errors. In Figure 2 we show the mean square error that is obtained if one compares the original function f with its approximation $\Phi_J * f$ for the different choices of kernels Φ_J (optimal kernel K_J, Shannon kernel Φ_J^{Sh}, and cp–kernel Φ_J^{cp}) relative to the truncation of the quadrature grid, which is further explained below. All approximations $\Phi_J * f$ were calculated at an equiangular grid $P \subset \Omega$ of 126×126 points $\xi \in \Omega$. For each value $(\Phi_J * f)(\xi)$, an equiangular integration grid (see the quadrature method by Driscoll and Healy [19]) $Q \subset \Omega$ of 502×502 points η_k was used to represent

$$\int_\Omega \Phi_J(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta) \approx \sum_{k=1}^{502^2} w_k \Phi_J\left(\xi\cdot\eta_k\right)F\left(\eta_k\right),$$

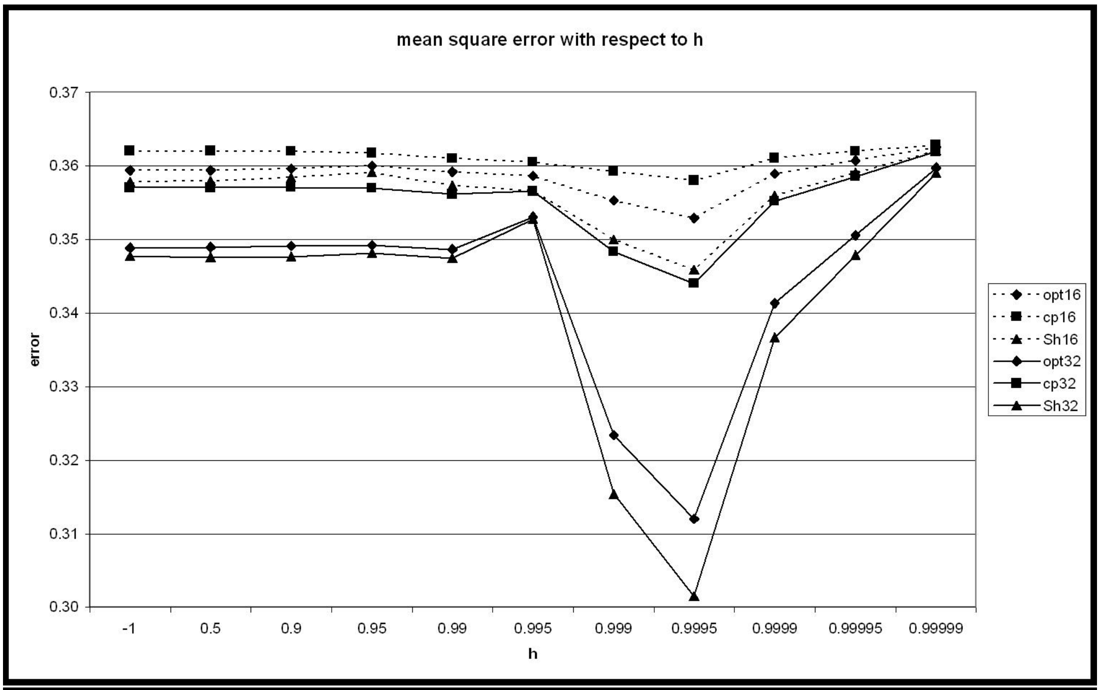
which theoretically yields an exact integration of polynomials up to degree 501. For purposes of acceleration, Φ_J was calculated via the Clenshaw algorithm (see [22]) at an equiangular grid of 40,001 points on $[-1,1]$ and linearly interpolated in between. To test the advantage of localization, we gradually omitted quadrature points η_k. More precisely, we constructed a spherical rectangle $R_{\xi,h}$ which covers a spherical cap given by $\{\eta \in \Omega | \xi\cdot\eta \geq h\}$ at the fixed point $\xi \in \Omega$ relative to a parameter $h \in [-1,1]$, see Figure 3 and Appendix B. This rectangle is used to accelerate the numerical integration on the sphere as follows:

$$\left(\Phi_J * f\right)(\xi) = \int_{S^2} \Phi_J(\xi\cdot\eta)f(\eta)\,\mathrm{d}\omega(\eta) \approx \sum_{\substack{k=1 \\ \eta_k\in R_{\xi,h}}}^{502^2} \alpha_k \Phi_J\left(\xi\cdot\eta_k\right)f\left(\eta_k\right), \tag{9}$$

i.e. we always only use values from the neighborhood of ξ since $\Phi_J(\xi\cdot\eta)$ should be almost 0 if η is far away from ξ. The restriction of the integration grid essentially accelerates the calculations. In Figure 4, the base 10 logarithm of the total number of used points relative to h is plotted. This value

$$\sum_{\xi\in P}\sum_{\eta\in Q\cap R_{\xi,h}} 1$$

is a measure for the calculation time, since it represents the total number of summands in (9). We observe that the list of values for h used for our computations corresponds approximately to an exponential decay of the calculation time. Figure 2 now shows that, as a general trend, the



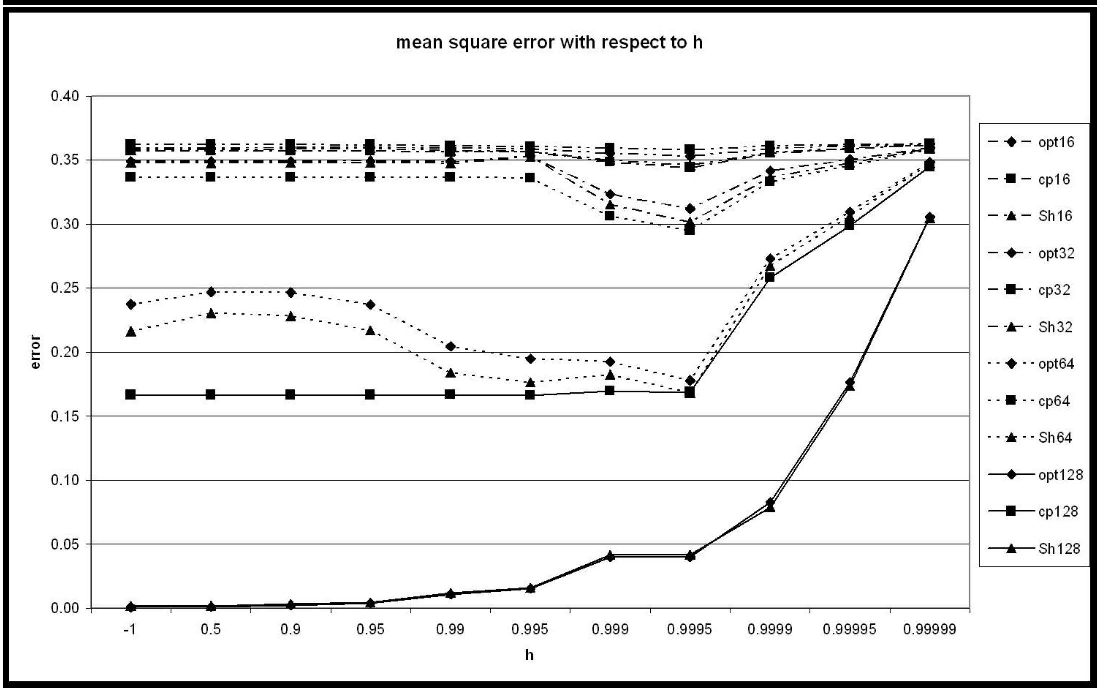


Figure 2: Root mean square error *function - truncated convolution* relative to h, note that $h = -1$ means that there is no spatial truncation in the quadrature rule.

restriction of the integration grid increases the approximation error, as one can expect. Obviously, the cp kernel produces the largest approximation error and the Shannon kernel generates the smallest one. Compared to the exponential decay of the calculation time, the increase of the error is relatively small. This holds, in particular, also for the new kernels.

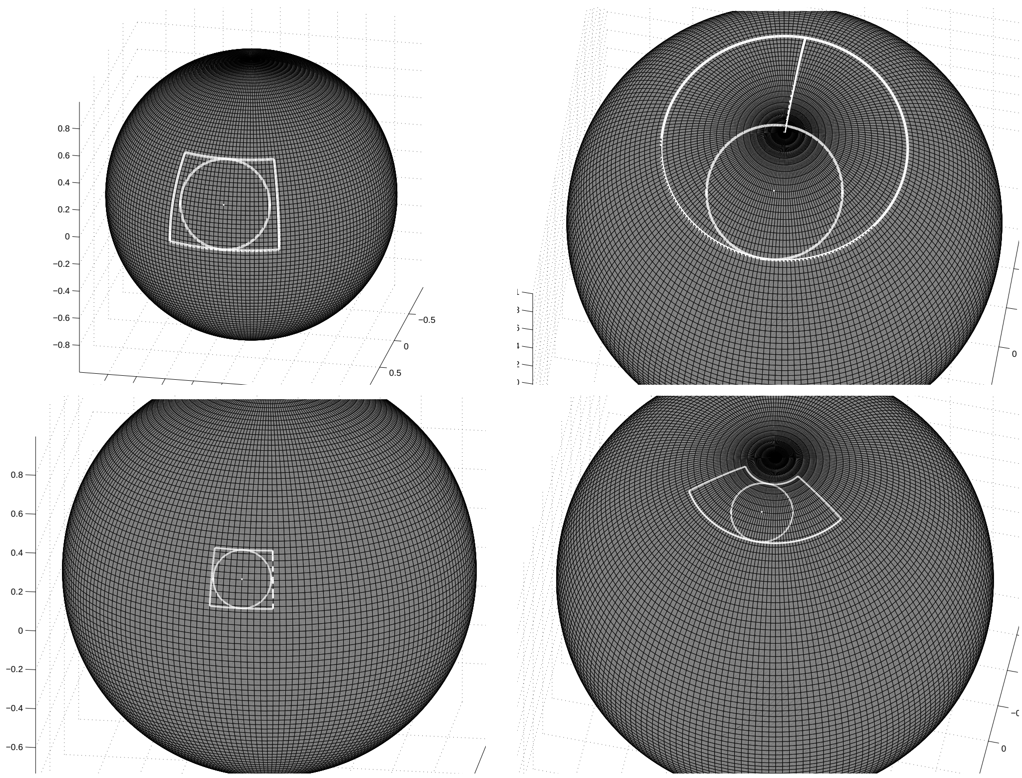


Figure 3: Illustration of the spherical cap and the larger spherical rectangle, defined via longitude and latitude limits, for $h = 0.95$ (top row) and $h = 0.99$ (bottom row) including the special case of caps including or near a pole (right column)

On the other hand, the approximation of a function by a convolution with kernels of different degrees can be regarded as a multiresolution analysis in terms of low–pass filters. Hence, the investigation of the convolution $\Phi_J * f$ itself can yield additional information. Therefore, we also investigate how close the restricted, i.e. accelerated, convolutions are to the function $\Phi_J * f$ of interest. In other words, we determine the pure truncation error, i.e. the difference between $\Phi_J * f$ obtained by always integrating over the whole sphere ($h = -1$) and its counterpart with restricted area ($h > -1$). The result is shown in Figure 5. We clearly see that the Shannon kernel shows the highest sensitivity with respect to the omission of quadrature points. The new kernels are less sensitive due to their good localization. The lowest sensitivity is observed in case of the cp kernels, which, however, yield an approximation which is most distant to f.

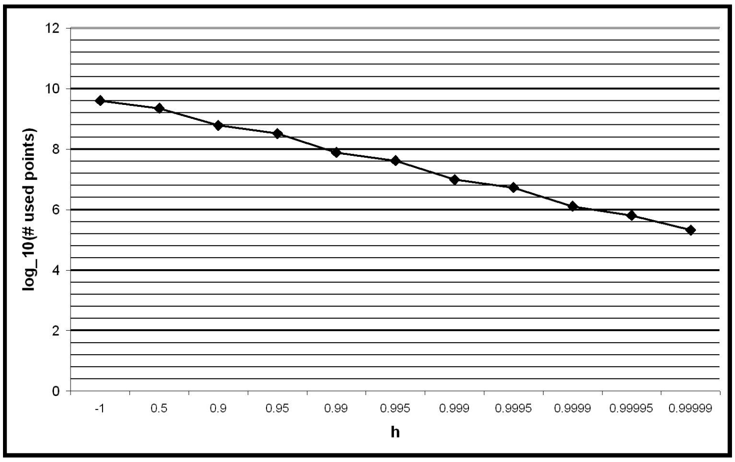


Figure 4: Base 10 logarithm of the total number of used quadrature points relative to h

6 Conclusions

A method for constructing optimally localizing approximate identities on the 2–sphere was constructed. Existence, uniqueness and convergence results were proved. The obtained example of kernels shows the desired trade-off of good approximation and fast calculations. The calculated approximations of the used test function are almost as good as those obtained by the Shannon scaling function and much better than for the cp–scaling function. Additionally, if this approximation process is accelerated by merely taking into account quadrature points in the neighborhood of the location of interest the new kernels show a smaller increase of the error than the Shannon kernel does.

Thus, the optimal kernels allow a good compromise between the (kind of) extremal situations of the Shannon kernel, which gives good approximations but reacts highly sensitive to a grid reduction, and the cp–kernel, which shows a low sensitivity to the omission of grid points but yields in general approximations with a higher error compared to the original function.

Therefore, the obtained kernels yield an interesting alternative to the presently known tools for approximation on the 2–sphere. They, in particular, allow fast approximations with a relatively low error.

7 Further Perspectives

Note that there are some interesting topics for further research. For instance, at least two theoretical problems remain. One is the construction of a corresponding sequence that can be proved to be an approximate identity in the non–bandlimited case. This would promise a further improved localization (see the uncertainty principle in [14]). And the second one is based on the observation that the numerically determined kernels of degrees $4, ..., 256, 300, 350, 400, 450, 500, 512$

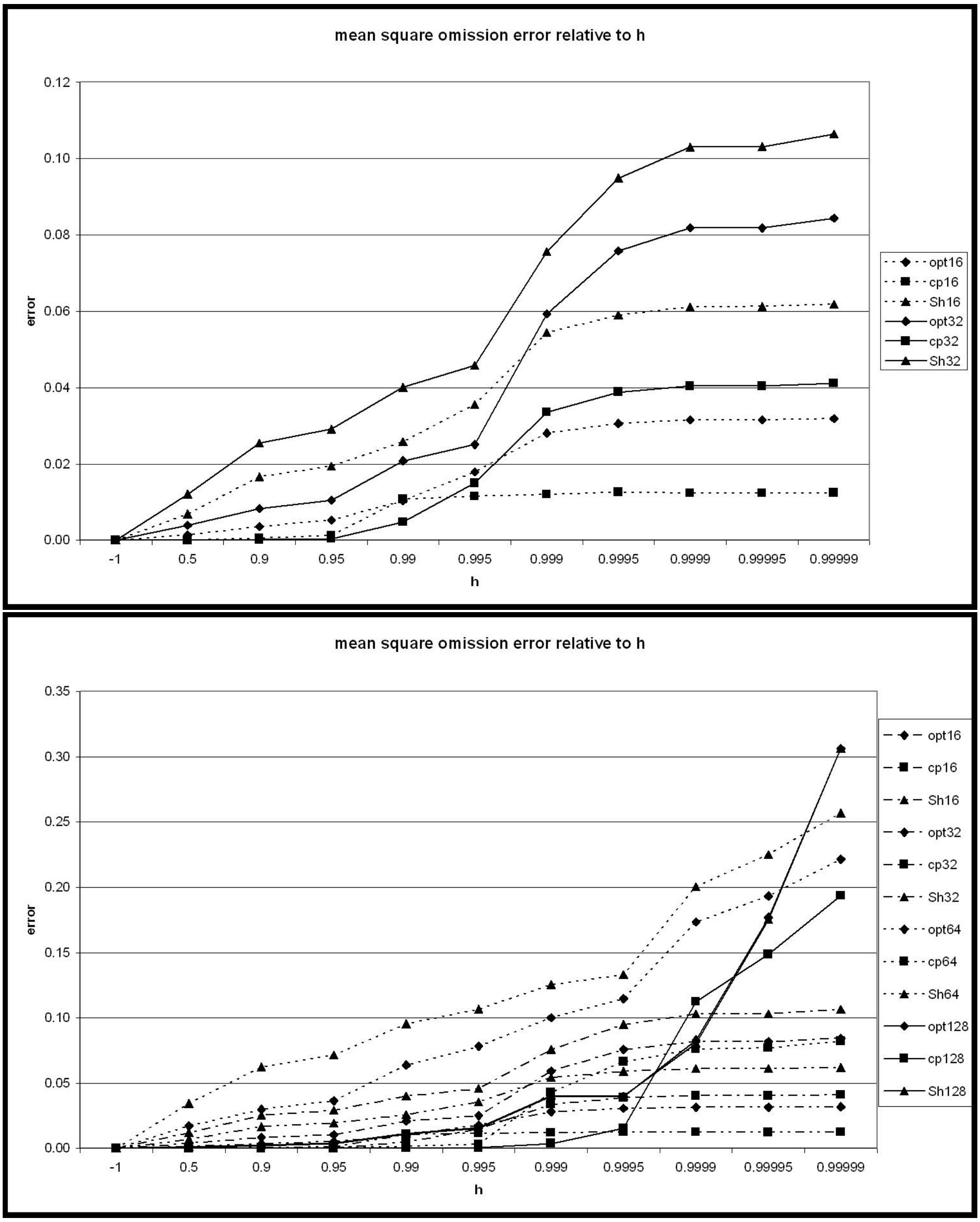


Figure 5: Root mean square omission error *convolution - truncated convolution* relative to h, note that $h = -1$ means that there is no spatial truncation in the quadrature rule.

satisfy the property that

$$\text{if } K_N^\wedge(n) \neq 0 \text{ then } K_M^\wedge(n) \neq 0 \text{ for all } M > N, \tag{10}$$

where here N corresponds to the degree of the kernel K_N. This implies that the corresponding scale spaces $V_N := \{K_N * F \mid F \in \mathrm{L}^2(\Omega)\}$ are nested. Hence, we have a multiresolution analysis. A mathematical proof that (10) holds in general still has to be found.
Moreover, in [12] approximate identities with exponential decay were derived for intervals. A combination with the spherical kernels obtained here can yield strongly localizing approximate identities on the 3–dimensional ball, which will find applications e.g. in the geosciences and in medical imaging.

8 Acknowledgments

The author gratefully acknowledges the financial support by the German Research Foundation (DFG), project MI 655/2-1 and the Stiftung Rheinland–Pfalz für Innovation, project 961-38 62 61 / 773.

A Calculation of the Matrix p

For calculating the matrix entries we have to numerically integrate $f := wP_nP_m$ on $[-1,1]$. For this purpose, the composite Newton–Cotes formula of order 4 with $4N+1 = 1,000,001$ grid points was chosen. We investigate here the error of this procedure. According to [23], p. 357 the error of integrating $F \in \mathrm{C}^{(6)}[a,b]$ with stepwidth $h = \frac{b-a}{4N}$ is estimated from above by $\frac{2(b-a)}{945}\, h^6 \|F^{(6)}\|_{\mathrm{C}[a,b]}$. We subdivide the integral here into two parts:

$$\begin{aligned} I_1 &:= \int_{1-4h}^{1} w(x)P_n(x)P_m(x)\,\mathrm{d}x, \\ I_2 &:= \int_{-1}^{1-4h} w(x)P_n(x)P_m(x)\,\mathrm{d}x. \end{aligned}$$

We discuss now the case $w(x) = \sqrt{1-x}$ and use the inequality $|P_k(t)| \leq 1$ which holds for all $k \in \mathbb{N}_0$ and all $t \in [-1,1]$. Then the quadrature error for the first integral can be estimated as follows:

$$\begin{aligned} &\left| I_1 - \frac{2h}{45}(7(f(1-4h)+f(1)) + 32(f(1-3h)+f(1-h)) + 12f(1-2h)) \right| \\ &\qquad \leq |I_1| + \frac{2h}{45} \cdot 90\, \|f\|_{\mathrm{C}[1-4h,1]} \\ &\qquad \leq (4+4)\, h\, \|f\|_{\mathrm{C}[1-4h,1]} \\ &\qquad \leq 8\, h\, \sqrt{4h} \cdot 1 \cdot 1 \\ &\qquad = 16\, h^{3/2} \\ &\qquad \approx 4.525 \cdot 10^{-8}. \end{aligned}$$

For the second integral the general error estimate above is useful. If $w(x) = \sqrt{1-x}$, the error can be estimated from above by

$$\frac{2\cdot(2-4h)}{945}\,h^6\left\|f^{(6)}\right\|_{\mathrm{C}[-1,1-4h]} \le \frac{4}{945}\,h^6\left\|f^{(6)}\right\|_{\mathrm{C}[-1,1-4h]}.$$

Without loss of generality we now assume that $n \ge m$. We calculate the sixth derivative of f as follows:

$$\begin{aligned} f^{(6)}(x) &= \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^6 \left(\sqrt{1-x}\,(P_n(x)P_m(x))\right) \\ &= \sum_{k=0}^{6}\binom{6}{k}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k (1-x)^{1/2}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{6-k}(P_nP_m)(x) \\ &= \sum_{k=0}^{6}\binom{6}{k}2^{-k}(-1)^{1-\delta_{0,k}}(2k-3)!!(1-x)^{(1-2k)/2}\,(P_nP_m)^{(6-k)}(x), \end{aligned}$$

where $k!! := k\cdot(k-2)!!$ for $k>1$ and $k!! := 1$ for $k \le 1$. Moreover, we find

$$(P_nP_m)^{(6-k)}(x) = \sum_{j=0}^{6-k}\binom{6-k}{j}P_n^{(j)}(x)P_m^{(6-k-j)}(x).$$

According to [24], we have

$$\left\|P_n^{(l)}\right\|_{\mathrm{C}[-1,1]} \le \left|P_n^{(l)}(1)\right| \le \left(\frac{1}{2}\right)^l \frac{(n+1)^{2l}}{l!}$$

for $l \in \mathbb{N}_0$. Hence, we can conclude that

$$\left\|(P_nP_m)^{(6-k)}\right\|_{\mathrm{C}[-1,1]} \le \sum_{j=0}^{6-k}\frac{(6-k)!}{(j!(6-k-j)!)^2}\,2^{-j-(6-k-j)}(n+1)^{2j+2(6-k-j)},$$

which implies that

$$\begin{aligned} \left\|f^{(6)}\right\|_{\mathrm{C}[-1,1-4h]} &\le \sqrt{2}\cdot\left(\sum_{j=0}^{6}\frac{6!}{(j!(6-j)!)^2}\right)2^{-6}(n+1)^{12} \\ &\quad+\sum_{k=1}^{6}\binom{6}{k}2^{-k}(2k-3)!!(4h)^{(1-2k)/2} \\ &\qquad\left(\sum_{j=0}^{6-k}\frac{(6-k)!}{(j!(6-k-j)!)^2}\right)2^{k-6}(n+1)^{12-2k}. \end{aligned}$$

The obtained error bounds for the quadrature of I_2 and the whole integral can be found in Table 2 for some selected values of n. Apparently, this accuracy is sufficient.

n	Error Bound for I_2	total error bound
64	$4.5139 \cdot 10^{-8}$	$9.0394 \cdot 10^{-8}$
128	$5.1388 \cdot 10^{-8}$	$9.6643 \cdot 10^{-8}$
256	$8.4336 \cdot 10^{-8}$	$1.2959 \cdot 10^{-7}$
512	$3.0013 \cdot 10^{-6}$	$3.0466 \cdot 10^{-6}$

Table 2: Error Bound for the quadrature of the second integral and the total integral

B Construction of the Spherical Rectangle $R_{\xi,h}$

Since the verification if an integration grid point is in the cap $C_{\xi,h} := \{\eta \in \Omega | \xi \cdot \eta \geq h\}$ is numerically expensive, we are looking for a covering spherical rectangle $R_{\xi,h} \supset C_{\xi,h}$ which is easy to determine. It should be given by bounds in polar coordinates $\underline{\varphi}_{\xi,h} \leq \varphi \leq \overline{\varphi}_{\xi,h}$, $\underline{\vartheta}_{\xi,h} \leq \vartheta \leq \overline{\vartheta}_{\xi,h}$. This allows the implementation of two nested for–loops (with the given bounds) per numerical integration in order to run through the remaining grid points on the sphere. For this reason, the total number of actually used grid points is a direct measure for the calculation time.
We use here the polar coordinate representation

$$\xi = \begin{pmatrix} \sin\vartheta_\xi \cos\varphi_\xi \\ \sin\vartheta_\xi \sin\varphi_\xi \\ \cos\vartheta_\xi \end{pmatrix}, \quad \eta = \begin{pmatrix} \sin\vartheta_\eta \cos\varphi_\eta \\ \sin\vartheta_\eta \sin\varphi_\eta \\ \cos\vartheta_\eta \end{pmatrix};$$

$\vartheta_\xi, \vartheta_\eta \in [0,\pi]$; $\varphi_\xi \in [0, 2\pi[$. The values of φ_η will be chosen below appropriately in the range $[-\pi, 3\pi]$. The rectangle we are interested in should, thus, be representable by $R_{\xi,h} = \{\eta \in \Omega | \underline{\varphi}_{\xi,h} \leq \varphi_\eta \leq \overline{\varphi}_{\xi,h}, \underline{\vartheta}_{\xi,h} \leq \vartheta_\eta \leq \overline{\vartheta}_{\xi,h}\}$. With this nomenclature the criterion $\xi \cdot \eta \geq h$ is equivalent to

$$\sin\vartheta_\xi \sin\vartheta_\eta \cos\varphi_\xi \cos\varphi_\eta + \sin\vartheta_\xi \sin\vartheta_\eta \sin\varphi_\xi \sin\varphi_\eta + \cos\vartheta_\xi \cos\vartheta_\eta \geq h. \tag{11}$$

If the longitudes of ξ and η coincide, i.e. $\varphi_\xi = \varphi_\eta$, then this inequality can be simplified to

$$\begin{aligned} \sin\vartheta_\xi \sin\vartheta_\eta + \cos\vartheta_\xi \cos\vartheta_\eta &\geq h \\ \Leftrightarrow \cos(\vartheta_\xi - \vartheta_\eta) &\geq h \\ \Leftrightarrow |\vartheta_\xi - \vartheta_\eta| &\leq \arccos h. \end{aligned}$$

Obviously, at different longitudes φ_η the cap covers a smaller range of latitudes, which leads us to the following bounds

$$\underline{\vartheta}_{\xi,h} := \max(0, \vartheta_\xi - \arccos h) \leq \vartheta_\eta \leq \min(\pi, \vartheta_\xi + \arccos h) =: \overline{\vartheta}_{\xi,h}$$

for the latitudes.
We now distinguish whether ξ is located at the northern or at the southern hemisphere, where it does not make any difference among which part the equator is. In the northern case, we have

a look at the lower (i.e. northwards located) bounding latitude $\underline{\vartheta}_{\xi,h}$ and consider $\tilde{\xi}$ with $\varphi_{\tilde{\xi}} = \varphi_\xi$ and $\vartheta_{\tilde{\xi}} = \underline{\vartheta}_{\xi,h}$. For $\vartheta_\eta = \underline{\vartheta}_{\xi,h}$, Inequality (11) becomes (with $\tilde{\xi}$ instead of ξ):

$$\begin{aligned} \sin^2 \underline{\vartheta}_{\xi,h} \left(\cos\varphi_\xi \cos\varphi_\eta + \sin\varphi_\xi \sin\varphi_\eta\right) + \cos^2 \underline{\vartheta}_{\xi,h} &\geq h \\ \Leftrightarrow \sin^2 \underline{\vartheta}_{\xi,h} \cos\left(\varphi_\xi - \varphi_\eta\right) + \cos^2 \underline{\vartheta}_{\xi,h} &\geq h. \end{aligned}$$

Formally, we shifted the cap by the latitude $\vartheta_\xi - \underline{\vartheta}_{\xi,h}$ northwards to the new center $\tilde{\xi}$. The idea is that the meridians get closer and closer towards the North pole. If we install a cap of equal radius at the upper bounding circle of latitude, then the range of longitudes covered by this cap at this circle of latitude must also cover the original cap. This also means, that the spherical rectangle which we construct here will have bounding meridians which need not (and will usually not) touch the boundary of the cap. However, we obtain a practical restriction of the sphere for our purposes.

We observe now a special case which occurs if

$$\sin^2 \underline{\vartheta}_{\xi,h} \cdot (-1) + \cos^2 \underline{\vartheta}_{\xi,h} \geq h,$$

which is equivalent to $\cos(2\underline{\vartheta}_{\xi,h}) \geq h$. If this holds true, then every choice of φ_η is admissible. This occurs if $\cos^2 \underline{\vartheta}_{\xi,h}$ is close to 1, i.e. if ξ is close to a pole (which is here the North pole). In this case, we set

$$\underline{\varphi}_{\xi,h} := \varphi_\xi - \pi \leq \varphi_\eta \leq \varphi_\xi + \pi =: \overline{\varphi}_{\xi,h}.$$

Otherwise, we get

$$|\varphi_\xi - \varphi_\eta| \leq \arccos\left(\frac{h - \cos^2 \underline{\vartheta}_{\xi,h}}{\sin^2 \underline{\vartheta}_{\xi,h}}\right),$$

i.e.

$$\underline{\varphi}_{\xi,h} := \varphi_\xi - \arccos\left(\frac{h - \cos^2 \underline{\vartheta}_{\xi,h}}{\sin^2 \underline{\vartheta}_{\xi,h}}\right) \leq \varphi_\eta \leq \varphi_\xi + \arccos\left(\frac{h - \cos^2 \underline{\vartheta}_{\xi,h}}{\sin^2 \underline{\vartheta}_{\xi,h}}\right) =: \overline{\varphi}_{\xi,h}.$$

The considerations for the southern hemisphere are analogous with $\overline{\vartheta}_{\xi,h}$ instead of $\underline{\vartheta}_{\xi,h}$.

References

[1] W. Freeden, T. Gervens, and M. Schreiner (1998). *Constructive Approximation on the Sphere - With Applications to Geomathematics*. Oxford University Press, Oxford.

[2] W. Freeden and K. Hesse (2002). On the multiscale solution of satellite problems by use of locally supported kernel functions corresponding to equidistributed data on spherical orbits, *Studia Scientiarum Mathematicarum Hungarica*. 39:37-74.

[3] V.A. Menegatto and A.C. Piantella (2005). Approximation on the sphere by weighted Fourier expansions. *Journal of Applied Mathematics*. 2005:321-339.

[4] J. Keiner, S. Kunis, and D. Potts (2006). Fast summation of radial functions on the sphere. *Computing*, 78:1-15.

[5] J. Keiner, S. Kunis, and D. Potts (2007). Efficient reconstruction of functions on the sphere from scattered data. *J. Fourier Anal. Appl.* 13:435-458.

[6] N. Laín Fernández (2003). *Polynomial Bases on the Sphere.* PhD Thesis, University of Lübeck, Department of Mathematics, Logos–Verlag, Berlin.

[7] N. Laín Fernández (2005). Localized polynomial basis on the sphere. *Electronic Transactions on Numerical Analysis.* 19:84-93.

[8] N. Laín Fernández (2007). Optimally space–localized band–limited wavelets on S^{q-1}. *Journal of Computational and Applied Mathematics.* 199:68-79.

[9] N. Laín Fernández and J. Prestin (2002). Localization of the spherical Gauss–Weierstrass kernel. In, *Constructive Function Theory*, (B. Bojanov, ed.), Varna, DARBA, Sofia.

[10] H.N. Mhaskar and J. Prestin (2005). On local smoothness classes of periodic functions. *Journal of Fourier Analysis and Applications.* 11:353-373.

[11] H.N. Mhaskar and J. Prestin (2005). Polynomial frames: a fast tour. In, *Approximation Theory XI: Gatlinburg 2004 (Brentwood)* (C.K. Chui, M. Neamtu and L. Shumaker, eds.), Nashboro Press, pp. 287-318.

[12] H.N. Mhaskar and J. Prestin (2006). Exponentially localized polynomial frames on compact subset of the real line and the euclidean sphere. Preprint.

[13] F.J. Simons, F.A. Dahlen, and M.A. Wieczorek (2006). Spatiospectral localization on a sphere. *SIAM Review.* 48:504-536.

[14] F.J. Narcowich and J.D. Ward (1996). Nonstationary wavelets on the m–sphere for scattered data. *Applied and Computational Harmonic Analysis (ACHA).* 3:324-336.

[15] G. Szegö (1967). *Orthogonal Polynomials.* American Mathematical Society Colloquium Publications, Vol. XXIII, Providence, Rhode Island.

[16] W. Freeden (1999). *Multiscale Modelling of Spaceborne Geodata.* B.G. Teubner Verlag, Stuttgart, Leipzig.

[17] V. Michel (2002). *A Multiscale Approximation for Operator Equations in Separable Hilbert Spaces — Case Study: Reconstruction and Description of the Earth's Interior.* Habilitation Thesis, Shaker Verlag, Aachen.

[18] C. Müller (1966). *Spherical Harmonics.* Lecture Notes in Mathematics, Vol. 17, Springer, Berlin, Heidelberg.

[19] J.R. Driscoll and R.M. Healy (1994). Computing Fourier transforms and convolutions on the 2-sphere. *Adv. Appl. Math.* 15:202-250.

[20] I.H. Sloan and R.S. Womersley (2004). Extremal systems of points and numerical integration on the sphere. *Advances in Computational Mathematics*. 21:107-125.

[21] A.P. Prudnikov, Yu.A. Brychkov, and O.I. Marichev (1986). *Integrals and Series, Volume 2: Special Functions*. Gordon and Breach Science Publishers, New York.

[22] P. Deuflhard (1975). On algorithms for the summation of certain special functions. *Computing*. 17:37-48.

[23] H.R. Schwarz (1993). *Numerische Mathematik*. B.G. Teubner, Stuttgart.

[24] C. Müller (1952). Über die ganzen Lösungen der Wellengleichung. *Mathematische Annalen*. 124:235-264.

Siegen Preprints on Geomathematics

The preprint series “Siegen Preprints on Geomathematics” was established in 2010. See www.geomathematics-siegen.de for details and a contact address. At present, the following preprints are available:

1. P. Berkel, D. Fischer, V. Michel: *Spline multiresolution and numerical results for joint gravitation and normal mode inversion with an outlook on sparse regularisation*, 2010.
2. M. Akram, V. Michel: *Regularisation of the Helmholtz decomposition and its application to geomagnetic field modelling*, 2010.
3. V. Michel: *Optimally Localized Approximate Identities on the 2-Sphere*, 2010.

Geomathematics Group Siegen
Prof. Dr. Volker Michel

Contact at:
Geomathematics Group
Department of Mathematics
University of Siegen
Walter-Flex-Str. 3
57068 Siegen
www.geomathematics-siegen.de