Handbook of Mathematical Geodesy
Geosystems Mathematics, 753–819

Spherical Harmonics Based Special Function Systems and Constructive Approximation Methods

Willi Freeden, Volker Michel, and Frederik J. Simons

Abstract. Special function systems are reviewed that reflect particular properties of the Legendre polynomials, such as spherical harmonics, zonal kernels, and Slepian functions. The uncertainty principle is the key to their classification with respect to their localization in space and frequency/momentum. Methods of constructive approximation are outlined such as spherical harmonic and Slepian expansions, spherical spline and wavelet concepts. Regularized Functional Matching Pursuit is described as an approximation technique of combining heterogeneous systems of trial functions to a kind of a 'best basis'.

Keywords. Spherical harmonics procedures, Slepian, spline and wavelet methods, regularized functional matching pursuit.

1. Introduction

Up until the present time, modeling geoscientific data is often performed on a global scale by orthogonal expansions in terms of spherical harmonics. However, in many aspects global spherical harmonic modeling cannot keep pace with the prospects and the expectations of the 'Earth system sciences'. In particular, there is an increasing need for high-precision modeling on local areas. As we shall discuss, Slepian functions are important tools for this purpose. For their part, zonal kernel functions – in the jargon of constructive approximation: radial basis functions – have become more and more important because of their space localizing properties (even in the vectorial and tensorial context).

The addition theorem for spherical harmonics enables us to express all types of zonal kernel functions in terms of a one-dimensional function, the *Legendre polynomial*. Weighted additive clustering of Legendre polynomials generates specific classes of space localizing zonal kernel functions, i.e., Legendre series expansions,

ready for approximation within the scalar, vectorial, and tensorial framework. The closer the Legendre series expansion is to the Dirac kernel, the more localized is the zonal kernel in space, and the more economical is its role in (spatial) local computation. In addition, the Funk–Hecke formula provides the natural tool for establishing convolutions of spherical fields against zonal kernels. Consequently, by specifying Dirac families, i.e., sequences of zonal functions tending to the Dirac kernel, (space-localized) filtered versions of (square-integrable) spherical fields are obtainable by convolution, leading to 'zooming-in', multiscale approximations. Altogether, the Legendre polynomial is the keystone of any work about special functions in the mathematical geosciences. It enables the transition from spherical harmonics via zonal kernels up to the Dirac kernel. The Funk–Hecke formula and its consequences for spherical convolutions open new methodological perspectives for global as well as local approximation in scalar, vectorial and tensorial applications.

In this paper, we discuss selected systems of trial functions on the sphere with a brief excursion to basis functions on the ball. These spherical function systems are investigated with respect to their localization in space and frequency/momemtum. Moreover, we briefly summarize a method of finding a best basis by Regularized Functional Matching Pursuit.

2. Special function systems on sphere and ball

Because of the nearly spherical shape of the Earth, spherical functions and concepts play an essential part in all of the geosciences. By a spherical variant of the Weierstraß theorem, spherical polynomials, the spherical harmonics, approximate continuous functions with respect to different topologies.

2.1. Spherical harmonics

Spherical harmonics are the analogues of trigonometric functions for Fourier expansion theory on the sphere. They were introduced to study gravitational theory [61, 62]. Early publications on the theory of spherical harmonics in their original physical interpretation as 'multipoles' are by Clebsch [16], Sylvester [93], Heine [54], Neumann [77], and Maxwell [66]. Global geomagnetic data and basic spherical harmonic expansions became available in the mid 1800s [50]. Today, the use of spherical harmonics in all geosciences is well established, particularly for the representation of scalar potentials. Reference models for the Earth's gravitational or magnetic fields are distributed as tables of coefficients for the spherical harmonic expansion of their potentials. In this approach, each spherical harmonic is a polynomial 'ansatz-function', corresponding to one 'degree', or in the jargon of signal processing, to exactly one 'frequency'. Thus, orthogonal (Fourier) expansion in terms of spherical harmonics amounts to the superposition of summands with an oscillating character determined by the degree of the Legendre polynomial (see Table 1). The more spherical harmonics are involved in the expansion, the more the oscillations grow in number, but the smaller are their amplitudes.

Weierstraß approximation theorem ↓ geophysical constraint of harmonicity
spherical harmonics $Y_{n,j}$ as restrictions of homogeneous harmonic polynomials $H_{n,j}$ to the unit sphere $\Omega \subset \mathbb{R}^3$
orthonormality and invariance ↓ addition theorem
one-dimensional Legendre polynomial P_n: $P_n(\xi \cdot \eta) = \frac{4\pi}{2n+1} \sum_{j=1}^{2n+1} Y_{n,j}(\xi) Y_{n,j}(\eta), \quad \xi, \eta \in \Omega$
convolution by the Legendre kernel ↓ Funk–Hecke formula
Legendre transform of F: $(P_n * F)(\xi) = \frac{2n+1}{4\pi} \int_\Omega P_n(\xi \cdot \eta) F(\eta) \, \mathrm{d}\omega(\eta), \quad \xi \in \Omega$
superposition over frequencies ↓ orthogonal series expansion
Fourier series of $F \in \mathrm{L}^2(\Omega)$: $F(\xi) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \int_\Omega P_n(\xi \cdot \eta) F(\eta) \, \mathrm{d}\omega(\eta), \quad \xi \in \Omega$

TABLE 1. Fourier expansion of square-integrable scalar functions on the sphere.

The geosciences deal with the space $\mathrm{L}^2(\Omega)$ of square-integrable functions on the unit sphere Ω. The quantity

$$\|F\|_{\mathrm{L}^2(\Omega)} = \left(\int_\Omega (F(\xi))^2 \, \mathrm{d}\omega(\xi) \right)^{1/2} \tag{1}$$

may be understood as the energy of the 'signal' $F \in \mathrm{L}^2(\Omega)$. The appropriate representation of a finite-energy signal in terms of a countable Hilbert basis is one of the most centrally important problems in the mathematical geosciences. The spherical harmonics form a Hilbert basis in $\mathrm{L}^2(\Omega)$. Suitable systems of spherical harmonics $\{Y_{n,k}\}_{n=0,1,\ldots;\ k=1,\ldots,2n+1}$ are often defined by the restriction of homogeneous

harmonic polynomials to the sphere. The polynomial structure has tremendous advantages. First, spherical harmonics of different degrees are orthogonal. Second, the space Harm_n of spherical harmonics of degree n is finite-dimensional: $\dim(\mathrm{Harm}_n) = 2n+1$. Therefore, the basis property of $\{Y_{n,k}\}_{n=0,1,\ldots;\ k=1,\ldots,2n+1}$ is equivalently characterized by the completion of the direct sum $\bigoplus_{n=0}^{\infty} \mathrm{Harm}_n$, i.e.:

$$\mathrm{L}^2(\Omega) = \overline{\bigoplus_{n=0}^{\infty} \mathrm{Harm}_n}^{\|\cdot\|_{\mathrm{L}^2(\Omega)}} . \tag{2}$$

This is the canonical reason why spherical harmonic (multipole) expansions underlie the classical approaches to geopotentials.

Fourier transform. More explicitly, any 'signal' $F \in \mathrm{L}^2(\Omega)$ can be split into 'orthogonal contributions' involving the Fourier transforms $F^\wedge(n,k)$ defined by

$$F^\wedge(n,k) = \int_\Omega F(\xi) Y_{n,k}(\xi)\, \mathrm{d}\omega(\xi), \tag{3}$$

in terms of $\mathrm{L}^2(\Omega)$-orthonormal spherical harmonics $\{Y_{n,k}\}_{\substack{n=0,1,\ldots \\ k=1,\ldots,2n+1}}$. Parseval's identity identifies the spatial energy of a signal with the spectral energy, decomposed orthogonally into single frequency contributions

$$\|F\|^2_{\mathrm{L}^2(\Omega)} = \langle F, F\rangle_{\mathrm{L}^2(\Omega)} = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \left(F^\wedge(n,k)\right)^2 .$$

This explains why the (global) geosciences work more often with the '*amplitude spectrum*' $\{F^\wedge(n,k)\}_{\substack{n=0,1,\ldots \\ k=1,\ldots,2n+1}}$ than with the 'original signal' $F \in \mathrm{L}^2(\Omega)$.

Inverse Fourier transform. The 'inverse Fourier transform'

$$F = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} F^\wedge(n,k) Y_{n,k} \tag{4}$$

allows the geoscientist to think of the function (signal) F as a sum of 'wave functions' $Y_{n,k}$ corresponding to different frequencies. One can think of measurements as operating on an 'input signal' F to produce an output signal $G = \Lambda F$, where Λ is an operator acting on $\mathrm{L}^2(\Omega)$. Fortunately, large portions of interest can be well approximated by linear rotation-invariant pseudodifferential operators (see, e.g., [33, 47, 92]). If Λ is such an operator on $\mathrm{L}^2(\Omega)$, this means that

$$\Lambda Y_{n,k} = \Lambda^\wedge(n) Y_{n,k}, n = 0, 1, \ldots; k = 1, \ldots, 2n+1, \tag{5}$$

where the 'symbol' $\{\Lambda^\wedge(n)\}_{n\in\mathbb{N}_0}$ is a sequence of real values (independent of the order k). Thus, we have the fundamental fact that the spherical harmonics are the eigenfunctions of the operator Λ. Different pseudodifferential operators Λ are characterized by their eigenvalues $\Lambda^\wedge(n)$. All eigenvalues $\{\Lambda^\wedge(n)\}_{n\in\mathbb{N}_0}$ are collected in the so-called symbol of Λ. The 'amplitude spectrum' $\{G^\wedge(n,k)\}$ of the response of Λ is described in terms of the amplitude spectrum of functions (signals) by a simple multiplication by the 'transfer' $\Lambda^\wedge(n)$.

Bandlimited/spacelimited functions. Physical devices do not transmit spherical harmonics of arbitrarily high frequency without severe attenuation. The 'transfer' $\Lambda^\wedge(n)$ usually tends to zero with increasing n. It follows that the amplitude spectra of the responses (observations) to functions (signals) of finite energy are also negligibly small beyond some finite frequency. Thus, both because of the frequency limiting nature of the devices used, and because of the nature of the 'transmitted signals', the geoscientist is soon led to consider bandlimited functions. These are the functions $F \in \mathrm{L}^2(\Omega)$ whose 'amplitude spectra' vanish for all $n > N$ ($N \in \mathbb{N}$ fixed). In other words, each bandlimited function $F \in \mathrm{L}^2(\Omega)$ can be written as a finite Fourier series. So, any function F of the form $F = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} F^\wedge(n,k) Y_{n,k}$ is said to be *bandlimited with the band* N, if $F^\wedge(N,k) \neq 0$ for at least one k. In analogous manner, $F \in \mathrm{L}^2(\Omega)$ is said to be *locally supported* (*spacelimited*) *with spacewidth* ρ around an axis $\eta \in \Omega$, if for some $\rho \in (-1,1)$ the function F vanishes on the set of all $\xi \in \Omega$ with $-1 \leq \xi \cdot \eta \leq \rho$ (where ρ is the largest number for which this is the case). Bandlimited functions are infinitely often differentiable everywhere. Moreover, it is clear that any bandlimited function F is an analytic function. From the analyticity, it follows immediately that a non-trivial bandlimited function cannot vanish on any (non-degenerate) subset of Ω. The only function that is both bandlimited and spacelimited is the zero function.

In addition to bandlimited but non-spacelimited functions, numerical analysis would like to deal with spacelimited functions. However, as we have seen, such a function (signal) of finite (space) support cannot be bandlimited, it must contain spherical harmonics of arbitrarily large frequencies. Thus, there is a dilemma of seeking functions that are somehow concentrated in both space and frequency (more accurately, angular momentum domain). There is a way of mathematically expressing the impossibility of simultaneous confinement of a function to space and angular momentum, namely the *uncertainty principle*.

2.2. Zonal kernel functions

To understand the *transition from the theory of spherical harmonics through zonal kernel functions to the Dirac kernel*, we have to realize the relative advantages of the classical Fourier expansion method by means of spherical harmonics, and this not only in the frequency domain, but also in the space domain. It is characteristic for Fourier techniques that the spherical harmonics as polynomial trial functions admit no localization in space domain, while in the frequency domain (or: angular momentum domain), they always correspond to exactly one degree, i.e., frequency, and therefore, are said to show ideal frequency localization. Because of the ideal frequency localization and the simultaneous absence of space localization, in fact, local changes of fields (signals) in the space domain affect the whole table of orthogonal (Fourier) coefficients. This, in turn, causes global changes of the corresponding (truncated) Fourier series in the space domain. Nevertheless, ideal frequency localization is often helpful for meaningful physical interpretations by relating the different observables of a geopotential to each other at a fixed frequency.

Taking these aspects on spherical harmonic modeling by Fourier series into account, trial functions which simultaneously show ideal frequency localization as well as ideal space localization would be a desirable choice. In fact, such an ideal system of trial functions would admit models of highest spatial resolution which were expressible in terms of single frequencies. However, from the uncertainty principle – the connection between space and frequency localization – we will see that both characteristics are mutually exclusive.

In conclusion, Fourier expansion methods are well suited to resolve low and medium frequency phenomena, i.e., the 'trends' of a signal, while their application to obtain high resolution in global or local models is critical. This difficulty is also well known to theoretical physics, e.g., when describing monochromatic electromagnetic waves or considering the quantum-mechanical treatment of free particles. There, plane waves with fixed frequencies (ideal frequency localization, no space localization) are the solutions of the corresponding differential equations, but they do certainly not reflect the physical reality. As a remedy, plane waves of different frequencies are superposed into 'wave-packages' that gain a certain amount of space localization, while losing their ideal spectral localization. In a similar way, a suitable superposition of polynomial functions leads to so-called zonal kernel functions, in particular to kernel functions with a reduced frequency, but increased space localization.

More concretely, any kernel function $K : \Omega \times \Omega \to \mathbb{R}$ that is characterized by the property that there exists a function $\tilde{K} : [0, 2] \to \mathbb{R}$ such that

$$K(\xi, \eta) = \tilde{K}(|\xi - \eta|) = \tilde{K}\left(\sqrt{2 - 2\xi \cdot \eta}\right) = \hat{K}(\xi \cdot \eta), \quad \xi, \eta \in \Omega, \tag{6}$$

is called a (spherical) *radial basis function* (at least in the theory of constructive approximation).

Zonal kernels. The application of a rotation (i.e., a 3×3 'orthogonal' matrix $\mathbf{t}$ with $\mathbf{t}^{\mathrm{T}} = \mathbf{t}^{-1}$) leads to $K(\mathbf{t}\xi, \mathbf{t}\eta) = \hat{K}((\mathbf{t}\xi) \cdot (\mathbf{t}\eta)) = \hat{K}(\xi \cdot (\mathbf{t}^{\mathrm{T}}\mathbf{t}\eta)) = \hat{K}(\xi \cdot \eta) = K(\xi, \eta)$. In particular, a rotation around the axis $\xi \in \Omega$ (i.e., $\mathbf{t}\xi = \xi$) yields $K(\xi, \eta) = K(\xi, \mathbf{t}\eta)$ for all $\eta \in \Omega$. Hence, $K(\xi, \cdot)$ possesses a rotational symmetry with respect to the axis ξ. In the theory of special functions of mathematical physics, a kernel $\hat{K} : \Omega \times \Omega \to \mathbb{R}$ satisfying $\hat{K}(\xi \cdot \eta) = \hat{K}(\mathbf{t}\xi \cdot \mathbf{t}\eta)$, $\xi, \eta \in \Omega$, for all orthogonal transformations $\mathbf{t}$ is known as a *zonal kernel function*. To highlight the reducibility of $\hat{K}$ to a function defined on the interval $[-1, 1]$, the notation $(\xi, \eta) \mapsto \hat{K}(\xi \cdot \eta)$, $(\xi, \eta) \in \Omega \times \Omega$, is used throughout this chapter (see also (6)).

From the theory of spherical harmonics we get a representation of any $\mathrm{L}^2(\Omega)$-zonal kernel function K in terms of a Legendre expansion

$$K(\xi \cdot) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} K^{\wedge}(n) P_n(\xi \cdot) \tag{7}$$

(in the $\|\cdot\|_{\mathrm{L}^2(\Omega)}$-sense), where the sequence $\{K^\wedge(n)\}_{n\in\mathbb{N}_0}$ given by

$$K^\wedge(n) = 2\pi \int_{-1}^{1} K(t) P_n(t)\,\mathrm{d}t \tag{8}$$

is called the *Legendre symbol* of the zonal kernel $K(\xi\cdot)$. A simple but extreme example (with optimal frequency localization and no space localization) is the Legendre kernel where $K^\wedge(n) = 1$ for one particular n and $K^\wedge(m) = 0$ for $m \neq n$, i.e., the Legendre kernel is given by

$$\Omega \times \Omega \ni (\xi, \eta) \mapsto \frac{2n+1}{4\pi} P_n(\xi \cdot \eta).$$

In other words, additive clustering of weighted Legendre kernels generates zonal kernel functions. It is of importance to distinguish bandlimited kernels (i.e., $K^\wedge(n) = 0$ for all $n \geq N$) and non-bandlimited ones, for which infinitely many numbers $K^\wedge(n)$ do not vanish. Non-bandlimited kernels show a much stronger space localization than their bandlimited counterparts. Empirically, if $K^\wedge(n) \approx K^\wedge(n+1) \approx 1$ for many successive large integers n, then the support of the series (7) in the space domain is small, i.e., the kernel is spacelimited (i.e., in the jargon of approximation theory 'locally supported'). This leads to the other extremal kernel (in contrast to the Legendre kernel) which is the Dirac kernel with optimal space localization but no frequency localization and $K^\wedge(n) = 1$ for all n, where, however, the Dirac kernel does not exist as a classical function in the mathematical sense. Nevertheless, it is well known that, if we have a family of kernels $\{K_J\}_{J=0,1,\ldots}$ where $\lim_{J\to\infty} K_J^\wedge(n) = 1$ for each n and an additional (technical) condition holds, then $K_J * F$ tends to F in the sense of $\mathrm{L}^2(\Omega)$ for all $F \in \mathrm{L}^2(\Omega)$.

Assuming $\lim_{n\to\infty} K^\wedge(n) = 0$, necessary to get a 'proper' function, the slower the sequence $\{K^\wedge(n)\}_{n=0,1,\ldots}$ converges to zero, the lower the frequency localization, and the higher the space localization. A unified scheme is found in Table 2. Zonal kernel function theory relies on the following principles:

(i) Weighted Legendre kernels are the summands of zonal kernel functions.
(ii) The Legendre kernel is ideally localized in frequency. The Dirac kernel is ideally localized in space.
(iii) The only frequency- *and* spacelimited zonal kernel is the zero function.

Legendre kernels	zonal kernels general case		Dirac kernel
	bandlimited	spacelimited	

TABLE 2. From Legendre kernels via zonal kernels to the Dirac kernel

2.3. Slepian functions

As we have seen, $\|F\|^2_{\mathrm{L}^2(\Omega)}$ is the energy of a certain function $F \in \mathrm{L}^2(\Omega)$. Suppose now that there is a particular region $C \subset \Omega$, and let us define the 'local' energy of that function as $\|F\|^2_{\mathrm{L}^2(C)}$. Functions F that are *band*limited,

$$F = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} F^{\wedge}(n,k) Y_{n,k} \tag{9}$$

cannot also be *space*limited, but they can be space*concentrated*.

Bandlimited/spaceconcentrated Slepian functions. By maximizing the spatial energy ratio

$$\lambda_C(F) = \frac{\|F\|^2_{\mathrm{L}^2(C)}}{\|F\|^2_{\mathrm{L}^2(\Omega)}}, \quad 0 < \lambda_C(F) < 1, \tag{10}$$

we obtain bandlimited spherical 'Slepian functions' [56, 67, 87], named in analogy with the prolate spheroidal wave functions of Slepian [90]. They are not, in general, zonal functions.

The Fourier coefficients of the Slepian functions are the $(N+1)^2$ orthogonal eigenvectors of the symmetric concentration matrix whose elements are the limited-domain inner-product terms $\langle Y_{m,j}, Y_{n,k} \rangle_{\mathrm{L}^2(C)}$, $0 \le m, n \le N$, i.e.,

$$\sum_{n=0}^{N} \sum_{k=1}^{2n+1} \langle Y_{m,j}, Y_{n,k} \rangle_{\mathrm{L}^2(C)} F^{\wedge}(n,k) = \lambda_C(F) F^{\wedge}(m,j). \tag{11}$$

We will give their associated eigenvalues superscripted labels and rank them in decreasing order of concentration, $1 > \lambda_C^{(1)}(F) \ge \lambda_C^{(\alpha)}(F) \ge \lambda_C^{((N+1)^2)}(F) > 0$. The bandlimited Slepian functions can alternatively be obtained by solving a Fredholm integral equation with a 'Shannon' concentration kernel:

$$\int_C \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta) F(\eta) \, \mathrm{d}\omega(\eta) = \lambda_C(F) F(\xi), \quad \xi \in \Omega. \tag{12}$$

Spacelimited/bandconcentrated Slepian functions. We can define space*limited* Slepian functions which are band*concentrated*. They are obtained by the restriction of the bandlimited Slepian functions F to the region of interest C, or, equivalently, their Fourier coefficients are

$$\sum_{n=0}^{N} \sum_{k=1}^{2n+1} \langle Y_{m,j}, Y_{n,k} \rangle_{\mathrm{L}^2(C)} F^{\wedge}(n,k), \tag{13}$$

extending the $F^{\wedge}(n,k)$ to *all* degrees $m = 0, 1, \ldots, \infty$ and order indices $j = 1, \ldots, 2m+1$.

A central concept is the *effective* dimension of functions that are 'essentially' space- and bandlimited. The *Shannon number* is the trace of the concentration operators in (11)–(12), given by (using the addition theoremin 14b–14c),

$$N_C = \sum_{\alpha=1}^{(N+1)^2} \lambda_C^{(\alpha)}(F) \tag{14a}$$

$$= \sum_{n=0}^{N} \sum_{k=1}^{2n+1} \langle Y_{n,k}, Y_{n,k} \rangle_{\mathrm{L}^2(C)} \tag{14b}$$

$$= \int_C \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\eta \cdot \eta)\, \mathrm{d}\omega(\eta) \tag{14c}$$

$$= (N+1)^2 \frac{\int_C \mathrm{d}\omega(\eta)}{4\pi}. \tag{14d}$$

The eigenvalue spectrum $\lambda_C^{(\alpha)}(F)$ has a characteristic step-like shape, with the property $\sum_{\alpha=1}^{(N+1)^2} \lambda_C^{(\alpha)}(F) \approx \sum_{\alpha=1}^{N_C} \lambda_C^{(\alpha)}(F)$ revealing that N_C will be close to the number of Slepian functions that usefully contribute to the approximation of arbitrary target functions on domains $C \subset \Omega$.

While computation can be carried out via either (11) or (12), when the region of interest C is a spherical cap (one whose boundary ∂C is a circle and whose half-opening angle is Θ), the integral equation (12) commutes with a Sturm–Liouville differential equation whose spectral-domain representation has an extremely simple analytical form, rendering the computation of Slepian functions of domains essentially trivial [51]. In that case, the Slepian functions degenerate to being the solutions of fixed-order (j) versions of equation (11), with a partial Shannon number given in terms of products of the associated Legendre functions and their derivatives (primed), namely

$$N_j = \frac{(N-j+1)!}{2(N+j)!} \int_{\cos\Theta}^{1} \left[P'_{N+1,j}(t) P_{N,j}(t) - P'_{N,j}(t) P_{N+1,j}(t) \right] \mathrm{d}t. \tag{15}$$

Only on circularly symmetric domains and when the spherical-harmonic order $j = 0$ are the Slepian functions zonal, and in that case, the fixed-order partial Shannon number is well approximated by Wieczorek and Simons [104] as

$$N_0 \approx 2\frac{\sqrt{N_C}}{\pi} \approx (N+1)\frac{\Theta}{\pi}. \tag{16}$$

Figure 1 shows examples of spherical-cap Slepian functions, their power spectra, and their eigenvalue spectra.

When the concentration domain is a spherical cap, the best-concentrated (highest-eigenvalue) bandlimited Slepian function is a zonal function that is close to optimally localized under the uncertainty principle (see Section 3). All the lower-eigenvalue zonal Slepian functions, and finally, all the non-zonal Slepian functions, together form a complete orthonormal basis for the space of functions on the unit

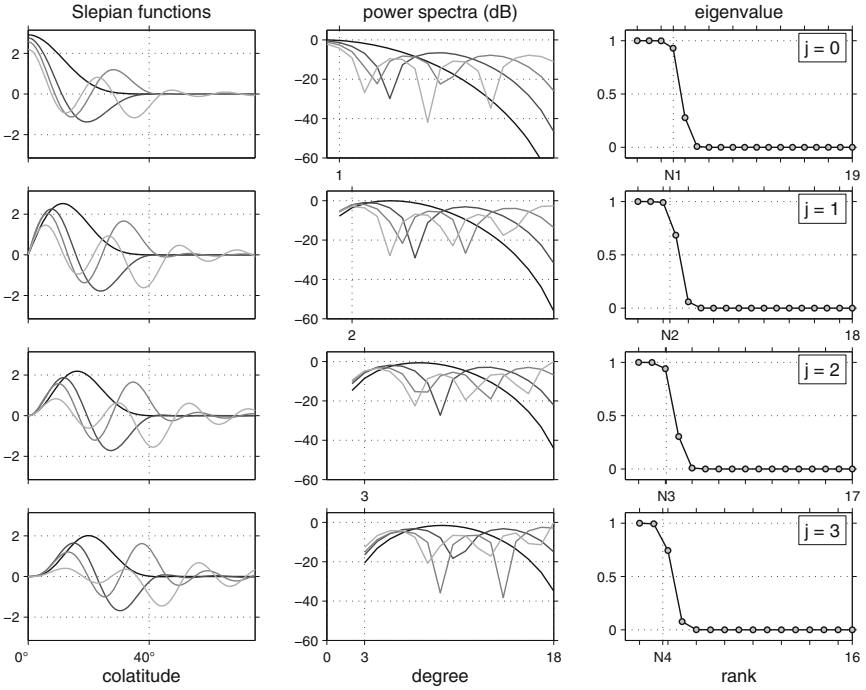


FIGURE 1. The four best-concentrated (in decreasing gray shading) fixed-order (top to bottom, $j = 0, 1, 2, 3$) Slepian functions and their power spectra, for a common bandwidth $N = 18$, with the domain C a 40° spherical cap. Also shown are the complete eigenvalue spectra with the fixed-order (partial) Shannon numbers indicated. Compare to [87] (their Figs. 5.1, 5.2 and 5.3.).

sphere Ω that are bandlimited to N. The partial Shannon numbers N_j sum to the full Shannon number N_C via

$$N_C = \sum_{j=0}^{2N+1} N_j. \tag{17}$$

2.4. From the scalar to the vector and tensor context

In the second half of the last century, a physically motivated approach for the decomposition of spherical vector and tensor fields was presented based on a spherical variant of the Helmholtz theorem, e.g., [6–8, 75]. Following this concept, the tangential part of a spherical vector field is split up into a curl-free and a divergence-free field by use of two differential operators, viz. the surface gradient and the surface curl gradient. Of course, an analogous splitting is valid in tensor theory.

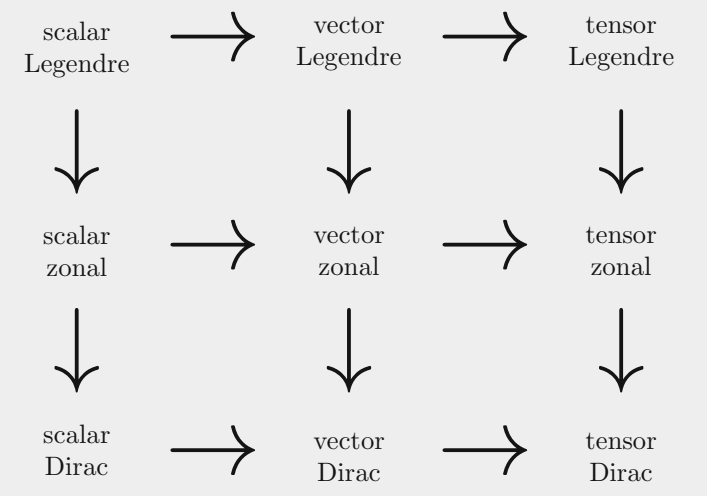


TABLE 3. From scalar via vectorial to tensorial kernels

In subsequent publications during the second half of the last century, however, the vector spherical harmonic theory was usually written in local coordinate expressions that make mathematical formulations lengthy and hard to read. Tensor spherical harmonic settings were even more difficult to understand. In addition, when using local coordinates within a global spherical concept, differential geometry tells us that there is no representation of vector and tensor spherical harmonics that is free of singularities. As a consequence, vector and tensor spherical harmonics have suffered from an inadequately complex and inconsistent literature. Absent coordinate-free explicit formulas, the orthogonal invariance based on specific vector/tensor extensions of the Legendre polynomials was not worked out suitably in a unifying scalar/vector/tensor framework, nor was the concept of zonal (kernel) functions adequately generalized to the spherical vector/tensor case.

All new structures for spherical functions in mathematical (geo)physics were developed by Freeden and Schreiner [43] and Freeden and Gutting [32]. Two fundamental transitions underlie their approach: one from spherical harmonics via zonal kernel functions to the Dirac kernels, and the other one from scalar to vector and tensor theory (see Table 3).

Helmholtz decomposition of spherical vector/tensor fields. To explain the *transition from the theory of scalar spherical harmonics to its vectorial and tensorial extensions*, Freeden and Schreiner [43] start from physically motivated dual pairs of operators (the reference space being always the space of signals with finite energy, i.e., the space of square-integrable fields). The pair $o^{(i)}, O^{(i)}, i \in \{1,2,3\}$, originates in the ingredients of the Helmholtz decomposition of a vector field, while

$o^{(i,k)}, O^{(i,k)}, i, k \in \{1,2,3\}$, take the analogous role for the Helmholtz decomposition of tensor fields (see, e.g., [7, 43, 47]).

For example, in vector theory, $o^{(1)}F$ is the normal field

$$\xi \mapsto o_\xi^{(1)} F(\xi) = F(\xi)\xi, \quad \xi \in \Omega,$$

whereas $o^{(2)}F$ is the surface gradient field

$$\xi \mapsto o_\xi^{(2)} F(\xi) = \nabla_\xi^* F(\xi), \quad \xi \in \Omega,$$

and $o^{(3)}F$ is the surface curl gradient field

$$\xi \mapsto o_\xi^{(3)} F(\xi) = \mathrm{L}_\xi^* F(\xi), \quad \mathrm{L}_\xi^* = \xi \wedge \nabla_\xi^*, \quad \xi \in \Omega,$$

applied to a scalar function F. In addition, $O^{(1)}f$ is the normal component

$$\xi \mapsto O_\xi^{(1)} f(\xi) = f(\xi) \cdot \xi, \quad \xi \in \Omega,$$

while $O^{(2)}f$ is the negative surface divergence

$$\xi \mapsto O_\xi^{(2)} f(\xi) = -\nabla_\xi^* \cdot f(\xi), \quad \xi \in \Omega,$$

and $O^{(3)}f$ is the negative surface curl

$$\xi \mapsto O_\xi^{(3)} f(\xi) = -\mathrm{L}_\xi^* \cdot f(\xi), \quad \xi \in \Omega,$$

taken over a vector-valued function f.

Clearly, the operators $o^{(i,k)}, O^{(i,k)}$ are also definable in orientation to the tensor Helmholtz decomposition theorem (for reasons of simplicity, however, their explicit description is omitted here). The pairs $o^{(i)}, O^{(i)}$ and $o^{(i,k)}, O^{(i,k)}$ of dual operators lead us to an associated palette of Legendre kernel functions, all of them generated by the classical one-dimensional Legendre polynomial P_n of degree n. To be more specific, three types of Legendre kernels occur in the vectorial as well as tensorial context (see Table 4).

The Legendre kernels $o^{(i)}P_n$ and $o^{(i)}o^{(i)}P_n$ pertain to the vector approach for spherical harmonics, whereas $o^{(i,k)}P_n$ and $o^{(i,k)}o^{(i,k)}P_n$, $i, k = 1,2,3$, form the analogues in tensorial theory. Corresponding to each Legendre kernel, we are led to two variants for representing square-integrable fields by orthogonal (Fourier) expansion, where the reconstruction – as in the scalar case – is undertaken by superposition over all frequencies.

In a unified notation, the formalism for vector/tensor spherical harmonic theory is based on the following principles (cf. [43]):

(i) The vector/tensor spherical harmonics involving the $o^{(i)}, o^{(i,k)}$-operators, respectively, are obtainable as restrictions of three-dimensional homogeneous harmonic vector/tensor polynomials, respectively.

(ii) The vector/tensor Legendre kernels are obtainable as the outcome of sums extended over a maximal orthonormal system of vector/tensor spherical harmonics of degree (frequency) n, respectively.

Scalar Legendre polynomial: $P_n = \frac{O^{(i)}O^{(i)}\mathbf{p}_n^{(i,i)}}{\mu_n^{(i)}} = \frac{O^{(i,k)}O^{(i,k)}\mathbf{P}_n^{(i,k)}}{\mu_n^{(i,k)}}$	
application of $o^{(i)}$ ↓↑ application of $O^{(i)}$	application of $o^{(i,k)}$ ↓↑ application of $O^{(i,k)}$
vector Legendre kernel	tensor Legendre kernel (order 2)
$p_n^{(i)} = \frac{o^{(i)}P_n}{\left(\mu_n^{(i)}\right)^{1/2}} = \frac{O^{(i)}\mathbf{p}_n^{(i,i)}}{\left(\mu_n^{(i)}\right)^{1/2}}$	$\mathbf{p}_n^{(i,k)} = \frac{o^{(i,k)}P_n}{\left(\mu_n^{(i,k)}\right)^{1/2}} = \frac{O^{(i,k)}\mathbf{P}_n^{(i,k)}}{\left(\mu_n^{(i,k)}\right)^{1/2}}$
application of $o^{(i)}$ ↓↑ application of $O^{(i)}$	application of $o^{(i,k)}$ ↓↑ application of $O^{(i,k)}$
tensor Legendre kernel (order 2)	tensor Legendre kernel (order 4)
$\mathbf{p}_n^{(i,i)} = \frac{o^{(i)}p_n^{(i)}}{\left(\mu_n^{(i)}\right)^{1/2}} = \frac{o^{(i)}o^{(i)}P_n}{\mu_n^{(i)}}$	$\mathbf{P}_n^{(i,k,i,k)} = \frac{o^{(i,k)}\mathbf{p}_n^{(i,k)}}{\left(\mu_n^{(i,k)}\right)^{1/2}} = \frac{o^{(i,k)}o^{(i,k)}P_n}{\mu_n^{(i,k)}}$
vectorial context	*tensorial context*

TABLE 4. Legendre scalar, vectorial, and tensorial kernel functions.

(iii) The vector/tensor Legendre kernels are zonal kernel functions, rotation-invariant (in vector/tensor sense, respectively) with respect to orthogonal transformations (leaving one point of the unit sphere Ω fixed).

(iv) Spherical harmonics of degree (frequency) n form an irreducible subspace of the reference space of (square-integrable) fields on Ω.

(v) Each Legendre kernel implies an associated Funk–Hecke formula that determines the constituting features of the convolution (filtering) of a square-integrable field against the Legendre kernel.

(vi) The orthogonal Fourier expansion of a square-integrable field is the sum of the convolutions of the field against the Legendre kernels being extended over all frequencies.

To summarize, the theory of spherical harmonics provides us with a framework to unify, review and supplement the different approaches in real scalar, vector, and tensor theory. The essential tools are the Legendre functions, used in orthogonal Fourier expansions and endowed with rotational invariance. The coordinate-free construction yields a number of formulas and theorems that previously were derived only in coordinate (e.g., polar) representations. Consequently, any kind of singularities is avoided at the poles. Finally, our transition from the scalar to

the vectorial as well as the tensorial case opens new promising perspectives of constructing important zonal classes of spherical trial functions by summing up Legendre kernel expressions, thereby providing (geo-)physical relevance and increasing local applicability [43]. Similar considerations apply to the construction of vector/tensor Slepian functions, e.g., [21, 79].

2.5. From the sphere to the ball

The modeling of structures inside the Earth requires basis functions on the ball $\mathcal{B} = \{x \in \mathbb{R}^3 : |x| \leq \beta\}$ with $\beta > 0$. Several approaches for the construction of such basis systems exist. Of course, from the mathematical point of view, one could easily take a basis $\{B_k\}_{k=0,1,\ldots}$ on the Cartesian domain $[-\beta, \beta]$ to construct a basis on the cube $[-\beta, \beta]^3$ by simply taking the tensor product basis $(x_1, x_2, x_3) \mapsto B_{k_1}(x_1)B_{k_2}(x_2)B_{k_3}(x_3)$, $k_1, k_2, k_3 = 0, 1, \ldots$ However, the Earth's interior is usually subdivided into structural layers that are approximately bounded by spheres. In view of this fact, the use of cartesian-coordinate-based trial functions appears to be inappropriate and the spherical harmonics also here play an essential role.

An intuitive approach is to look for basis functions of the form

$$G_{m,n,k}(r\xi) = F_{m,n}(r)Y_{n,k}(\xi), \quad \xi \in \Omega,\ r \in [0, \beta],$$

for $m, n = 0, 1, \ldots$ and $k = 1, \ldots, 2n + 1$. Also here, orthogonality appears to be useful, which leads to the requirement that

$$\begin{aligned}
&\int_{\mathcal{B}} G_{m_1,n_1,k_1}(x)\, G_{m_2,n_2,k_2}(x)\, \mathrm{d}x \\
&\quad = \int_0^\beta r^2 F_{m_1,n_1}(r)\, F_{m_2,n_2}(r)\, \mathrm{d}r \int_\Omega Y_{n_1,k_1}(\xi)\, Y_{n_2,k_2}(\xi)\, \mathrm{d}\omega(\xi) \\
&\quad = \int_0^\beta r^2 F_{m_1,n_1}(r)\, F_{m_2,n_1}(r)\, \mathrm{d}r = 0,
\end{aligned} \tag{18}$$

if $m_1 \neq m_2$ or $n_1 \neq n_2$ or $k_1 \neq k_2$. The weight function r^2 in the radial integral in (18) suggests the use of the Jacobi polynomials as building blocks for $F_{m,n}$. However, there is a notable degree of freedom in the choice of (e.g., polynomial) functions for $F_{m,n}$. This degree of freedom can be used to construct the $G_{m,n,k}$ in a manner such that they characterize the non-uniqueness of solutions of tomographic inverse problems in the geosciences or medical imaging. For further details, see [9, 63, 67, 68, 70, 97] and the contribution by Leweke, Michel, and Telschow (this book, pp. 883–919). Note that some of the obtained systems become discontinuous or even singular at the origin $0 \in \mathcal{B}$ but in a way such that they are still elements of $\mathrm{L}^2(\mathcal{B})$.

The fact that such orthonormal basis functions on the ball arise from the spherical harmonics as orthonormal basis functions on the sphere yields a way to formulate analogies regarding the methodologies and the associated properties – though often further difficulties occur due to the additional radial coordinate.

Particular analogies exist with respect to the space and 'frequency' localization of kernels

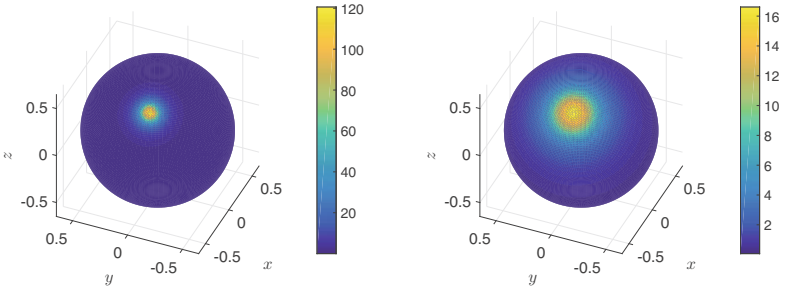
$$K(x,y) = \sum_{m,n=0}^{\infty} \sum_{k=1}^{2n+1} K^\wedge(m,n)\, G_{m,n,k}(x)\, G_{m,n,k}(y)$$
$$= \sum_{m,n=0}^{\infty} K^\wedge(m,n) F_{m,n}(|x|) F_{m,n}(|y|) \frac{2n+1}{4\pi} P_n\left(\frac{x}{|x|} \cdot \frac{y}{|y|}\right), \qquad (19)$$

$x, y \in \mathcal{B} \setminus \{0\}$. It should be noted, however, that most of the choices of $F_{m,n}$ do *not* lead to functions $G_{m,n,k}$ which are algebraic polynomials in x_1, x_2, x_3. Nevertheless, the properties of the Jacobi polynomials and the spherical harmonics imply that the $G_{m,n,k}$ also show an increasing oscillatory behavior for increasing m or n. Furthermore, the Dirac kernel can also here be associated to the case where all coefficients satisfy $K^\wedge(m,n) = 1$.

One of the advantages of this approach – in contrast to a cartesian setup – is that the obtained kernels inherit the rotational invariance from the spherical kernels in the sense that

$$K(r\mathbf{t}\xi, s\mathbf{t}\eta) = \sum_{n=0}^{\infty} K^\wedge(m,n) F_{m,n}(r) F_{m,n}(s) \frac{2n+1}{4\pi} P_n((\mathbf{t}\xi)\cdot(\mathbf{t}\eta))$$
$$= K(r\xi, s\eta)$$

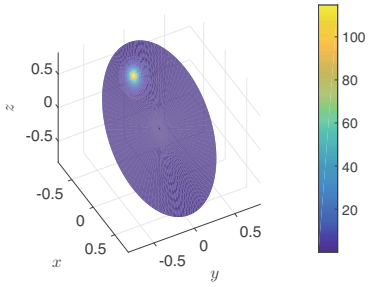
for all 3×3-orthogonal matrices (i.e., rotations) $\mathbf{t}$, see Figure 2.



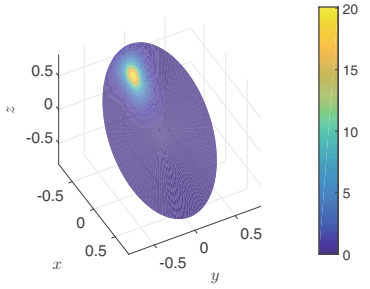
(A) Kernel with $K^\wedge(m,n) = 0.8^m 0.9^n$ (B) Kernel with $K^\wedge(m,n) = 0.8^m 0.7^n$

FIGURE 2. The figures show localized trial functions $\mathcal{B} \ni y \mapsto K(x,y)$ based on a kernel of the kind in (19) with a fixed point $x = (-0.4, -0.1, 0.5)^{\mathrm{T}}$. The functions are plotted on the sphere with radius $|x|$. Each function is a hat function concentrated around x. Its restriction to a sphere around 0 is a rotationally symmetric function, as it is known for the case of spherical kernels. Note that the series representations were truncated at $n = m = 400$ in the numerical implementation.

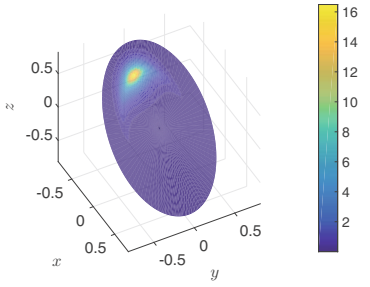
Moreover, the localization with respect to the radius (or the 'depth') can be separated from the localization with respect to the angular coordinates by taking, for example, symbols of the form $K^\wedge(m,n) = A_m B_n$ like $K^\wedge(m,n) = h_{\mathrm{r}}^m h_{\mathrm{ang}}^n$ for parameters $h_{\mathrm{r}}, h_{\mathrm{ang}} \in (0,1)$, see Figure 3. This is useful, e.g., for tomographic problems where it is known that the solution has a finer structure in the angular domain than in the radial domain (or vice versa).



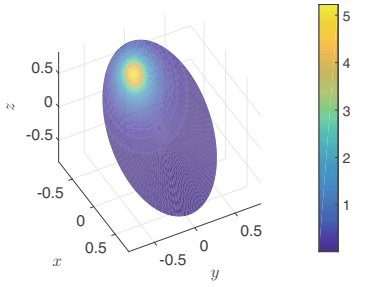
(A) Kernel with $h_{\mathrm{r}} = 0.8$ and $h_{\mathrm{ang}} = 0.9$



(B) Kernel with $h_{\mathrm{r}} = 0.5$ and $h_{\mathrm{ang}} = 0.9$



(C) Kernel with $h_{\mathrm{r}} = 0.8$ and $h_{\mathrm{ang}} = 0.7$



(D) Kernel with $h_{\mathrm{r}} = 0.5$ and $h_{\mathrm{ang}} = 0.7$

FIGURE 3. The figures show localized trial functions $\mathcal{B} \ni y \mapsto K(x,y)$ based on a kernel of the kind in (19) with a fixed point $x = (-0.4, -0.1, 0.5)^{\mathrm{T}}$. The functions are plotted on the planar cross section with normal vector $(1,1,1)^{\mathrm{T}}$. By choosing a symbol $K^\wedge(m,n) = h_{\mathrm{r}}^m h_{\mathrm{ang}}^n$, the localization in radial and in angular domain can be controlled separately. Note that the series representations were truncated at $n = m = 400$ in the numerical implementation.

3. Spherical uncertainty principle

As pointed out in Section 2, four classes of zonal kernel functions can be distinguished, namely bandlimited and non-bandlimited, spacelimited and non-spacelimited ones. In addition, Slepian functions exist in bandlimited and spacelimited varieties. What is the right kernel function for the purpose of local approximation? Of course, the user of a mathematical method is interested in knowing the trial system which fits 'adequately' to the problem. When several choices are possible or an optimal choice cannot be found it is necessary to choose the trial systems in close adaptation to the data extent or density, and the required smoothness of the field to be approximated. This, however, is often a local property, i.e., the data density can be high in one area and low in another. In addition, the field to be approximated can have a high-detail structure in some parts of the sphere (e.g., over mountainous regions) and a low-detail structure elsewhere (e.g., over the oceans). This makes the selection of appropriate trial functions even more challenging.

3.1. Derivation and basic theory

An uncertainty principle that specifies the degree of space and frequency localization is helpful to serve as a decisive criterion. The essential outcome is a better understanding of the classification of zonal kernel functions, and Slepian functions, based on the development of suitable bounds for their quantification with respect to space and frequency localization.

Localization in space. Assume F is of class $\mathrm{L}^2(\Omega)$ with energy

$$\|F\|_{\mathrm{L}^2(\Omega)} = \left(\int_\Omega \left(F(\eta)\right)^2 \, \mathrm{d}\omega(\eta) \right)^{1/2} = 1.$$

We associate to F the normal (radial) field $\eta \mapsto \eta F(\eta) = o_\eta^{(1)} F(\eta)$, $\eta \in \Omega$. This function maps $\mathrm{L}^2(\Omega)$ into the associated set of normal fields on Ω. The 'center of gravity' of F is the *expectation of the normal operator* $o^{(1)}$ *on* Ω,

$$g_F^{o^{(1)}} = \int_\Omega \left(o_\eta^{(1)} F(\eta) \right) F(\eta) \, \mathrm{d}\omega(\eta) = \int_\Omega \eta (F(\eta))^2 \, \mathrm{d}\omega(\eta) \in \mathbb{R}^3 \tag{20}$$

thereby interpreting $(F(\eta))^2 \, \mathrm{d}\omega(\eta)$ as surface mass distribution over the sphere Ω embedded in Cartesian space $\mathbb{R}^3$. It is clear that $g_F^{o^{(1)}}$ lies in the closed inner space $\overline{\Omega^{\mathrm{int}}}$ of Ω: $|g_F^{o^{(1)}}| \leq 1$. The variance of the operator $o^{(1)}$ is understood in the canonical sense as the *variance in the space domain*,

$$\begin{aligned} \left(\sigma_F^{o^{(1)}} \right)^2 &= \int_\Omega \left(\left(o_\eta^{(1)} - g_F^{o^{(1)}} \right) F(\eta) \right)^2 \, \mathrm{d}\omega(\eta) \\ &= \int_\Omega \left(\eta - g_F^{o^{(1)}} \right)^2 \left(F(\eta)\right)^2 \, \mathrm{d}\omega(\eta) \in \mathbb{R}. \end{aligned} \tag{21}$$

Observing the identity $(\eta - g_F^{o^{(1)}})^2 = 1 + (g_F^{o^{(1)}})^2 - 2\eta \cdot g_F^{o^{(1)}}$, $\eta \in \Omega$, it follows immediately that $(\sigma_F^{o^{(1)}})^2 = 1 - (g_F^{o^{(1)}})^2$. Naturally, $0 \leq (\sigma_F^{o^{(1)}})^2 \leq 1$.

Since we are particularly interested in zonal functions, some simplifications can be made. Let K be of class $\mathrm{L}^2[-1,1]$ and $\|K\|_{\mathrm{L}^2[-1,1]} = 1$, where $\|F\|_{\mathrm{L}^2[-1,1]} = (2\pi \int_{-1}^{1} (F(t))^2 \,\mathrm{d}t)^{1/2}$ for $F \in \mathrm{L}^2[-1,1]$. Then the corresponding center of gravity can be computed readily as follows ($\varepsilon^3 = (0,0,1)^{\mathrm{T}}$):

$$g^{o^{(1)}}_{K(\cdot\varepsilon^3)} = \int_\Omega \eta \left(K\left(\eta \cdot \varepsilon^3\right)\right)^2 \,\mathrm{d}\omega(\eta) = \left(2\pi \int_{-1}^{1} t\,(K(t))^2 \,\mathrm{d}t\right) \varepsilon^3. \tag{22}$$

Letting $t^{o^{(1)}}_K = \left| g^{o^{(1)}}_{K(\cdot\varepsilon^3)} \right| = 2\pi \left| \int_{-1}^{1} t\,(K(t))^2 \,\mathrm{d}t \right| \in \mathbb{R}$ we find for the variance

$$\begin{aligned} \left(\sigma^{o^{(1)}}_K\right)^2 &= \int_\Omega \left(\eta - g^{o^{(1)}}_{K(\cdot\varepsilon^3)}\right)^2 \left(K\left(\eta \cdot \varepsilon^3\right)\right)^2 \,\mathrm{d}\omega(\eta) \\ &= 1 - \left(t^{o^{(1)}}_K\right)^2 = 1 - \left(g^{o^{(1)}}_{K(\cdot\varepsilon^3)}\right)^2 \in \mathbb{R}. \end{aligned} \tag{23}$$

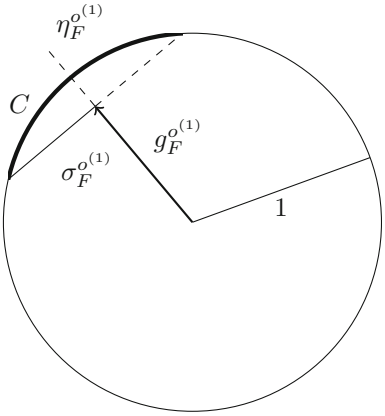


FIGURE 4. Localization in a spherical cap.

Figure 4 gives a geometric interpretation of $g^{o^{(1)}}_F$ and $\sigma^{o^{(1)}}_F$. We associate to $g^{o^{(1)}}_F$, $g^{o^{(1)}}_F \neq 0$, and its projection $\eta^{o^{(1)}}_F$ onto the sphere Ω the spherical cap $C = \{\eta \in \Omega \mid 1 - \eta \cdot \eta^{o^{(1)}}_F \leq 1 - |g^{o^{(1)}}_F|\}$. Then the boundary ∂C is a circle with radius $\sigma^{o^{(1)}}_F$. Thinking of a zonal function F as a 'spherical window function' on Ω, the window is determined by C, and its width is given by $\sigma^{o^{(1)}}_F$.

Localization in frequency ('momentum space'). The 'expectation in the frequency domain' is introduced as the *expectation of the surface curl operator* $o^{(3)}$ on Ω. Then, for $F \in \mathrm{H}^{(2l)}(\Omega)$, $l \in \mathbb{N}$, i.e., for all $F \in \mathrm{L}^2(\Omega)$ such that there exists a function $G \in \mathrm{L}^2(\Omega)$ with $G^\wedge(n,k) = (-n(n+1))^l F^\wedge(n,k)$ for all $n = 0,1,\ldots$; $k = 1,\ldots,2n+1$, we have

$$g^{o^{(3)}}_F = \int_\Omega \left(o^{(3)}_\eta F(\eta)\right) F(\eta) \,\mathrm{d}\omega(\eta) = 0 \in \mathbb{R}^3. \tag{24}$$

	operator	expectation value
space	$o^{(1)}$	$g_F^{o^{(1)}} = \int_\Omega \left(o_\eta^{(1)} F(\eta)\right) F(\eta)\, \mathrm{d}\omega(\eta)$
frequency	$o^{(3)}$	$g_F^{o^{(3)}} = \int_\Omega \left(o_\eta^{(3)} F(\eta)\right) F(\eta)\, \mathrm{d}\omega(\eta)$

	operator	variance
space	$o^{(1)}$	$\left(\sigma_F^{o^{(1)}}\right)^2 = \int_\Omega \left(\left(o_\eta^{(1)} - g_F^{o^{(1)}}\right) F(\eta)\right)^2 \mathrm{d}\omega(\eta)$
frequency	$o^{(3)}$	$\left(\sigma_F^{o^{(3)}}\right)^2 = \int_\Omega \left(\left(o_\eta^{(3)} - g_F^{o^{(3)}}\right) F(\eta)\right)^2 \mathrm{d}\omega(\eta)$

TABLE 5. Localization in terms of the normal and curl operators $o^{(1)}$ and $o^{(3)}$.

Correspondingly, the *variance in the frequency domain* is given by

$$\left(\sigma_F^{o^{(3)}}\right)^2 = \int_\Omega \left(\left(o_\eta^{(3)} - g_F^{o^{(3)}}\right) F(\eta)\right)^2 \mathrm{d}\omega(\eta) \in \mathbb{R}. \tag{25}$$

The surface theorem of Stokes shows us that

$$\begin{aligned}\left(\sigma_F^{o^{(3)}}\right)^2 &= \int_\Omega \left(o_\eta^{(3)} F(\eta)\right) \cdot \left(o_\eta^{(3)} F(\eta)\right) \mathrm{d}\omega(\eta) \\ &= \int_\Omega \left(-\Delta_\eta^* F(\eta)\right) F(\eta)\, \mathrm{d}\omega(\eta) \quad = \quad g_F^{-\Delta^*}.\end{aligned} \tag{26}$$

Expressed in terms of spherical harmonics we get via the Parseval identity

$$\left(\sigma_F^{o^{(3)}}\right)^2 = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} n(n+1) \left(F^\wedge(n,k)\right)^2. \tag{27}$$

Note that we require $\|F\|^2_{\mathrm{L}^2(\Omega)} = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \left(F^\wedge(n,k)\right)^2 = 1$. The meaning of $\sigma_F^{o^{(3)}}$ as measure of 'frequency localization' is as follows: the range of $\sigma_F^{o^{(3)}}$ is the interval $[0, \infty]$; a large value of $\sigma_F^{o^{(3)}}$ occurs if many Fourier coefficients contribute to $\sigma_F^{o^{(3)}}$. In conclusion, relating any spherical harmonic to a 'single wavelength', a large value $\sigma_F^{o^{(3)}}$ informs us that F is spread out widely in 'frequency domain'. In contrast, a small value $\sigma_F^{o^{(3)}}$ indicates that only a few number of Fourier coefficients is significant (see Table 5).

Again we reformulate our quantities in the specific context of zonal functions. Let $K(\cdot\varepsilon^3)$ be of class $\mathrm{H}^{(2)}(\Omega)$ satisfying $\|K(\cdot\varepsilon^3)\|_{\mathrm{L}^2(\Omega)} = 1$, then

$$\begin{aligned}\left(\sigma^{o^{(3)}}_{K(\cdot\varepsilon^3)}\right)^2 &= -\int_\Omega \Delta^*_\eta K\left(\eta\cdot\varepsilon^3\right) K\left(\eta\cdot\varepsilon^3\right)\,\mathrm{d}\omega(\eta)\\ &= -2\pi\int_{-1}^{1} K(t)\mathrm{L}_t K(t)\,\mathrm{d}t \end{aligned} \tag{28}$$

where L_t denotes the Legendre operator as given by $\mathrm{L}_t = \frac{\mathrm{d}}{\mathrm{d}t}(1-t^2)\frac{\mathrm{d}}{\mathrm{d}t}$.

Uncertainties and the uncertainty principle. The square roots of the variances, i.e., $\sigma^{o^{(1)}}$ and $\sigma^{o^{(3)}}$, are called the *uncertainties* in $o^{(1)}$ and $o^{(3)}$, respectively. For these quantities we get (see [43, 76]) an estimate given by $(\sigma^{o^{(1)}}_F)^2(\sigma^{o^{(3)}}_F)^2 \geq |g^{o^{(1)}}_F|^2$. We summarize our results in Theorem 1. For details on the proof the reader is referred to [29].

Theorem 1. *Let $F \in \mathrm{H}^{(2)}(\Omega)$ satisfy $\|F\|_{\mathrm{L}^2(\Omega)} = 1$. Then*

$$\left(\sigma^{o^{(1)}}_F\right)^2\left(\sigma^{o^{(3)}}_F\right)^2 \geq \left|g^{o^{(1)}}_F\right|^2. \tag{29}$$

If $g^{o^{(1)}}_F$ is non-vanishing, then

$$\Delta^{o^{(1)}}_F \Delta^{o^{(3)}}_F \geq 1, \tag{30}$$

where we have used the abbreviations

$$\Delta^{o^{(1)}}_F = \frac{\sigma^{o^{(1)}}_F}{\left|g^{o^{(1)}}_F\right|}, \quad \Delta^{o^{(3)}}_F = \sigma^{o^{(3)}}_F. \tag{31}$$

The *uncertainty relation* measures the tradeoff between 'space localization' and 'frequency localization' ('spread in frequency'). It states that *sharp localization in space and frequency are mutually exclusive.*

An immediate consequence of Theorem 1 is its reformulation for zonal functions $K(\varepsilon^3\cdot) : \eta \mapsto K(\varepsilon^3\cdot\eta)$, $\eta \in \Omega$.

Corollary 2. *Let $K(\varepsilon^3\cdot) \in \mathrm{H}^{(2)}(\Omega)$ satisfy $\|K\|_{\mathrm{L}^2[-1,1]} = 1$. If $t^{o^{(1)}}_K$ is non-vanishing, then*

$$\Delta^{o^{(1)}}_K \Delta^{o^{(3)}}_K \geq 1, \tag{32}$$

where

$$\Delta^{o^{(1)}}_K = \frac{\sigma^{o^{(1)}}_K}{t^{o^{(1)}}_K}, \quad \Delta^{o^{(3)}}_K = \sigma^{o^{(3)}}_K. \tag{33}$$

The interpretation of $(\sigma^{o^{(3)}}_K)^2$ as variance in 'total angular momentum' helped us to prove Theorem 1. However, this interpretation shows two essential drawbacks: first, the expectation of the surface curl gradient is a vector which seems to be inadequate in 'momentum localization' in terms of scalar spherical harmonics, and secondly the value of $g^{o^{(3)}}_F$ vanishes for all candidates F. This means that the

'center of gravity of the spherical window' in 'momentum domain' is independent of the function F under consideration. Therefore, we are finally interested in the variance of the operator $-\Delta^*$

$$\left(\sigma_F^{-\Delta^*}\right)^2 = \int_\Omega \left| \left((-\Delta_\eta^*) - g_F^{-\Delta^*} \right) F(\eta) \right|^2 \, \mathrm{d}\omega(\eta) \tag{34}$$

which is a measure for the 'spread in momentum'. Now the corresponding expectation value $g_F^{-\Delta^*}$ is scalar-valued and non-vanishing. It can be easily seen that

$$\left(\sigma_F^{-\Delta^*}\right)^2 = g_F^{(-\Delta^*)^2} - \left(g_F^{-\Delta^*}\right)^2. \tag{35}$$

In connection with Theorem 1 this leads to the following result.

Theorem 3. *Let F be of class $\mathrm{H}^{(4)}(\Omega)$ such that $\|F\|_{\mathrm{L}^2(\Omega)} = 1$. Then*

$$\left(\sigma_F^{o^{(1)}}\right)^2 \left(\sigma_F^{-\Delta^*}\right)^2 \geq \left| g_F^{o^{(1)}} \right| \, \frac{g_F^{(-\Delta^*)^2} - \left(g_F^{-\Delta^*}\right)^2}{g_F^{-\Delta^*}} \tag{36}$$

provided that $g_F^{-\Delta^*} \neq 0$. If the right-hand side of (36) *is non-vanishing, then*

$$\Delta_F^{o^{(1)}} \Delta_F^{-\Delta^*} \geq 1, \tag{37}$$

where

$$\Delta_F^{-\Delta^*} = \left(\frac{\left(\sigma_F^{-\Delta^*}\right)^2}{\frac{g_F^{(-\Delta^*)^2} - \left(g_F^{-\Delta^*}\right)^2}{g_F^{-\Delta^*}}} \right)^{1/2} = \left(g_F^{-\Delta^*}\right)^{1/2} = \Delta_F^{o^{(3)}}. \tag{38}$$

3.2. Classification of examples

We continue with some examples of particular interest for geoscientific research.

Localization of the spherical harmonics. We know that

$$\int_\Omega (Y_{n,k}(\xi))^2 \, \mathrm{d}\omega(\xi) = 1 \ . \tag{39}$$

One can prove that

$$g_{Y_{n,k}}^{o^{(1)}} = 0, \qquad \sigma_{Y_{n,k}}^{o^{(1)}} = 1. \tag{40}$$

Moreover, we have

$$g_{Y_{n,k}}^{-\Delta^*} = n(n+1), \qquad \sigma_{Y_{n,k}}^{-\Delta^*} = 0. \tag{41}$$

In other words, *spherical harmonics show an ideal frequency localization, but no space localization* (see Figure 5 for an illustration of space and frequency localization for the Legendre polynomials).

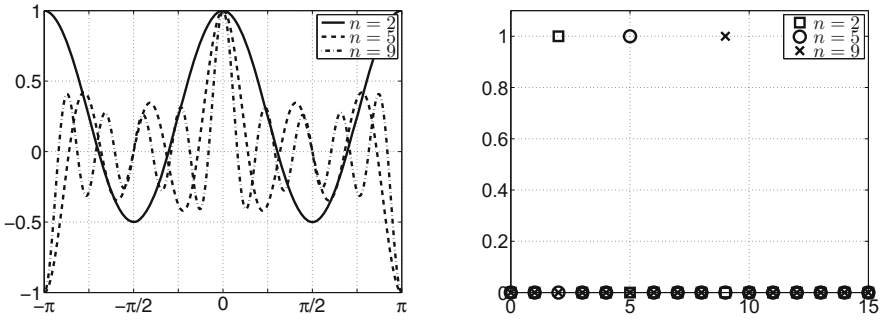


FIGURE 5. The Legendre kernel P_n for $n = 2, 5, 9$, space representation for $\vartheta \mapsto P_n(\cos(\vartheta))$ (left), and frequency representation $m \mapsto (P_n)^\wedge(m)$ (right).

Localization of the ideally bandlimited Legendre kernel. We have, with $P_n^* = \sqrt{\frac{2n+1}{4\pi}} P_n$,

$$\int_\Omega (P_n^*(\xi \cdot \zeta))^2 \, \mathrm{d}\omega(\zeta) = 1 \tag{42}$$

for all $\xi \in \Omega$, such that

$$g^{o^{(1)}}_{P_n^*(\xi\cdot)} = 0, \qquad \sigma^{o^{(1)}}_{P_n^*(\xi\cdot)} = 1, \tag{43}$$

$$g^{-\Delta^*}_{P_n^*(\xi\cdot)} = n(n+1), \quad \sigma^{-\Delta^*}_{P_n^*(\xi\cdot)} = 0. \tag{44}$$

Localization of the bandlimited Shannon kernel. The Shannon kernel Φ_ρ, $\rho > 0$, given by

$$\Phi_\rho(\xi \cdot \eta) = \sum_{n \le \rho^{-1}} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta), \quad \xi, \eta \in \Omega, \tag{45}$$

may be interpreted as a truncated Dirac kernel. It is not surprising that the Shannon kernel as a 'finite polynomial kernel' shows strong oscillations in space. This is the price to be paid for the sharp separation in frequency space.

The investigation of the uncertainty properties of the Shannon kernel starts from (cf. [43])

$$\|\Phi_\rho\|^2_{\mathrm{L}^2(\Omega)} = \sum_{n=0}^{\lfloor \rho^{-1} \rfloor} \frac{2n+1}{4\pi} = \frac{1}{4\pi} \left((\lfloor \rho^{-1} \rfloor + 1) + \lfloor \rho^{-1} \rfloor \lfloor \rho^{-1} + 1 \rfloor \right), \tag{46}$$

where, as usual, $\lfloor \rho^{-1} \rfloor$ is the largest integer which is less or equal ρ^{-1}. Observing this result, we introduce the normalized Shannon kernel by

$$\tilde{\Phi}_\rho = \frac{1}{\|\Phi_\rho\|_{\mathrm{L}^2(\Omega)}} \Phi_\rho. \tag{47}$$

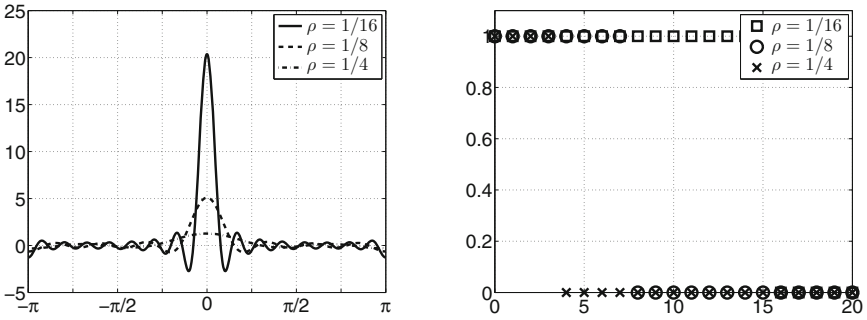


FIGURE 6. The Shannon scaling function Φ_ρ for $\rho = 1/16,\ 1/8,\ 1/4$. Space representation $\vartheta \mapsto \Phi_\rho(\cos(\vartheta))$ (left) and frequency representation $n \mapsto (\Phi_\rho)^\wedge(n)$ (right).

Its localization in space satisfies

$$\begin{aligned}\left(\sigma^{o(1)}_{\tilde{\Phi}_\rho}\right)^2 &= 1 - \frac{1}{\|\Phi_\rho\|^2}\left(\sum_{n=1}^{\lfloor \rho^{-1}-1\rfloor} \frac{2n+2}{4\pi}\right)^2 \\ &= 1 - \left(\frac{2\lfloor \rho^{-1}-1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}-1\rfloor}{\lfloor \rho^{-1}+1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor}\right)^2, \end{aligned} \tag{48}$$

so that

$$\Delta^{o(1)}_{\tilde{\Phi}_\rho} = \sqrt{\frac{1 - \left(\frac{2\lfloor \rho^{-1}-1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}-1\rfloor}{\lfloor \rho^{-1}+1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor}\right)^2}{\frac{2\lfloor \rho^{-1}-1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}-1\rfloor}{\lfloor \rho^{-1}+1\rfloor + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor}}}. \tag{49}$$

Moreover, we find

$$\begin{aligned}\left(\sigma^{o(3)}_{\tilde{\Phi}_\rho}\right)^2 &= \frac{4\pi}{\lfloor \rho^{-1}\rfloor + 1 + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor} \sum_{n=0}^{\lfloor \rho^{-1}\rfloor} \frac{2n+1}{4\pi}\, n(n+1) \\ &= \frac{1}{2}\frac{\lfloor \rho^{-1}\rfloor(1+\lfloor \rho^{-1}\rfloor)^2(2+\lfloor \rho^{-1}\rfloor)}{\lfloor \rho^{-1}\rfloor + 1 + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor} \end{aligned} \tag{50}$$

such that

$$\Delta^{o(3)}_{\tilde{\Phi}_\rho} = \sqrt{\frac{1}{2}\frac{\lfloor \rho^{-1}\rfloor(1+\lfloor \rho^{-1}\rfloor)^2(2+\lfloor \rho^{-1}\rfloor)}{\lfloor \rho^{-1}\rfloor + 1 + \lfloor \rho^{-1}\rfloor\lfloor \rho^{-1}+1\rfloor}}. \tag{51}$$

The results are graphically illustrated in Figure 7.

Localization of the non-bandlimited/non-spacelimited Abel–Poisson kernel. Let us consider the function $Q_h : [-1,1] \to \mathbb{R}$, $h < 1$, given by

$$Q_h(t) = \frac{1}{4\pi}\frac{1-h^2}{(1+h^2-2ht)^{3/2}} = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, h^n P_n(t). \tag{52}$$

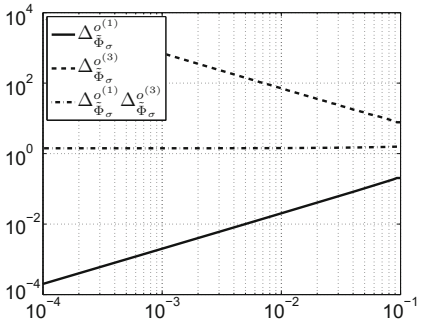


FIGURE 7. Uncertainty classification of the normalized Shannon Dirac family $\tilde{\Phi}_\rho$. Shown are $\Delta^{o^{(1)}}_{\tilde{\Phi}_\rho}$, $\Delta^{o^{(3)}}_{\tilde{\Phi}_\rho}$, and the product $\Delta^{o^{(1)}}_{\tilde{\Phi}_\rho}\Delta^{o^{(3)}}_{\tilde{\Phi}_\rho}$ as functions of ρ in a double logarithmic setting.

An easy calculation gives us

$$\|Q_h\|_{\mathrm{L}^2[-1,1]} = (Q_{h^2}(1))^{1/2} = \left(\frac{1+h^2}{4\pi}\right)^{1/2} \frac{1}{1-h^2}. \tag{53}$$

Furthermore, for $\tilde{Q}_h(t) = \|Q_h\|^{-1}_{\mathrm{L}^2[-1,1]}\ Q_h(t)$, $t \in [-1,1]$, we obtain after an elementary calculation (see also Figure 8)

$$\Delta^{o^{(1)}}_{\tilde{Q}_h} = \frac{1-h^2}{2h}, \quad \Delta^{-\Delta^*}_{\tilde{Q}_h} = \frac{\sqrt{6}h}{1-h^2}. \tag{54}$$

Thus, we finally obtain

$$\Delta^{o^{(1)}}_{\tilde{Q}_h}\Delta^{-\Delta^*}_{\tilde{Q}_h} = \frac{\sqrt{6}}{2} = \sqrt{\frac{3}{2}} > 1. \tag{55}$$

Here, the value $\Delta^{o^{(1)}}_{\tilde{Q}_h}\Delta^{-\Delta^*}_{\tilde{Q}_h}$ is independent of h. All intermediate cases of 'space-frequency localization' are realized by the Abel–Poisson kernel, but the Abel–Poisson kernel does not satisfy a minimum uncertainty state.

Localization of the spacelimited Haar kernel. Let k be a non-negative integer, i.e., $k \in \mathbb{N}_0$. The (smoothed) Haar kernel $\{B^{(k)}_h\}_{h\in(0,1)} \subset \mathrm{C}^{(k-1)}[-1,1]$ is defined by

$$B^{(k)}_h(t) = \begin{cases} 0 & , \quad t \in [-1,h) \\ \dfrac{(t-h)^k}{(1-h)^k} & , \quad t \in [h,1]. \end{cases} \tag{56}$$

By definition, $B^{(k)}_h$ is non-negative and has the support $[h,1]$. Obviously, the function $B^{(0)}_h$, $h \in (-1,1)$, represents the (classical) Haar function (cf. [53]). The

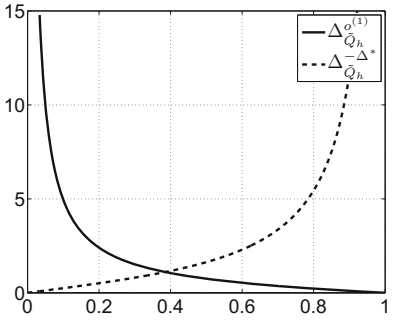


FIGURE 8. Abel–Poisson kernel uncertainty classification. The curves graphically illustrate the functions $h \mapsto \Delta^{o^{(1)}}_{\tilde{Q}_h}$ and $h \mapsto \Delta^{-\Delta^*}_{\tilde{Q}_h}$.

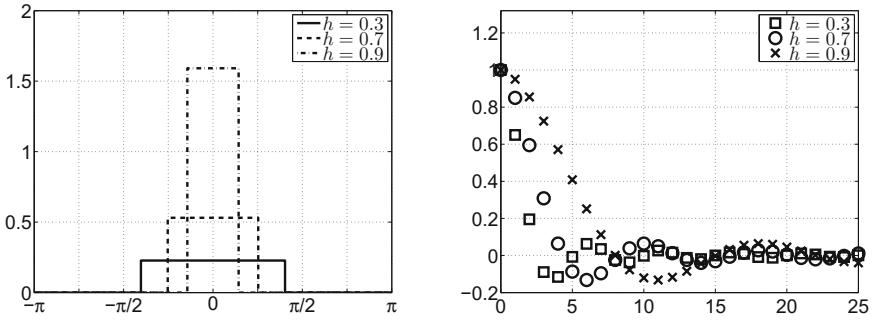


FIGURE 9. The Haar kernel$B^{(0)}_h$ for $h = 0.3,\ 0.7,\ 0.9$. Space representation $\vartheta \mapsto B^{(0)}_h(\cos(\vartheta)), \vartheta \in [-\pi, \pi]$, (left) and frequency representation $n \mapsto (B^{(0)}_h)^\wedge(n)$ (right).

Legendre coefficients of $B^{(k)}_h$, $h \in (-1, 1)$, $k \in \mathbb{N}_0$, can be calculated recursively (cf. [47]):

$$\left(B^{(k)}_h\right)^\wedge(0) = 2\pi \frac{1-h}{k+1} \neq 0, \tag{57}$$

$$\left(B^{(k)}_h\right)^\wedge(1) = 2\pi \frac{1-h}{k+1}\left(1 - \frac{1-h}{k+2}\right), \tag{58}$$

$$\left(B^{(k)}_h\right)^\wedge(n+1) = \frac{2n+1}{n+k+2}\, h\, \left(B^{(k)}_h\right)^\wedge(n) + \frac{k+1-n}{n+k+2}\, \left(B^{(k)}_h\right)^\wedge(n-1). \tag{59}$$

An elementary calculation shows

$$\begin{aligned} \left\| B^{(k)}_h \right\|^2_{\mathrm{L}^2(\Omega)} &= 2\pi \int_{-1}^{1} \left[B^{(k)}_h(t)\right]^2 \mathrm{d}t \\ &= 2\pi \frac{1-h}{2k+1}. \end{aligned} \tag{60}$$

We define the kernel

$$\tilde{B}_h^{(k)} = \sqrt{\frac{2k+1}{2\pi(1-h)}}\, B_h^{(k)}, \tag{61}$$

since the uncertainty properties are normally defined for kernels with norm one. We find

$$g^{o^{(1)}}_{\tilde{B}_h^{(k)}(\,\cdot\varepsilon^3)} = 2\pi \int_{-1}^{1} t \left(\tilde{B}_h^{(k)}(t)\right)^2 \mathrm{d}t\, \varepsilon^3 = \frac{1+h+2k}{2+2k}\, \varepsilon^3. \tag{62}$$

Consequently,

$$\left(\sigma^{o^{(1)}}_{\tilde{B}_h^{(k)}}\right)^2 = 1 - \left(\frac{1+h+2k}{2+2k}\right)^2 = \frac{(1-h)(h+4k+3)}{(2k+2)^2}. \tag{63}$$

Using (31), we finally arrive at

$$\Delta^{o^{(1)}}_{\tilde{B}_h^{(k)}} = \frac{1}{1+h+2k}\sqrt{(1-h)(h+4k+3)}. \tag{64}$$

For the localization in frequency, we assume $k \geq 2$. We have

$$\begin{aligned}\left(\sigma^{o^{(3)}}_{\tilde{B}_h^{(k)}(\,\cdot\varepsilon^3)}\right)^2 &= -2\pi \int_{-1}^{1} \tilde{B}_h^{(k)}(t)\, \mathrm{L}_t \tilde{B}_h^{(k)}(t)\, \mathrm{d}t \\ &= \frac{2k+1}{2\pi(1-h)} \frac{-2\pi}{(1-h)^{2k}} \int_h^1 (t-h)^k \mathrm{L}_t (t-h)^k\, \mathrm{d}t \\ &= \frac{k(h+2k)}{(1-h)(2k-1)},\end{aligned} \tag{65}$$

so that

$$\Delta^{o^{(3)}}_{\tilde{B}_h^{(k)}} = \sqrt{\frac{k(h+2k)}{(1-h)(2k-1)}}. \tag{66}$$

The application of L_t requires that the kernel is twice differentiable. However, using integration by parts, the results immediately carry over to the case $k = 1$. Figure 10 gives a graphical impression of these results for the particular cases $k = 1$ and $k = 3$.

Localization of the ideally spacelimited Dirac kernel. Letting h formally tend to 1 in the results provided by the uncertainty principle for the Abel–Poisson kernel function we are able to interpret the localization properties of the Dirac kernel on Ω satisfying $\delta^\wedge(n) = 1$ for all $n \in \mathbb{N}_0$:

$$\delta(\xi \cdot \eta) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta), \quad \xi, \eta \in \Omega, \tag{67}$$

where the convergence is understood in distributional sense. As a matter of fact, letting h tend to 1 shows us that the variances in the space domain take the constant value 0. On the other hand, the variances in the frequency domain converge to ∞. Hence, the Dirac kernel shows ideal space localization, but no frequency localization.

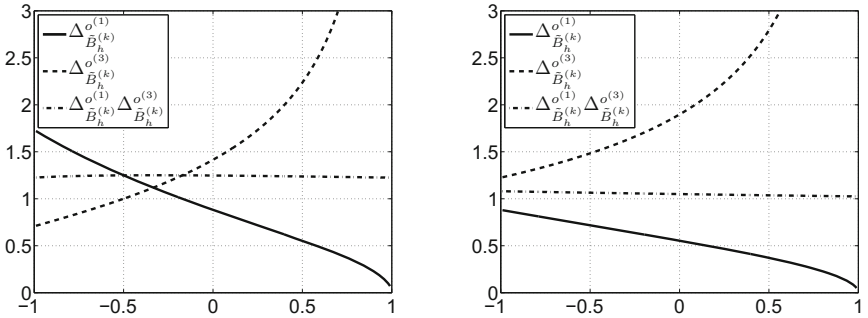


FIGURE 10. Uncertainty classification of the normalized smoothed Haar scaling function $\tilde{B}^{(k)}_h$ ($k = 1$, left; $k = 3$ right). $\Delta^{o^{(1)}}_{\tilde{B}^{(k)}_h}$, $\Delta^{o^{(3)}}_{\tilde{B}^{(k)}_h}$ and the product $\Delta^{o^{(1)}}_{\tilde{B}^{(k)}_h}\Delta^{o^{(3)}}_{\tilde{B}^{(k)}_h}$ are shown as functions of h.

Localization of the non-bandlimited/non-spacelimited Gaussian function. The minimum uncertainty state within the uncertainty relation is provided by the Gaussian probability density function (see [47, 59]). Consider the function G_λ given by

$$G_\lambda(t) = e^{-(\lambda/2)(1-t)}, \qquad t \in [-1, 1], \quad \lambda > 0. \tag{68}$$

An elementary calculation shows us that

$$\tilde{G}_\lambda(t) = \gamma(\lambda) e^{-(\lambda/2)(1-t)}, \tag{69}$$

with

$$\gamma(\lambda) = \left(1/\sqrt{4\pi}\right)\left(\frac{1}{2\lambda}\left(1 - e^{-2\lambda}\right)\right)^{-1/2}, \tag{70}$$

satisfies $\|\tilde{G}_\lambda\|_{\mathrm{L}^2[-1,1]} = 1$. It is not difficult to deduce that $\Delta^{o^{(1)}}_{\tilde{G}_\lambda}\Delta^{-\Delta^*}_{\tilde{G}_\lambda} \to 1$ as $\lambda \to \infty$: the best value of the uncertainty principle (Theorem 3) is 1.

Localization of Slepian functions. The bandlimited Slepian functions solve the concentration criterion (10) on general domains $C \subset \Omega$. If we restrict our attention to spherical caps as in Figures 1 and 4, the solutions degenerate and equations (11) and (12) can be solved for fixed spherical-harmonic orders j, with twice-repeated eigenvalues for the nonzonal functions at the same nonzero absolute orders.

While the Slepian functions do not formally optimize the uncertainty relation (38), calculations by Wieczorek and Simons [104] reveal that, again on spherical caps C of various opening half-angles Θ, the values attained by the largest-eigenvalue ($\lambda^{(\alpha)}_C \approx 1$ for $\alpha = 1$) zonal Slepian functions of varying bandwidths N are very close to satisfying the bounds (38) for Shannon numbers $N_0 = (N+1)\Theta/\pi$ (see (16)) greater than about 2. Furthermore, for increasing Shannon numbers, the uncertainty products for the αth best-concentrated Slepian function, when $N_0 \geq \alpha + 1$, tend to $2\alpha - 1$. This favorable behavior was illustrated by Wieczorek

and Simons [104], see their Figures 5 and 6b. Subsequent work by, among others, Guilloux et al. [52] and Khalid et al. [57], has substantiated and elaborated on these early analyses.

Slepian functions vs. the Gaussian. Another way by which the spatiospectral localization properties of the Slepian functions may be appreciated is by comparing how close they are to the family of minimum-uncertainty 'squeezed' *coherent states* (e.g., [15, 58]), a common root for many later developments in spline, Slepian function, and wavelet analysis [20]. This is of importance because in practical problems in the geosciences (e.g., [17, 84]), as in cosmology (e.g., [94]), we place as much value on the precise bandwidth, or bandwidth resolution, of our observations as on the spatial domain of interest. The Gaussian (68) may satisfy the uncertainty lower bound exactly, but it is not a bandlimited kernel. In contrast, the Slepian functions (11–12) can be bandlimited and spaceconcentrated at the same time. Formally, they are the optimizers of (10), though not of (38).

That they get close is shown in Figure 11. Inspired by Bluhm et al. [14] we determine the squeeze factor, s that renders the suitably normalized function

$$G_s(\cos\theta) = \gamma(s) e^{s\cos\theta}, \quad 0 \le \theta \le \pi, \tag{71}$$

as close as possible, in the mean-squared sense, to the best-concentrated bandlimited zonal Slepian function, concentrated to a spherical cap of a certain radius Θ, and whereby the tradeoff between spatial (the area of the spherical cap) and spectral concentration (the bandwidth N) is parameterized via the partial Shannon number $N_0 = (N+1)\Theta/\pi$.

3.3. Closing remarks

The uncertainty principle represents a trade-off between two 'spreads', in position and in frequency. *Sharp localization in space and in frequency are mutually exclusive.* The reason for the validity of the uncertainty relation (Theorem 1) is that the normal and curl operators $o^{(1)}$ and $o^{(3)}$ do not commute, hence, they cannot be sharply defined simultaneously. Extremal members of the uncertainty relation are polynomials (spherical harmonics) and Dirac function(al)s. An asymptotically optimal kernel is the Gaussian function.

Corollary 2 allows a quantitative classification and a hierarchy of the space and frequency localization properties of kernel functions of the form

$$K(t) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} K^{\wedge}(n) P_n(t), \qquad t = \xi \cdot \eta, \quad (\xi, \eta) \in \Omega \times \Omega. \tag{72}$$

In view of their space/frequency localization, it is also important to distinguish bandlimited kernels (i.e., $K^{\wedge}(n) = 0$ for all $n \ge N \in \mathbb{N}_0$) and non-bandlimited ones ($K^{\wedge}(n) \ne 0$ for an infinite number of integers n). Non-bandlimited kernels show a much stronger space localization than their bandlimited counterparts. It is

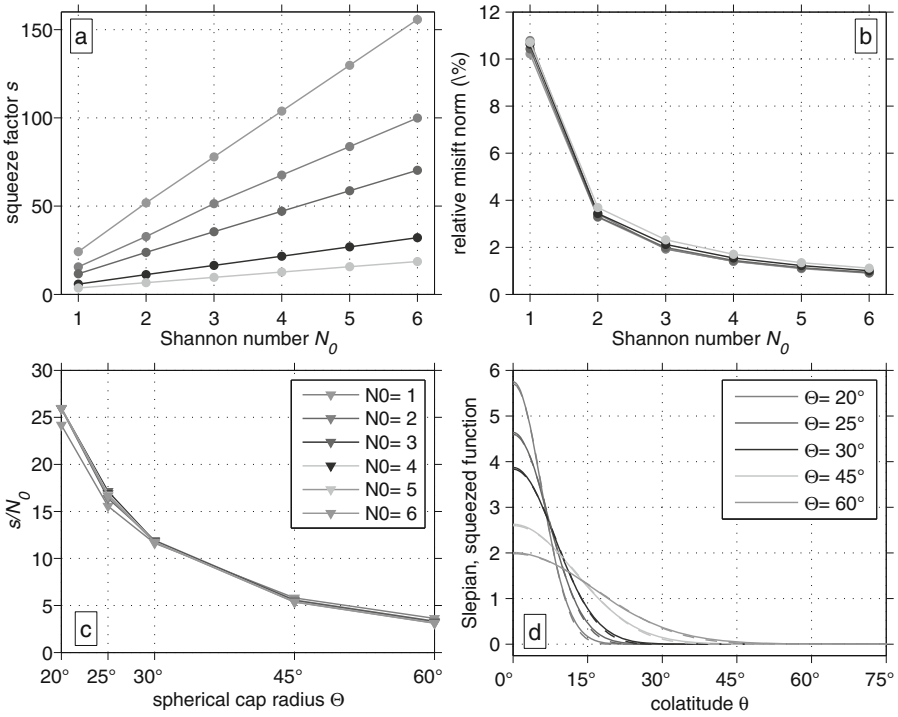


FIGURE 11. Slepian functions compared to Gaussian 'squeezed coherent states'. The chosen squeeze factors render functions of the type (71) as close as possible to the best-concentrated zonal Slepian functions in the relative mean-squared sense, for a variety of spherical cap sizes Θ and Slepian-function bandwidths N, linked through the Shannon number N_0.

not difficult to prove that, if $K \in \mathrm{L}^2[-1,1]$ with $\|K(\xi\cdot\,)\|_{\mathrm{L}^2(\Omega)} = 1$,

$$\left(\sigma^{o\,(1)}_{K(\xi\cdot\,)}\right)^2 = 1 - \left(\sum_{n=1}^{\infty} \frac{2n+1}{4\pi}\, K^\wedge(n) K^\wedge(n+1)\right)^2. \tag{73}$$

If $K^\wedge(n) \approx K^\wedge(n+1) \approx 1$ for many successive integers n, the space-domain support of $K(t)$ in (72) is small.

Space/frequency localization on the sphere can also be illustrated directly from (72). Choosing $K^\wedge(n) = \delta_{nk}$ we obtain a Legendre kernel of degree k, on the left in our scheme (Table 6). Setting $K^\wedge(n) = 1$ for $n = 0, 1, \ldots$, we obtain the Dirac kernel. The slower the sequence $\{K^\wedge(n)\}_{n=0,1,\ldots}$ converges to zero, the lower the frequency localization, but the higher the space localization.

Altogether, Table 6 gives a qualitative illustration of the consequences of the uncertainty principle in the theory of zonal kernel functions on the sphere: on the

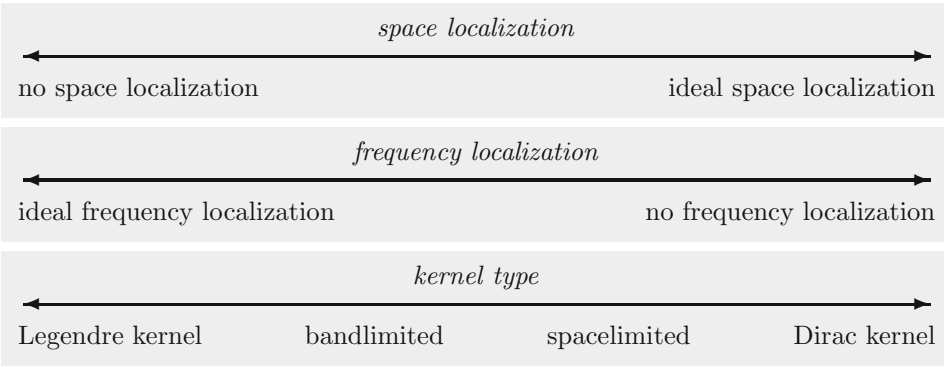


TABLE 6. The uncertainty principle and its consequences.

left end of this scheme, we have the Legendre kernels with their ideal frequency (momentum) localization. However, they show no space localization, as they are of polynomial nature. Thus, the present standard way in applications of increasing the accuracy in spherical harmonic (Fourier) expansions is to increase the maximum degree of the spherical harmonics expansions under consideration. On the right end of the scheme, there is the Dirac kernel which maps a function to its value at a certain point. Hence, this (generalized) function has an ideal space localization but no frequency localization. Consequently, it can be used in a finite pointset approximation.

4. Constructive approximation on the sphere

In Section 4.1, we discuss an approach using Slepian functions, Section 4.2 is an approach based on splines, and Section 4.3 treats the case of wavelets. Section 4.4 helps combine benefits of various approaches.

4.1. Approximation by Slepian functions

Given a certain region of interest C on the unit sphere Ω and a certain bandwidth N (a limiting spherical-harmonic degree in the sense of (9)), optimization of a *concentration criterion* yields linear combinations of spherical harmonics that we call *Slepian functions*. In Section 2.3, we gave their formulation in terms of band*limited* functions that are space*concentrated*. We shall denote these functions from now on as $G_N^C(\xi)$. Of course, we can equally well ask for space*limited* functions that are band*concentrated* – see [84, 87] for details. We shall denote those functions from now on as $H_C^N(\xi)$. The Fourier coefficients of the H_C^N can be calculated from those of the G_N^C by extension as in (13). We refer to [60] for an extensive discussion on the properties of what are, essentially, cases *intermediate* between these two endmembers, for functions defined on the real line.

If we introduce the space of all square-integrable scalar spherical functions that are spacelimited to the region C as $\mathcal{S}_C$, and the space of all square-integrable spherical functions that are bandlimited to the spherical-harmonic degree N as $\mathcal{S}_N$, then it is implied that $H_C^N \in \mathcal{S}_C$ and $G_N^C \in \mathcal{S}_N$.

Reproducing properties. We can show that the spectral-domain kernel that we first encountered in bandlimited form in (11), and which we now extend to $0 \leq m, n < \infty$,

$$d^C_{(m,j),(n,k)} = \langle Y_{m,j}, Y_{n,k} \rangle_{\mathrm{L}^2(C)} = \int_C Y_{m,j}(\xi) Y_{n,k}(\xi) \, \mathrm{d}\omega(\xi), \tag{74}$$

is a reproducing kernel in the space $\mathcal{S}_C$. Indeed, for any function $F \in \mathcal{S}_C$,

$$\begin{aligned} \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} d^C_{(m,j),(n,k)} F^{\wedge}(n,k) &= \int_C Y_{m,j}(\xi) \left(\sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} Y_{n,k}(\xi) F^{\wedge}(n,k) \right) \mathrm{d}\omega(\xi) \\ &= \int_C Y_{m,j}(\xi) F(\xi) \, \mathrm{d}\omega(\xi) \\ &= F^{\wedge}(m,j). \end{aligned} \tag{75}$$

At the same time, the spatial-domain Shannon kernel that we encountered in (12), and which we rebaptize

$$D_N(\xi, \eta) = \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta) = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} Y_{n,k}(\xi) Y_{n,k}(\eta), \tag{76}$$

is a reproducing kernel in the space $\mathcal{S}_N$, since, for any function $F \in \mathcal{S}_N$,

$$\begin{aligned} &\int_{\Omega} \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta) F(\eta) \, \mathrm{d}\omega(\eta) \\ &\quad = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} Y_{n,k}(\xi) \int_{\Omega} Y_{n,k}(\eta) F(\eta) \, \mathrm{d}\omega(\eta) \\ &\quad = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} Y_{n,k}(\xi) F^{\wedge}(n,k) \\ &\quad = F(\xi). \end{aligned} \tag{77}$$

Equations (75) and (77) hold the key to the approximation properties of the Slepian functions, since they imply that the spacelimited Slepian functions H_C^N provide a complete basis for all spacelimited functions in $\mathcal{S}_C \subset \mathrm{L}^2(\Omega)$, whereas the bandlimited Slepian functions G_N^C are a complete basis for all bandlimited functions in $\mathcal{S}_N \subset \mathrm{L}^2(\Omega)$.

Mercer's theorem. A second set of properties that solidifies these notions is established through an identity known as *Mercer's theorem*, which in this context takes the form

$$D_N(\xi,\eta) = \sum_{\alpha=1}^{(N+1)^2} \left(G_N^C\right)_\alpha(\xi)\left(G_N^C\right)_\alpha(\eta), \tag{78}$$

for all the α-indexed bandlimited Slepian functions G_N^C, with eigenvalues $\lambda_C^{(\alpha)}$, from which we establish, using (76) and as in (14d), that

$$D_N(\xi,\xi) = \sum_{\alpha=1}^{(N+1)^2} \left(G_N^C\right)_\alpha^2(\xi) = \frac{(N+1)^2}{4\pi} = \frac{N_C}{\int_C \mathrm{d}\omega(\eta)}, \tag{79}$$

recovering the spherical Shannon number N_C and the area of the domain of interest, $\int_C \mathrm{d}\omega(\eta)$.

A useful corollary is that the eigenvalue-weighted sum of squares of the bandlimited Slepian eigenfunctions closely *approximates* the value $N_C/\int_C \mathrm{d}\omega(\eta)$ when $\xi \in C$, and vanishes otherwise,

$$\sum_{\alpha=1}^{(N+1)^2} \lambda_C^{(\alpha)}\left(G_N^C\right)_\alpha^2 \approx \sum_{\alpha=1}^{N_C} \lambda_C^{(\alpha)}\left(G_N^C\right)_\alpha^2 \approx \begin{cases} N_C/\int_C \mathrm{d}\omega(\eta) & \text{if } \xi \in C \\ 0 & \text{otherwise,} \end{cases}$$

which is a consequence of the step-shaped eigenvalue spectrum that we saw in Figure 1. Eq. (80) testifies to the fact that the *effective* dimension of the space $\mathcal{S}_N$ of bandlimited functions that are *also* spaceconcentrated to C, is reduced from the canonical $(N+1)^2$ to the Shannon number N_C. It is our first clue to the approximation qualities of the Slepian functions, e.g., for (linear) signal estimation from regionally available data [85].

Power spectrum. If we furthermore define the power spectrum or degree variance of the bandlimited Slepian functions as

$$\mathscr{P}_n^{(\alpha)} = \frac{1}{2n+1}\sum_{k=1}^{2n+1}\left|\left(G_N^C\right)_\alpha^\wedge(n,k)\right|^2 \tag{80}$$

we get, via the spectral theorem, equation (74), and the addition theorem a spectral-domain equation equivalent to (79)–(80), namely,

$$\begin{aligned}\sum_{\alpha=1}^{(N+1)^2} \lambda_C^{(\alpha)}\mathscr{P}_n^{(\alpha)} &= \frac{1}{2n+1}\sum_{k=1}^{2n+1} d^C_{(n,k),(n,k)} = \frac{\int_C \mathrm{d}\omega(\eta)}{4\pi} \\ &\approx \sum_{\alpha=1}^{N_C} \lambda_C^{(\alpha)}\mathscr{P}_n^{(\alpha)},\end{aligned} \tag{81}$$

which is suggestive of the *spectral*-domain approximation properties of the Slepian functions, as arises, e.g., in the theory of (quadratic) power-spectral estimation from regionally available data [17].

Equations (80) and (81) together, show that the set of $N_C < (N+1)^2$ Slepian functions provide essentially uniform coverage over the spatial domain C and spectral bandwidth N. This is of interest when estimating (interpolating, approximating) functions from observations, as is common to a large number of research fields, not limited to the geosciences.

Alternative Mercer theorem. We note for completeness, and since the relevant identities have not been published before, that an alternative version of Mercer's theoremwould have transformed (12) and (76) from

$$\int_C D_N(\xi,\eta)F(\eta)\,\mathrm{d}\omega(\eta) = \lambda_C(F)F(\xi) \tag{82}$$

into the full-domain

$$\int_\Omega D^N(\xi,\eta)F(\eta)\,\mathrm{d}\omega(\eta) = \lambda_C(F)F(\xi), \tag{83}$$

which have the same eigenfunctions, but where we have defined

$$D^N(\xi,\eta) = \sum_{m=0}^{N}\sum_{j=1}^{2m+1}\sum_{n=0}^{N}\sum_{k=1}^{2n+1} d^C_{(m,j),(n,k)}Y_{m,j}(\xi)Y_{n,k}(\eta). \tag{84}$$

In that case, the equivalent to (78) is the to some more familiar expression

$$D^N(\xi,\eta) = \sum_{\alpha=1}^{(N+1)^2} \lambda_C^{(\alpha)} \left(G_N^C\right)_\alpha(\xi)\left(G_N^C\right)_\alpha(\eta). \tag{85}$$

Approximation 0: Noiseless data (interpolation). Imagine a certain function is 'known' as a spherical-harmonic expansion. Clearly, considering such a situation is merely postponing the problem of how to estimate an *unknown* function from observations. However, it is a common occurrence in the geosciences that, for example, space agencies perform exhaustive satellite data reductions that end up in the official release of spherical harmonic 'models' (typically of gravity or magnetic fields) that are then available for further research [103]. Another situation is where spectral forward-modeling codes deliver 'simulations' that are subsequently in need of interpretation and evaluation [102].

Whatever the source, and however large the bandlimit, the key property of the Slepian function basis is that the function expansion coefficients can be obtained by a simple transformation. If indeed the known function is F, then it is immaterial whether it is expressed in the spherical-harmonic basis, or in a bandlimited Slepian basis designed for *whichever* region C of interest, *as long as its bandwidth N matches the original*:

$$F = \sum_{n=0}^{N}\sum_{k=1}^{2n+1} F^\wedge(n,k)Y_{n,k} = \sum_{\alpha=1}^{(N+1)^2} F^\wedge(\alpha)\left(G_N^C\right)_\alpha. \tag{86}$$

The Slepian-function expansion coefficients $F^\wedge(\alpha)$, $\alpha = 1,\ldots,(N+1)^2$ are simply obtained from the spherical-harmonic expansion coefficients $F^\wedge(n,k)$,

$n = 0, 1, \ldots, N$ and order indices $k = 1, \ldots, 2n+1$, by the (orthogonal) transformation [85]

$$F^{\wedge}(\alpha) = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} \left(G_N^C\right)_{\alpha}^{\wedge}(n,k) F^{\wedge}(n,k). \tag{87}$$

A linear basis transformation (87) is exact and thus, strictly speaking 'uninteresting'. However, the properties of the Slepian functions designed for a region C are such that after a *partial* Slepian expansion to $J < (N+1)^2$ terms, denoted $F_J(\xi)$, equation (86) will hold *approximately* in the region of interest:

$$F(\xi) = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} F^{\wedge}(n,k) Y_{n,k}(\xi) \approx \sum_{\alpha=1}^{J} F^{\wedge}(\alpha) \left(G_N^C\right)_{\alpha}(\xi), \quad \xi \in C. \tag{88}$$

Clearly, a truncation of the spherical-harmonic series to its first J terms, however ordered, would generally result in poor approximations, *precisely* because of the non-localized spatial behavior of the basis functions. The eigenvalue-ranked Slepian transformation (87), on the other hand, has reordered the basis such that its first J functions increasingly uniformly 'cover' the spatial region of interest while providing an increasingly complete coverage over the entire spectral band, see (80) and (81). As a measure of approximation quality we take the area-weighted relative mean-squared error. It can be easily shown to depend on the truncation level in the manner

$$\frac{\|F - F_J\|^2_{\mathrm{L}^2(C)}}{\|F\|^2_{\mathrm{L}^2(C)}} = \frac{\sum_{\alpha>J}^{(N+1)^2} \left(F^{\wedge}(\alpha)\right)^2 \lambda_C^{(\alpha)}}{\sum_{\alpha=0}^{(N+1)^2} \left(F^{\wedge}(\alpha)\right)^2 \lambda_C^{(\alpha)}}. \tag{89}$$

Given the universally favorable decay of the eigenvalue spectrum of the spatio-spectral concentration problem (11), in this noiseless case, the Shannon number N_C is an obvious practical first choice for the truncation level J, although (89) of course shows the role played by the spectrum of the signal itself. An illustrative numerical example is given by Simons et al. [88], their Figure 3.

Approximation 1: Noisy data. We finally turn to the approximation problem that is most familiar in geophysical inverse theory, namely that of the *estimation* of a certain unknown signal from noisily observed data. We will briefly discuss the traditional spherical-harmonics based approach, and then clarify the beneficial role that localized basis functions (here: *Slepian functions*) may play in this context. We adhere to the continuous viewpoint for notational convenience and to lay bare the structure of the solutions. In practice, all datasets will be sample values at discrete geographic locations. As a consequence, the properties derived for constructive approximation by Slepian functions will themselves hold only approximately – to the degree by which continuous integrals are (hopefully, well) approximated by their Riemann sums [17, 84].

However, therein lies the crux of the Slepian-function method: if the data are regionally (in some region C) and densely (warranting a certain 'Nyquist' bandlimit at spherical-harmonic degree N) available, computing the Slepian basis for the idealized acquisition geometry *ahead of time* is what will lead to manageably sized inverse problems (on the order of the Shannon number N_C, and $N_C \ll (N+1)^2$ when $|C| \ll |\Omega|$) that solve for the unknown signal from which we assume the data to have been sampled.

Such a viewpoint, in a sense, embodies a strict *geographical prior*, and is very different from the splines and wavelets that will be discussed in the remaining Sections 4.2 and 4.3. Indeed, in contrast to Slepian functions, splines and wavelets made from zonal kernel functions do not *strictly* select for *particular* regions of interest, although of course, when particular combinations of any of those constructions are sought by optimization, as they are in Section 4.4, *effectively*, they do. Simons et al. [89] discuss a hybrid situation termed 'Slepian trees', as well as an alternative spherical wavelet transform obtained via a simple 'cubed-sphere' mapping of the 'usual' separable Cartesian discrete wavelet transforms [20].

The most detailed and up-to-date discussion of approximation by Slepian functions (both scalar and vector-valued, and for geomathematics problems involving measurements made by satellites at altitude) is found in the works by Simons and Plattner [80, 81, 86]. From these references, we retain and present a few essential points.

Suppose that we have 'data', M, consisting of a superposition of 'signal', F, and 'noise', E. What is F? The measurements are only available over some closed region C of the unit sphere Ω, i.e.,

$$M(\xi) = \begin{cases} F(\xi) + E(\xi) & \text{if } \xi \in C \\ \text{unknown/undesired} & \text{if } \xi \in \Omega \setminus C. \end{cases} \tag{90}$$

We assume that both signal and noise can be represented via an infinite spherical harmonic expansion as in (4), and we furthermore assume that they are uncorrelated realizations of zero-mean Gaussian random processes. Paying no heed to the structure of the noise (i.e., without explicit prior information that could be weighted into the norms in the form of a noise covariance) we elect to seek solutions to the optimization problem that results in a regularized bandlimited (to N, which remains to be determined) *estimate* of the signal, $\hat{F}$, in the form of equation (9), and which solves

$$\|\hat{F} - M\|^2_{\mathrm{L}^2(C)} + \lambda \|\hat{F}\|^2_{\mathrm{L}^2(\Omega\setminus C)} = \text{minimum}, \tag{91}$$

where $\lambda \geq 0$ is a regularization (damping) parameter forcing the solution to vanish outside of the observation domain. In the following two paragraphs, we distinguish solutions $\hat{F}_N$ and $\hat{F}_J$, both bandlimited.

Approximation 2: Regularized spherical-harmonic expansions. Simons and Dahlen [85] give the Fourier coefficients that solve equation (91) as

$$\hat{F}_N^{\wedge}(m,j) = \sum_{n=0}^{N} \sum_{k=1}^{2n+1} \left(d^{C}_{(m,j),(n,k)} + \lambda\, d^{(\Omega\setminus C)}_{(m,j),(n,k)} \right)^{-1} \langle M, Y_{n,k} \rangle_{\mathrm{L}^2(C)} \,. \tag{92}$$

We note from equation (74) that $d^{C}_{(m,j),(n,k)} + d^{(\Omega\setminus C)}_{(m,j),(n,k)}$ is the identity matrix. Regularization is unavoidable: as we have seen, the eigenvalues of $d^{C}_{(m,j),(n,k)}$ trail off quickly to nearly zero, see Figure 1. Restricted-region data availability is the prime reason for our inverse problem to be *ill posed* – even if no downward continuation from satellite height is required and if no internal density distributions (in the case of gravimetry) are being sought.

How well are we doing when accepting (92) as our solution? Rewriting the inverse Slepian eigenvalues $\lambda_C^{(\alpha)}$ with the damping parameter λ as

$$\left(\lambda_C^{(\alpha)}\right)^{*}(\lambda) = \left[\lambda_C^{(\alpha)} + \lambda\left(1-\lambda_C^{(\alpha)}\right)\right]^{-1}, \tag{93}$$

[85] derive the regional relative mean-squared error, the *expected value* of the ratio of approximation-error to signal norms as

$$\frac{E\left\{\left\|\hat{F}_N - F\right\|^2_{\mathrm{L}^2(C)}\right\}}{E\left\{\|F\|^2_{\mathrm{L}^2(C)}\right\}} = \sum_{\alpha=1}^{(N+1)^2} \frac{\lambda_C^{(\alpha)}}{N_C} \left[\left(\lambda_C^{(\alpha)}\right)^{*}(\lambda)\right]^2 \left[\mathscr{R}^{-1}\lambda_C^{(\alpha)} + \lambda^2\left(1-\lambda_C^{(\alpha)}\right)^2\right]. \tag{94}$$

In the expression above, both signal and noise were assumed to be characterized by a white (flat) power spectrum (defined in (80)), and we introduced $\mathscr{R}$, the signal-to-noise ratio. Valid only for this admittedly idealized case, (94) nevertheless contains all the elements by which the quality of the approximation can be appreciated: the bandwidth N and the size and shape of the region C enter through the eigenvalues $\lambda_C^{(\alpha)}$ and the Shannon number N_C, and of course the dependence on the signal-to-noise ratio $\mathscr{R}$ and the damping parameter λ are important controlling factors. Minimization of the relative error norm provides an implicit criterion for the regularization parameter:

$$\lambda_{\mathrm{opt}} = \mathscr{R}^{-1} \frac{\sum_{\alpha=1}^{(N+1)^2} \left[\left(\lambda_C^{(\alpha)}\right)^{*}(\lambda)\right]^3 \left(\lambda_C^{(\alpha)}\right)^2 \left(1-\lambda_C^{(\alpha)}\right)}{\sum_{\alpha=1}^{(N+1)^2} \left[\left(\lambda_C^{(\alpha)}\right)^{*}(\lambda)\right]^3 \left(\lambda_C^{(\alpha)}\right)^2 \left(1-\lambda_C^{(\alpha)}\right)^2}. \tag{95}$$

At high signal-to-noise ratios, (95) is well approximated by $\lambda_{\mathrm{opt}} \approx \mathscr{R}^{-1}$.

Approximation 3: Truncated Slepian expansions. Where did the Slepian functions go? We solved (91) using spherical harmonics, but we discussed the statistics of the solution (92) in terms of the eigenvalues of the Slepian concentration problem. The link, of course, is that the spherical-harmonic solution is derived via the

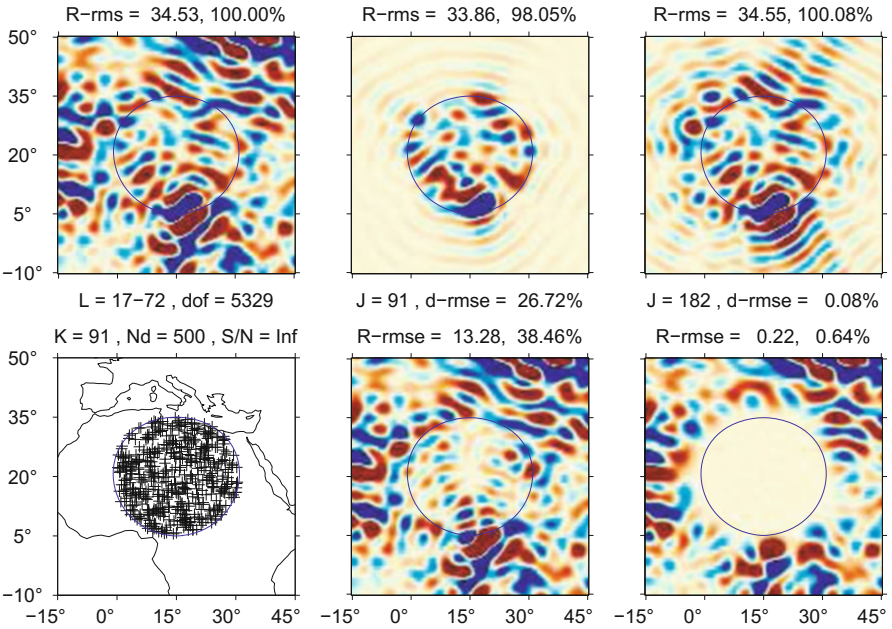


FIGURE 12. Example of Slepian-basis (Shannon number K = 91) approximation of a non-white bandlimited (bandpass, spherical-harmonic degrees L = 17 − 72) geomagnetic field from Nd = 500 noiseless data, for two truncation levels, J = 91 and J = 182, over a circular domain R. Top: the field, and the two reconstructions. Bottom: the location of the data points, and the difference between the truth and the approximation. The relative regional root-mean square signal, reconstruction and error strengths are indicated.

intermediary of the inverse of the Slepian localization matrix $d^{C}_{(m,j),(n,k)}$ and, with regularization, its complement, $d^{(\Omega\setminus C)}_{(m,j),(n,k)}$. Both of these are large, full (though banded) matrices whose inverses (especially at large spherical-harmonic degrees N) are computed at significant cost. We have previously seen how a partial set of Slepian functions provides excellent regional approximations in noiseless cases. To conclude this section, we thus propose an estimator for the situation of the form (90), where we attempt to reconstruct the unknown signal F from a regionally observed set of noisy measurements, M.

This time, our estimator does not take the form of a spherical-harmonic expansion that needs to be *regularized* (sometimes at great computational cost), but rather of a Slepian-function expansion which can be *truncated* (usually without any difficulty at all). In the context of equation (88): we prefer the approximate identity over the equality which may well furnish us with a 'complete' expansion, but whose coefficients we can only calculate approximately, after regularization.

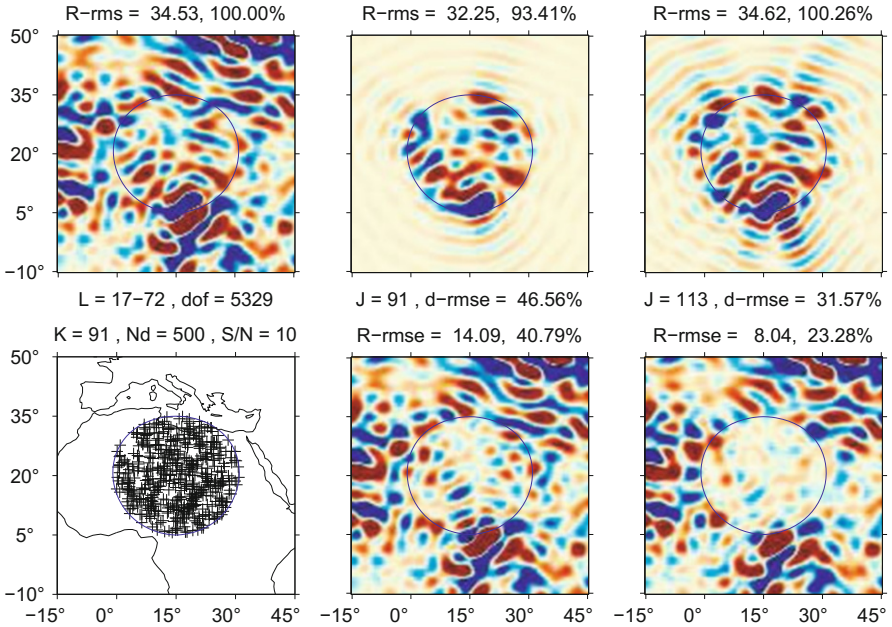


FIGURE 13. Example of Slepian-basis (Shannon number 91) approximation of a bandlimited field from data with a signal-to-noise ratio of 10, for two truncation levels, 91 and 113. Layout as in Figure 12.

In the framework of Slepian-function estimation, truncation *is* our regularization. The Slepian-basis solutions to the 'unregularized' ($\lambda = 0$) problem (91) are, quite simply,

$$\hat{F}^{\wedge}(\alpha) = \left(\lambda_C^{(\alpha)}\right)^{-1} \left\langle M, \left(G_N^C\right)_\alpha \right\rangle_{\mathrm{L}^2(C)} . \tag{96}$$

Truncation means that we only compute J of them, which gives us the freedom to avoid the blowup of the inverse eigenvalues, i.e., the estimate in the Slepian basis is given by

$$\hat{F}_J = \sum_{\alpha=1}^{J} \hat{F}^{\wedge}(\alpha) \left(G_N^C\right)_\alpha . \tag{97}$$

By the same metric of (94), we evaluate the quality of this solution as

$$\frac{E\left\{ \left\| \hat{F}_J - F \right\|^2_{\mathrm{L}^2(C)} \right\}}{E\left\{ \|F\|^2_{\mathrm{L}^2(C)} \right\}} = \mathscr{R}^{-1} \frac{J}{N_C} + \frac{1}{N_C} \sum_{\alpha>J}^{(N+1)^2} \lambda_C^{(\alpha)} . \tag{98}$$

As (94), but unlike (89), again (98) is only applicable in the case of white noise and white signal with a signal-to-noise power ratio $\mathscr{R}$. Of course, the signal contained in the neglected terms of what should be a complete Slepian expansion exerts

a controlling factor on the mean squared error behavior. [85] show how, in the Slepian basis, the neglected terms positively affect the variance of the estimate, but negatively the bias; the mean-squared error being the combination of the two. Minimization of (98) to determine the optimal truncation level for these circumstances yields it in terms of the Slepian eigenvalue and the signal-to-noise level, namely

$$\lambda_C^{(J_{\mathrm{opt}})} \approx \mathscr{R}^{-1}. \tag{99}$$

In other words, we include Slepian functions in the expansion until their ranked eigenvalues drop below the noise-to-signal ratio.

We have ignored that in order to 'solve' data-driven approximation problems, we need to determine an optimal bandwidth N *and* an optimal truncation level J for data situations that are more involved than just being given by white noise and white signal. Such vital practical matters are discussed by Slobbe et al. [91] and Plattner and Simons [81]. The solution procedures involved are always cumbersome – but the computational complexity, and the overall size of the numerical problem, of the truncated Slepian-function approach is always smaller than via regularized spherical-harmonics. Slepian functions lend themselves well to solving approximation problems involving noisy and partially observed data on the sphere.

Two realistic examples of truncated Slepian-basis approximation problems are given in Figures 12 and 13.

4.2. Approximation by splines

Only relatively recently have zonal kernel function techniques such as spline interpolation/approximation and wavelet analysis been playing a fundamental role in numerical analysis on the sphere. *Spherical splines* (independently introduced by Freeden [28] andWahba [99] in 1981) are canonical generalizations of 'spherical polynomials' (spherical harmonics) which have desirable characteristics as interpolating, smoothing, and best approximating functions (see also [100]). By spline interpolation we mean the variational problem of minimizing an 'energy'-norm of a suitable Sobolev space. Depending on the chosen norm, bandlimited and non-bandlimited splines are distinguished. Spherical splines have been successfully applied to many areas of application in particular in geodesy for gravitational field determination, radio occultation, ocean flow, etc. (for more details see [29, 36, 47] and the references therein). Spherical splines, especially their counterparts on the ball, have been applied to tomographic inverse problems in geophysics ([1–3, 10–12, 22, 74]) and in medical imaging ([27]).

To understand spherical splines, we adopt the idea of one-dimensional cubic splines to the sphere. Cubic splines in one-dimension are well known for having minimal 'bending energy' (roughly, minimal 'curvature energy' understood in a linearized sense). More concretely, among all interpolating functions of the Sobolev space $\mathrm{H}^{(2)}([a,b])$, the integral $\int_a^b |F''(x)|\,\mathrm{d}x$ becomes minimal, where F may be physically interpreted as the deflection normal to the rest position which is supposed to be horizontal. The physical model is suggested by the classical inter-

pretation of the potential energy of a statically deflected thin beam which indeed is proportional to the integral taken over the square of the linearized curvature of the elastic beam. Analogously, the concept can be applied to the sphere by choosing $\int_\Omega |\Delta_\xi^* F(\xi)|^2 \, \mathrm{d}\omega(\xi)$, where F now denotes the deflection of a thin membrane normal to the rest position supposed to be spherical. In other words, the second derivative canonically takes on the form of the Beltrami operator Δ^*. Indeed, our interest now is to state that the interpolating spline to a given dataset has minimum 'bending energy' for all interpolants within the Sobolev space $\mathrm{H}^{(2)}(\Omega)$. Furthermore, the spline functions defined in this section are able to simultaneously interpolate and smooth the data. Hence, we can decide in our spline application, which knots of the input data should be strictly interpolated and which ones should be 'near' the interpolating function, i.e., the points subjected to smoothing.

Reproducing kernel Hilbert reference space. As usual (see, e.g., [47]), we introduce the Sobolev space $\mathrm{H}^{(2)}(\Omega)$ as the completion of $\mathrm{C}^{(2)}(\Omega)$ with respect to a specific scalar product thereby specifying $\mathrm{H}^{(2)}(\Omega)$ as a certain reproducing kernel space. In more detail, the inner product $\langle \cdot, \cdot \rangle_{\mathrm{H}^{(2)}(\Omega)}$ is defined by

$$\begin{aligned}\langle F, G \rangle_{\mathrm{H}^{(2)}(\Omega)} = & \underbrace{\int_\Omega F(\eta) Y_{0,1}(\eta) \, \mathrm{d}\omega(\eta) \int_\Omega G(\eta) Y_{0,1}(\eta) \, \mathrm{d}\omega(\eta)}_{= \langle F, G \rangle_{\mathrm{H}_0}} \\ & + \underbrace{\sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} (n(n+1))^2 \int_\Omega F(\eta) Y_{n,j}(\eta) \, \mathrm{d}\omega(\eta) \int_\Omega G(\eta) Y_{n,j}(\eta) \, \mathrm{d}\omega(\eta)}_{= \langle F, G \rangle_{\mathrm{H}_0^\perp}}, \end{aligned} \tag{100}$$

which is equivalent in accordance with Parseval's identity to

$$\begin{aligned}\langle F, G \rangle_{\mathrm{H}^{(2)}(\Omega)} = & \underbrace{\int_\Omega F(\eta) Y_{0,1}(\eta) \, \mathrm{d}\omega(\eta) \int_\Omega G(\eta) Y_{0,1}(\eta) \, \mathrm{d}\omega(\eta)}_{= \langle F, G \rangle_{H_0}} \\ & + \underbrace{\int_\Omega (\Delta_\eta^* F(\eta))(\Delta_\eta^* G(\eta)) \, \mathrm{d}\omega(\eta)}_{= \langle F, G \rangle_{\mathrm{H}_0^\perp}} \\ = & \langle F, G \rangle_{\mathrm{H}_0} + \langle F, G \rangle_{\mathrm{H}_0^\perp} \end{aligned} \tag{101}$$

for all $F, G \in \mathrm{C}^{(2)}(\Omega)$. The Sobolev space $\mathrm{H}^{(2)}(\Omega)$ as defined in Section 3.1 is the completion of $\mathrm{C}^{(2)}(\Omega)$ under the norm $\| \cdot \|_{\mathrm{H}^{(2)}(\Omega)}$, i.e.,

$$\mathrm{H}^{(2)}(\Omega) = \overline{\mathrm{C}^{(2)}(\Omega)}^{\| \cdot \|_{\mathrm{H}^{(2)}(\Omega)}}, \tag{102}$$

where $\|F\|_{\mathrm{H}^{(2)}(\Omega)} = \sqrt{\langle F, F \rangle_{\mathrm{H}^{(2)}(\Omega)}}$.

Consider the kernel $K : (\xi, \eta) \mapsto K(\xi, \eta)$, $(\xi, \eta) \in \Omega^2 = \Omega \times \Omega$ given in the form

$$K(\xi, \eta) = Y_{0,1}(\xi) Y_{0,1}(\eta) + \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \frac{1}{(n(n+1))^2} Y_{n,j}(\xi) Y_{n,j}(\eta). \tag{103}$$

Then $K(\xi, \cdot)$, $\xi \in \Omega$ fixed, is a member of $\mathrm{H}^{(2)}(\Omega)$. Inserting K into the inner product, we see via the orthogonal properties of the spherical harmonics that

$$\langle F, K(\xi, \cdot) \rangle_{\mathrm{H}^{(2)}(\Omega)} = F(\xi), \qquad \xi \in \Omega, \tag{104}$$

for all $F \in \mathrm{H}^{(2)}(\Omega)$. Hence, $K(\cdot, \cdot)$ is the unique reproducing kernel of the Hilbert space $\mathrm{H}^{(2)}(\Omega)$. The reproducing kernel $K(\cdot, \cdot)$ can be decomposed into the reproducing kernels of the spaces H_0 and $\mathrm{H}_0^{\perp}$, respectively, via

$$K(\xi, \eta) = \underbrace{Y_{0,1}(\xi) Y_{0,1}(\eta)}_{= K_0(\xi, \eta) = \frac{1}{4\pi}} + \underbrace{\sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \frac{1}{(n(n+1))^2} Y_{n,j}(\xi) Y_{n,j}(\eta)}_{= K_0^{\perp}(\xi, \eta)}. \tag{105}$$

Applying the spherical-harmonic addition theorem and comparing with (105), we get

$$\begin{aligned} K_0^{\perp}(\xi, \eta) &= \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} \frac{1}{(n(n+1))^2} Y_{n,j}(\xi) Y_{n,j}(\eta) \\ &= \frac{1}{4\pi} \sum_{n=1}^{\infty} \frac{2n+1}{(n(n+1))^2} P_n(\xi \cdot \eta) \\ &= G((\Delta^*)^2; \xi, \eta) \end{aligned} \tag{106}$$

where $G((\Delta^*)^2, \cdot, \cdot)$ is the Green function with respect to the iterated Beltrami operator $(\Delta^*)^2 = \Delta^* \Delta^*$ (see [28]). Summarizing our results we, therefore, see that

$$K(\xi, \eta) = \underbrace{Y_{0,1}(\xi) Y_{0,1}(\eta)}_{= K_0(\xi, \eta) = \frac{1}{4\pi}} + \underbrace{G\left((\Delta^*)^2; \xi, \eta\right)}_{= K_0^{\perp}(\xi, \eta)}, \qquad \xi, \eta \in \Omega, \tag{107}$$

is the uniquely determined reprokernel of the space $(\mathrm{H}^{(2)}(\Omega), \langle \cdot, \cdot \rangle_{\mathrm{H}^{(2)}(\Omega)})$, i.e.:

(i) For each fixed $\xi \in \Omega$, $K(\xi, \eta)$, a function of η, is an element of $\mathrm{H}^{(2)}(\Omega)$.
(ii) For every function $F \in \mathrm{H}^{(2)}(\Omega)$ and for every point $\xi \in \Omega$, the reproducing property holds:

$$F(\xi) = \langle F, K(\xi, \cdot) \rangle_{\mathrm{H}^{(2)}(\Omega)}. \tag{108}$$

Explicit representation of the reproducing kernel. Keeping the reprostructure of $\mathrm{H}^{(2)}(\Omega)$ in mind, we are able to handle our announced spline interpolation and smoothing problem. To this end, we follow the concept presented by Freeden [28] and observe, in addition, the explicit representation of $G((\Delta^*)^2; \cdot, \cdot)$ known from

[41]. In fact, Green's function corresponding to the iterated Beltrami operator $(\Delta^*)^2$ is continuous on $\Omega \times \Omega$ and admits the explicit formulation:

$$G((\Delta^*)^2;\xi,\eta) = \begin{cases} \frac{1}{4\pi}, & 1-\xi\cdot\eta = 0 \\ \frac{1}{4\pi}(1-\ln(1-\xi\cdot\eta)(\ln(1+\xi\cdot\eta)-\ln(2)) & \\ -\mathfrak{L}_2(\frac{1-t}{2}) - (\ln(2))^2 + \ln(2)\ln(1+\xi\cdot\eta)), & 1\pm\xi\cdot\eta \neq 0 \\ \frac{1}{4\pi} - \frac{\pi}{24}, & 1+\xi\cdot\eta = 0, \end{cases} \tag{109}$$

where the function $\mathfrak{L}_2(x)$ is the dilogarithm given by

$$\mathfrak{L}_2(x) = -\int_0^x \frac{\ln(1-t)}{t}\,\mathrm{d}t = \sum_{k=1}^{\infty}\frac{x^k}{k^2}. \tag{110}$$

Spline concept. We have come to the definition of spherical splines corresponding to one-dimensional cubic splines (a more general concept involving pseudo-differential operators is known from [47]). Let $M_1,\ldots,M_n$ be a linearly independent system of bounded linear functionals on $\mathrm{H}^{(2)}(\Omega)$. Any function $S \in \mathrm{H}^{(2)}(\Omega)$ of the form

$$S(\eta) = c_0 Y_{0,1}(\eta) + \sum_{i=1}^{n} a_i M_i G((\Delta^*)^2;\eta,\cdot), \quad a_i \in \mathbb{R}, \quad \eta \in \Omega \tag{111}$$

with

$$\sum_{i=1}^{n} a_i M_i(Y_{0,1}) = 0 \tag{112}$$

is called a *spherical spline* in $\mathrm{H}^{(2)}(\Omega)$ relative to $M_1,\ldots,M_n$.

By virtue of (111) and (112), we are now prepared to formulate the following uniqueness result: let $M_1,\ldots,M_n$ be a system of bounded linear functionals on the Sobolev space $\mathrm{H}^{(2)}(\Omega)$ such that the $((n+1)\times(n+1))$-matrix

$$\begin{pmatrix} \alpha & \kappa \\ \kappa^{\mathrm{T}} & 0 \end{pmatrix} \tag{113}$$

is non-singular, where the matrix α and the vector κ are given as follows:

$$\alpha = \left(M_i M_j G((\Delta^*)^2;\cdot,\cdot)\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}, \tag{114}$$

$$\kappa = \left(M_i Y_{0,1}(\cdot)\right)_{i=1,\ldots,n}. \tag{115}$$

Then, there exists a unique spline in $\mathrm{H}^{(2)}(\Omega)$ relative to $M_1,\ldots,M_n$ that solves the interpolation problem $M_i S = \mu_i$, $i=1,\ldots,n$. This spline is called the *interpolating spline*. The proof easily follows by inserting the representation (111) into $M_i S = \mu_i$, $i=1,\ldots,n$, resulting in a linear system for the coefficients a_i, c_0, whose coefficient matrix is given by (113). Since the matrix is assumed to be non-singular, the coefficients are uniquely determined.

The key to spline approximation is the so-called *spline integration formula*

$$\int_\Omega \Delta_\eta^* S(\eta)\ \Delta_\eta^* F(\eta)\,\mathrm{d}\omega(\eta) = \sum_{k=1}^{n} a_k M_k F, \tag{116}$$

valid for the uniquely determined interpolating spline S and all members $F \in \mathrm{H}^{(2)}(\Omega)$, provided that the constraint $\kappa^{\mathrm{T}} a = 0$ is fulfilled. The proof is a direct conclusion of the reproducing kernel property. Its idea is to inspect the sum $\sum_{k=1}^{n} a_k M_k F$ and substitute F by the reproducing kernel property

$$\sum_{k=1}^{n} a_k M_k F = \sum_{k=1}^{n} a_k M_k \langle F(\cdot), K(\eta,\cdot)\rangle_{\mathrm{H}^{(2)}(\Omega)}. \tag{117}$$

Evaluating the inner product by inserting the reproducing kernel function leads to the desired result.

Next, we turn to dealing with the *'minimum energy property' of strict spline interpolation.*

Theorem 4. *Let $M_1, \ldots, M_n$ be a linearly independent system of bounded linear functionals on $\mathrm{H}^{(2)}(\Omega)$. Let S be the unique spline which solves the interpolation problem $M_i S = \mu_i$, $i = 1, \ldots, n$. Then, for all twice continuously differentiable functions F on Ω, which interpolate the given data, i.e., $M_i F = \mu_i$, $i = 1, \ldots, n$, the following inequality*

$$\int_\Omega (\Delta_\eta^* S(\eta))^2\,\mathrm{d}\omega(\eta) \le \int_\Omega (\Delta_\eta^* F(\eta))^2\,\mathrm{d}\omega(\eta) \tag{118}$$

holds true with equality if and only if $S = F$.

The proof easily follows from arguments given by Freeden [28]. Theorem 4 tells us that the 'bending energy' (the integral over the second derivative) of the spline is minimal among all functions in $\mathrm{H}^{(2)}(\Omega)$ interpolating the data.

Combined spline interpolation and smoothing. Theorem 4 allows an extension to include smoothing at predefined points while interpolating the remaining pointset (in accordance with [46]). This technique was used by Blick and Freeden [13] to visualize radio occultation data collected by the satellite CHAMP.

Given $n = p + q$ data points, where the data points μ_i, $i = 1, \ldots, p$, are subjected to smoothing and the points ν_i, $i = 1, \ldots, q$, are subjected to strict interpolation, we are lead to the following result.

Theorem 5. *Suppose that δ and $\beta_1^2, \ldots, \beta_p^2$ are prescribed positive weights and that μ_i, $i = 1, \ldots, p$; ν_j, $j = 1, \ldots, q$ are given data points. Let $M_1, \ldots, M_p$ and $N_1, \ldots, N_q$ be systems of bounded linear functionals on $\mathrm{H}^{(2)}(\Omega)$ such that the $((p+q)+1) \times ((p+q)+1)$-matrix*

$$\begin{pmatrix} \alpha & \beta & \kappa \\ \beta^{\mathrm{T}} & \gamma & \zeta \\ \kappa^{\mathrm{T}} & \zeta^{\mathrm{T}} & 0 \end{pmatrix} \tag{119}$$

is non-singular, where the matrices $\alpha,\ \beta,\ \gamma,\ \kappa,\ \zeta$ *are given as follows*

$$\alpha = \left(M_i M_j G\left(\left(\Delta^*\right)^2;\cdot,\cdot\right) + \delta\beta_i^2\delta_{ij}\right)_{\substack{i=1,\ldots,p\\ j=1,\ldots,p}}, (\textit{Kronecker } \delta_{ij}) \tag{120}$$

$$\beta = \left(M_i N_j G\left(\left(\Delta^*\right)^2;\cdot,\cdot\right)\right)_{\substack{i=1,\ldots,p\\ j=1,\ldots,q}}, \tag{121}$$

$$\gamma = \left(N_i N_j G\left(\left(\Delta^*\right)^2;\cdot,\cdot\right)\right)_{\substack{i=1,\ldots,q\\ j=1,\ldots,q}}, \tag{122}$$

$$\kappa = \left(M_i Y_{0,1}(\cdot)\right)_{i=1,\ldots,p}, \tag{123}$$

$$\zeta = \left(N_j Y_{0,1}(\cdot)\right)_{j=1,\ldots,q}. \tag{124}$$

Then the smoothing spline function S *of the form*

$$S(\zeta) = c_0 Y_{0,1}(\xi) + \sum_{i=1}^{p} a_i M_i G\left(\left(\Delta^*\right)^2;\xi,\cdot\right) + \sum_{j=1}^{q} b_j N_j G\left(\left(\Delta^*\right)^2;\xi,\cdot\right), \xi\in\Omega, \tag{125}$$

with coefficients $a\in\mathbb{R}^p$, $a^{\mathrm{T}} = (a_1,\ldots,a_p)$; $b\in\mathbb{R}^q$, $b^{\mathrm{T}} = (b_1,\ldots,b_q)$ *and* $c_0\in\mathbb{R}$ *subjected to the constraint*

$$\sum_{i=1}^{p} a_i M_i(Y_{0,1}) + \sum_{j=1}^{q} b_j N_j(Y_{0,1}) = 0 \tag{126}$$

is the unique solution of the interpolation and smoothing problem given by

$$\begin{aligned} M_i S + \delta\beta_i^2 a_i &= \mu_i, \quad i = 1,\ldots,p,\\ N_j S &= \nu_j, \quad j = 1,\ldots,q, \end{aligned}$$

corresponding to the data points μ_i, $i = 1,\ldots,p$; ν_j, $j = 1,\ldots,q$ *and represents the only element of* $\mathrm{H}^{(2)}(\Omega)$ *satisfying*

$$\sum_{i=1}^{p}\left(\frac{M_i S - \mu_i}{\beta}\right)^2 + \delta\langle S,S\rangle_{H_0^\perp} \le \sum_{i=1}^{p}\left(\frac{M_i F - \mu_i}{\beta_i}\right)^2 + \delta\langle F,F\rangle_{\mathrm{H}_0^\perp} \tag{127}$$

for all $F\in\mathrm{H}^{(2)}(\Omega)$ *with* $N_j F = \nu_j,\ j = 1,\ldots,q$.

As already mentioned, the proof can be given in parallel to the arguments stated by Freeden and Witte [46]. Moreover, Theorem 4 leads us to the following comments:

(i) The values $\mu_1,\ldots,\mu_p,\ \nu_1,\ldots,\nu_q$ are regarded as the observed quantities, e.g., geodetic observations and measurements.

(ii) The spline function $S\in\mathrm{H}^{(2)}(\Omega)$ satisfies that $M_i S$ is 'near' μ_i, $i = 1,\ldots,p$ and $N_j S$ is equal to ν_j, $j = 1,\ldots,q$. The 'nearness' of the values $M_i S$ to μ_i, $i = 1,\ldots,p$ can be controlled by choosing the constant δ in a suitable way. A small value of δ emphasizes fidelity to the observed data at the expense of smoothness, while a large value does the opposite.

(iii) Taking $\delta = 0$ yields $M_i S = \mu_i$, $i = 1,\ldots,p$, i.e., the combined smoothing and interpolation procedure leads back to strict interpolation.

(iv) For numerical purposes, it is advantageous to adapt the quantities $\beta_1^2,\ldots,\beta_p^2$ to the standard deviations of the measured values.

4.3. Approximation by wavelets

As already pointed out, the context of the spectral representation of a square-integrable function by means of spherical harmonics is essential to solving many problems in today's applications. In future research, however, orthogonal (Fourier) expansions in terms of spherical harmonics $\{Y_{n,j}\}$ will not be the only way of representing a square-integrable function. In order to explain this in more detail, we think of a square-integrable function as a signal in which the spectrum evolves over space in significant way. We imagine that, at each point on the sphere Ω, the function refers to a certain combination of frequencies, and that these frequencies are continuously changing. This space-evolution of the frequencies, however, is not reflected in the Fourier expansion in terms of non-space localizing spherical harmonics, at least not directly. Therefore, in theory, any member F of the space $\mathrm{L}^2(\Omega)$ can be reconstructed from its Fourier transforms, i.e., the 'amplitude spectrum' $\{F^\wedge(n,j)\}_{\substack{n=0,1,\dots,\\ j=1,\dots,2n+1}}$, but the Fourier transform contains information about the frequencies of the function over all positions instead of showing how the frequencies vary in space.

Dirac families. In what follows, we present a two-parameter, i.e., scale- and space-dependent method of achieving a reconstruction of a function $F \in \mathrm{L}^2(\Omega)$ involving (scalar) zonal kernel functions which we refer to as a *Dirac family* $\{\Phi_\rho\}_{\rho\in(0,\infty)}$ converging to the (zonal) Dirac kernel δ. In other words, a Dirac family is a set of zonal kernels $\Phi_\rho : [-1,1] \to \mathbb{R}$, $\rho \in (0,\infty)$, of the form

$$\Phi_\rho(\xi\cdot\eta) = \sum_{n=0}^{\infty} \Phi_\rho^\wedge(n)\,\frac{2n+1}{4\pi}\,P_n(\xi\cdot\eta), \quad \xi,\eta\in\Omega, \tag{128}$$

converging to the 'Dirac kernel' δ as $\rho\to 0$, $\rho>0$. Consequently, if $\{\Phi_\rho\}_{\rho\in(0,\infty)}$ is a Dirac family, its 'symbol' $\{\Phi_\rho^\wedge(n))\}_{n=0,1,\dots}$ constitutes a sequence satisfying the limit relation

$$\lim_{\rho\to 0,\,\rho>0} \Phi_\rho^\wedge(n) = 1, \quad n = 0,1,\dots\ . \tag{129}$$

Accordingly, if $\{\Phi_\rho\}_{\rho\in(0,\infty)}$ is a scaling kernel function, the convolution integrals

$$(\Phi_\rho * F)(\xi) = \int_\Omega \Phi_\rho(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta), \quad \xi\in\Omega, \tag{130}$$

converge (in a certain topology) to the limit

$$F(\xi) = (\delta * F)(\xi) = \int_\Omega \delta(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta), \quad \xi\in\Omega, \tag{131}$$

for all $\xi\in\Omega$ as ρ tends to 0 (from the positive side). In more detail, if F is a function of class $\mathrm{L}^2(\Omega)$ and $\{\Phi_\rho\}$ is a (suitable) Dirac family (tending to the Dirac kernel), then the following limit relation holds true:

$$\lim_{\rho\to 0,\,\rho>0} \|F - \Phi_\rho * F\|_{\mathrm{L}^2(\Omega)} = 0. \tag{132}$$

There is a large number of Dirac families that is of interest for geoscientific application (for more details, the reader is referred to, e.g., [39, 44] and the

references therein). Only three prototypes of Dirac families should be mentioned here: the bandlimited Shannon family, the neither bandlimited nor spacelimited Abel–Poisson and Gauss–Weierstraß families, and the spacelimited Haar family.

It should be noted that an approximate convolution identity (132) acts as a space and frequency localization procedure in the following way. As $\{\Phi_\rho\}_{\rho\in(0,\infty)}$ is a Dirac family of zonal scalar kernel functions tending to the Dirac kernel, the function $\Phi_\rho(\eta\cdot)$, is highly concentrated around the point $\eta \in \Omega$, if the 'scale parameter' is a small positive value. Moreover, as ρ tends to infinity, $\Phi_\rho(\eta\cdot)$ becomes more and more localized in frequency. Correspondingly, the uncertainty principle states that the space localization of $\Phi_\rho(\eta\cdot)$ becomes more and more decreasing. In conclusion, the products $\eta \mapsto \Phi_\rho(\xi\cdot\eta)F(\eta)$, $\eta \in \Omega$, $\xi \in \Omega$, for each fixed value ρ, display information in $F \in \mathrm{L}^2(\Omega)$ at various levels of spatial resolution or frequency bands. Consequently, as ρ approaches ∞, the convolution integrals $\Phi_\rho * F = \int_\Omega \Phi_\rho(\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta)$ display coarser, lower-frequency features. As ρ approaches 0, the integrals give sharper and sharper spatial resolution. Thus, the convolution integrals can measure the space-frequency variations of spectral components, but they have a different space-frequency resolution.

Scaling and wavelet functions. Next we come to the bilinear theory of scaling and wavelet functions (note that we only deal with the bilinear theory, for basic aspects of the linear case the reader is referred to, e.g., [39, 44]).

The point of departure for our multi-scale approach is a particular type of a Dirac family: a *scaling* (*kernel*) *function* $\{\Phi_\rho^{(2)}\}_{\rho\in(0,\infty)}$ is a set of zonal kernels $\Phi_\rho^{(2)} = \Phi_\rho * \Phi_\rho : [-1,1] \to \mathbb{R}$, $\rho \in (0,\infty)$, of the form

$$\Phi_\rho^{(2)}(\xi\cdot\eta) = \sum_{n=0}^{\infty} \Phi_\rho^{(2)\wedge}(n)\,\frac{2n+1}{4\pi}\,P_n(\xi\cdot\eta), \quad \xi,\eta\in\Omega, \tag{133}$$

with

$$\lim_{\rho\to 0,\,\rho>0} \Phi_\rho^{(2)\wedge}(n) = \lim_{\rho\to 0,\,\rho>0} (\Phi_\rho^\wedge(n))^2 = 1, \quad n = 0,1,\dots\ . \tag{134}$$

and

$$\Phi_\rho^{(2)\wedge}(0) = 1. \tag{135}$$

Accordingly, the convolution integrals

$$\left(\Phi_\rho^{(2)} * F\right)(\xi) = \int_\Omega \Phi_\rho^{(2)}(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta), \quad \xi\in\Omega, \tag{136}$$

converge (in a certain topology) to the limit

$$F(\xi) = (\delta * F)(\xi) = \int_\Omega \delta(\xi\cdot\eta)F(\eta)\,\mathrm{d}\omega(\eta), \quad \xi\in\Omega, \tag{137}$$

for all $\xi \in \Omega$ as ρ tends to 0 (from the positive side). In other words, if F is a function of class $\mathrm{L}^2(\Omega)$ and $\{\Phi_\rho^{(2)}\}$ is a certain Dirac family (tending to the Dirac

kernel), then the approximate identity

$$\lim_{\rho\to 0,\,\rho>0} \left\| F - \Phi_\rho^{(2)} * F \right\|_{\mathrm{L}^2(\Omega)} = 0 \tag{138}$$

holds true.

Each scale approximation $\Phi_\rho^{(2)} * F$ of a function $F \in \mathrm{L}^2(\Omega)$ must be made directly by computing the relevant convolution integrals. In doing so, however, it is inefficient to use no information from the approximation $\Phi_\rho^{(2)} * F$ within the computation of $\Phi_{\rho'}^{(2)} * F$ provided that $\rho' < \rho$. In fact, the efficient construction of multiscale approximation based on Dirac families usually begins by a *multiresolution analysis* in terms of wavelets, i.e., a recursive method which is efficient for computation, but not all economic multiscale approaches constitute multiresolution procedures (see, e.g., [35, 36, 38, 40–43, 47] and the references therein).

Let $\Psi_\rho(\xi,\eta)$, $(\xi,\eta) \in \Omega \times \Omega$, be defined via the series expansion

$$\Psi_\rho(\xi,\eta) = \sum_{n=0}^{\infty} {\Psi_\rho}^\wedge(n)\, \frac{2n+1}{4\pi}\, P_n(\xi\cdot\eta), \quad (\xi,\eta) \in \Omega\times\Omega, \tag{139}$$

such that the symbol $\{{\Psi_\rho^{(2)}}^\wedge(n)\}_{n=0,1,\ldots}$ of $\Psi_\rho^{(2)} = \Psi_\rho * \Psi_\rho$ is derived from ${\Phi_\rho^{(2)}}^\wedge(n)$ via the differential equation ('scale equation')

$${\Psi_\rho^{(2)}}^\wedge(n) = -\rho\, \frac{\mathrm{d}}{\mathrm{d}\rho} {\Phi_\rho^{(2)}}^\wedge(n). \tag{140}$$

As immediate consequences, we obtain from (135) the properties

$${\Psi_\rho}^\wedge(0) = 0 \tag{141}$$

and

$$\lim_{\rho\to 0,\,\rho>0} \Psi_\rho^\wedge(n) = 0$$

for $n = 1, 2, \ldots$ As in classical one-dimensional theory, the condition (135), therefore, justifies the notion *wavelet* of order 0.

Typically, within wavelet nomenclature, we may write

$$\Psi_{\rho;\eta} : \xi \mapsto \Psi_{\rho;\eta}(\xi) = \Psi_\rho(\xi\cdot\eta) = R_\eta D_\rho \Psi(\cdot\xi), \quad \xi \in \Omega,$$

to indicate $\Psi_{\rho;\eta}$ as generated by two parameters, namely the '*η-rotation operator*' R_η and the '*ρ-dilation operator*' D_ρ, respectively, given by

$$R_\eta : \Psi(\cdot\xi) \mapsto R_\eta \Psi(\cdot\xi) = \Psi(\eta\cdot\xi), \tag{142}$$

$$D_\rho : \Psi(\cdot\xi) \mapsto D_\rho \Psi(\cdot\xi) = \Psi_\rho(\cdot\xi). \tag{143}$$

The function $\Psi = \Psi_1$ (i.e., $\rho = 1$) is called the *mother wavelet*.

The *wavelet transform* WT is defined as the $\mathrm{L}^2(\Omega)$-inner product (convolution) of $F \in \mathrm{L}^2(\Omega)$ with the set of 'rotations' and 'dilations' of F

$$(\mathrm{WT})(F)(\rho;\eta) = (\Psi_{\rho;\eta}, F)_{\mathrm{L}^2(\Omega)} = \int_\Omega \Psi_{\rho;\eta}(\xi) F(\xi)\, \mathrm{d}\omega(\xi), \tag{144}$$

i.e., the wavelet transform acts as a space and frequency localization operator. The wavelet transform (WT) is invertible on the space of functions $F \in \mathrm{L}^2(\Omega)$ satisfying $F^\wedge(0,1) = 0$, i.e.,

$$F = \int_\Omega \int_0^\infty (\mathrm{WT})(F)(\rho;\eta)\Psi_{\rho;\eta}(\cdot)\,\frac{\mathrm{d}\rho}{\rho}\,\mathrm{d}\omega(\eta) \tag{145}$$

holds true (in the sense of $\|\cdot\|_{\mathrm{L}^2(\Omega)}$) for all $F \in \mathrm{L}^2(\Omega)$ satisfying $F^\wedge(0,1) = 0$.

The reconstruction formula (145), in fact, is based on the simple idea of dilation and rotation of the mother wavelet.

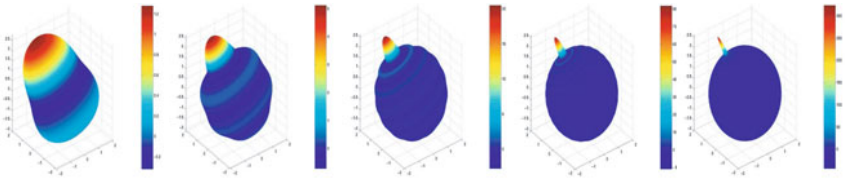


FIGURE 14. Shannon scaling (kernel) functions for decreasing scales ρ.

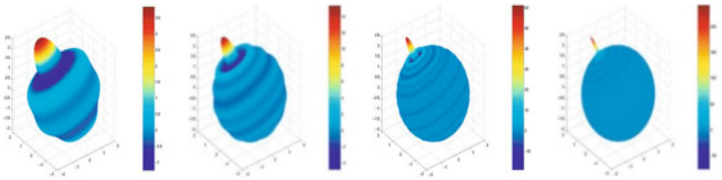


FIGURE 15. Shannon wavelet (kernel) functions for decreasing scales ρ.

Spectral interrelation between Fourier and wavelet transform. In terms of filtering, $\{\Phi_\rho\}_{\rho\in(0,\infty)}$ and $\{\Psi_\rho\}_{\rho\in(0,\infty)}$ may be interpreted (cf. Figures 14 and 15) as lowpass filter and bandpass filter, respectively. Correspondingly, the convolution operators are given by

$$\Phi_\rho * F, \quad F \in \mathrm{L}^2(\Omega), \tag{146}$$

$$\Psi_\rho * F, \quad F \in \mathrm{L}^2(\Omega). \tag{147}$$

The Fourier transforms read as follows:

$$(\Phi_\rho * F)^\wedge(n,j) = F^\wedge(n,j)\Phi_\rho^\wedge(n), \tag{148}$$

$$(\Psi_\rho * F)^\wedge(n,j) = F^\wedge(n,j)\Psi_\rho^\wedge(n). \tag{149}$$

These formulas provide the transition from the wavelet transform to the Fourier transform. Since all scales ρ are used, the reconstruction is highly redundant.

If $F, G \in \mathrm{L}^2(\Omega)$ have vanishing moments of order 0, i.e., if the property $F^\wedge(0,1) = G^\wedge(0,1) = 0$ is satisfied, then it follows from

$$\int_0^\infty (\Psi_\rho^\wedge(n))^2\,\frac{\mathrm{d}\rho}{\rho} = 1 \tag{150}$$

and the Parseval identity of the theory of spherical harmonics that

$$\begin{aligned}
&\int_\Omega \int_0^\infty \langle F, \Psi_{\rho;\eta}\rangle_{\mathrm{L}^2(\Omega)} \langle G, \Psi_{\rho;\eta}\rangle_{\mathrm{L}^2(\Omega)} \frac{\mathrm{d}\rho}{\rho}\,\mathrm{d}\omega(\eta) \\
&\quad = \int_0^\infty \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} F^\wedge(n,j) G^\wedge(n,j) (\Psi_\rho^\wedge(n))^2 \frac{\mathrm{d}\rho}{\rho} \\
&\quad = \sum_{n=1}^{\infty} \sum_{j=1}^{2n+1} F^\wedge(n,j) G^\wedge(n,j) \\
&\quad = \langle F, G\rangle_{\mathrm{L}^2(\Omega)}. \qquad (151)
\end{aligned}$$

Denote by $\mathrm{L}^2((0,\infty)\times\Omega)$ the space of all integrable functions $H : (0,\infty)\times\Omega \to \mathbb{R}$ such that

$$\int_\Omega \int_0^\infty |H(\rho;\eta)|^2 \frac{\mathrm{d}\rho}{\rho}\,\mathrm{d}\omega(\eta) < \infty. \qquad (152)$$

On the space $\mathrm{L}^2((0,\infty)\times\Omega)$, an inner product $\langle\cdot,\cdot\rangle_{\mathrm{L}^2((0,\infty)\times\Omega)}$ can be imposed corresponding to the norm

$$\|H\|_{\mathrm{L}^2((0,\infty)\times\Omega)} = \left(\int_\Omega \int_0^\infty |H(\rho;\eta)|^2 \frac{\mathrm{d}\rho}{\rho}\,\mathrm{d}\omega(\eta)\right)^{1/2}. \qquad (153)$$

From (151), it follows that

$$\left\langle \langle F, \Psi_{\cdot,\cdot}\rangle_{\mathrm{L}^2(\Omega)}, \langle G, \Psi_{\cdot,\cdot}\rangle_{\mathrm{L}^2(\Omega)} \right\rangle_{\mathrm{L}^2((0,\infty)\times\Omega)} = \langle F, G\rangle_{\mathrm{L}^2(\Omega)} \qquad (154)$$

and

$$\| \langle F, \Psi_{\cdot,\cdot}\rangle_{\mathrm{L}^2(\Omega)} \|^2_{\mathrm{L}^2((0,\infty)\times\Omega)} = \|F\|^2_{\mathrm{L}^2(\Omega)}. \qquad (155)$$

In other words, the total energy of a signal can be continuously distributed by the wavelet transform into scale and spatially dependent 'signal subenergy'.

Least energy representation. WT is a transformation from the one-parameter space $\mathrm{L}^2(\Omega)$ into the two-parameter space $\mathrm{L}^2((0,\infty)\times\Omega)$. Thus, it is clear that (WT) is not surjective on $\mathrm{L}^2((0,\infty)\times\Omega)$. That means that $\mathcal{W} = (\mathrm{WT})(\mathrm{L}^2(\Omega))$ is a proper subspace of $\mathrm{L}^2((0,\infty)\times\Omega)$:

$$\mathcal{W} \subsetneqq \mathrm{L}^2((0,\infty)\times\Omega). \qquad (156)$$

Thus, the problem is to characterize $\mathcal{W}$ within the framework of $\mathrm{L}^2((0,\infty)\times\Omega)$. For that purpose, we consider the operator $P : \mathrm{L}^2((0,\infty)\times\Omega) \to \mathcal{W}$ given by

$$(PH)(\rho';\eta') = \int_0^\infty \int_\Omega K(\rho';\eta' \mid \rho;\eta) H(\rho;\eta)\,\mathrm{d}\omega(\eta)\,\frac{\mathrm{d}\rho}{\rho}, \qquad (157)$$

where

$$K(\rho';\eta' \mid \rho;\eta) = \int_\Omega \Psi_{\rho';\eta'}(\xi)\Psi_{\rho;\eta}(\xi)\,\mathrm{d}\omega(\xi).$$

$\mathcal{W}$ is characterized as follows: $H \in \mathcal{W}$ if and only if

$$H(\rho';\eta') = \int_0^\infty \int_\Omega K(\rho';\eta' \mid \rho;\eta) H(\rho;\eta)\, \mathrm{d}\omega(\eta)\, \frac{\mathrm{d}\rho}{\rho}. \tag{158}$$

It can easily be seen that $K(\rho';\eta' \mid \cdot;\cdot) \in \mathcal{W}$ and $K(\cdot;\cdot \mid \rho;\eta) \in \mathcal{W}$. The kernel $K(\rho';\eta' \mid \rho;\eta)$ is the reproducing kernel in $\mathcal{W}$. The reproducing property (158) can also be understood in such a way that $H \in \mathcal{W}$ is calculable by superpositions of itself. This shows that there is a kind of linear dependence, which can be interpreted as redundancy. Although it might seem inefficient, such redundancy has certain advantages. Unlike a non-redundant expansion, errors can be detected and corrected.

The tendency for correcting errors is expressed in the next result (see [35, 47]):

– Let H be an arbitrary element of $\mathrm{L}^2((0,\infty)\times\Omega)$. Then the unique function $F_H \in \mathrm{L}^2(\Omega)$ which satisfies the property

$$\left\| H - \tilde{F}_H \right\|_{\mathrm{L}^2((0,\infty)\times\Omega)} = \inf_{F\in\mathrm{L}^2(\Omega)} \left\| H - \tilde{F} \right\|_{\mathrm{L}^2((0,\infty)\times\Omega)}$$

(with $\tilde{F}_H = (\mathrm{WT})(F_H)$) is given by

$$F_H(\xi) = \int_0^\infty \int_\Omega H(\rho;\eta) \Psi_{\rho;\eta}(\xi)\, \mathrm{d}\omega(\eta)\, \frac{\mathrm{d}\rho}{\rho}.$$

Indeed, $\tilde{F}_H$ is the orthogonal projection of H onto $\mathcal{W}$, which explains the aforementioned statement.

The linear dependence of $\tilde{F} \in \mathcal{W}$ leads to the effect that the coefficients in $\mathrm{L}^2((0,\infty)\times\Omega)$ for reconstructing a function $F \in \mathrm{L}^2(\Omega)$ are not unique. This can be easily seen from the following identity:

$$F(\xi) = \int_0^\infty \int_\Omega \left(\tilde{F}(\rho;\eta) + \tilde{F}^\perp(\rho;\eta) \right) \Psi_{\rho;\eta}(\xi)\, \mathrm{d}\omega(\eta)\, \frac{\mathrm{d}\rho}{\rho}$$

where $\tilde{F} = (\mathrm{WT})(F)$ and $\tilde{F}^\perp$ is an arbitrary member of $\mathcal{W}^\perp$. Nevertheless, we are able to deal with the following question: given an arbitrary $H(\rho;\xi) = (\mathrm{WT})(F)(\rho;\xi)$, $\rho \in (0,\infty)$, and $\xi \in \Omega$, for some $F \in \mathrm{L}^2(\Omega)$, how can we reconstruct F? The answer (see [35, 47]) is provided by the so-called *least-energy representation*:

– Of all possible functions $H \in \mathrm{L}^2((0,\infty)\times\Omega)$ for $F \in \mathrm{L}^2(\Omega)$, the function $H = (\mathrm{WT})(F)$ is unique in that it minimizes the 'energy' $\|H\|^2_{\mathrm{L}^2((0,\infty)\times\Omega)}$. More explicitly,

$$\|(\mathrm{WT})(F)\|_{\mathrm{L}^2((0,\infty)\times\Omega)} = \inf_{\substack{H\in\mathrm{L}^2((0,\infty)\times\Omega) \\ (\mathrm{WT})^{-1}(H)=F}} \|H\|_{\mathrm{L}^2((0,\infty)\times\Omega)}.$$

Wavelet variants. The construction of *spherical wavelets* has seen an enormous increase of activity in the last few years. Three features are essential in the thinking about georelevant wavelets: basis property, decorrelation, and fast computation.

First, wavelets are building blocks for general datasets derived from functions. By virtue of the basis property, each element of a general class of functions (e.g., a geopotential seen as a member of a set of potentials within a Sobolev space framework) can be expressed in stable way as a linear combination of dilated and shifted copies of a 'mother function' (see [29, 31, 35, 36] and the references therein). The role of the wavelet transform as a mapping from the class of functions into an associated two-parameter family of space and scale dependent functions is properly characterized by least squares properties.

Second, wavelets have the power to decorrelate. In other words, the representation of data in terms of wavelets is somehow 'more compact' than the original representation. We search for an accurate approximation by only using a small fraction of the original information of a function. Typically, the decorrelation is achieved by building wavelets which have a compact support (localization in space), which are smooth (decay towards high frequencies), and which have vanishing moments (decay towards low frequencies). Different types of wavelets can be found from certain constructions of space/momentum localization. The uncertainty principle tells us that sharp localization in 'space and momentum' are mutually exclusive. Nevertheless, decay towards long and short wavelengths (i.e., bandpass filtering) can be assured without any difficulty. Moreover, vanishing moments of wavelets (see, e.g., [45, 47]) enable us to combine (polynomial) outer harmonic expansions (responsible for the long-wavelength part of a function) with wavelet multiscale expansions (responsible for the medium-to-short-wavelengths contributions).

Third, the main question of recovering a function on the sphere, e.g., the Earth's gravitational potential, is how to decompose the function into wavelet coefficients, and how to reconstruct efficiently the potential from the coefficients. There is a 'tree algorithm' or 'pyramid algorithm' (cf. [29, 47]) that makes these steps simple and fast. In this respect, it is desirable to switch between the original representation of the data and its wavelet representation in a time proportional to the size of the data. In fact, the fast decorrelation power of wavelets is the key to applications such as data compression, fast data transmission, noise cancelation, signal recovery, etc.

In the last years, wavelets on the sphere have been the focus of several research groups which led to different wavelet approaches. Common to all these proposals is a multiresolution analysis which enables a balanced amount of both frequency (more accurately, angular momentum) and space localization (see, e.g., [18, 64, 82, 83, 101]).

A group theoretical approach to a continuous wavelet transform on the sphere is followed by Antoine and Vandergheynst [5], Antoine et al. [4], and Holschneider [55]. The parameter choice of their continuous wavelet transform is the product of SO(3) (for the motion on the sphere) and $\mathbb{R}^+$ (for the dilations). A continuous

wavelet transform approach for analyzing functions on the sphere is presented by Dahlke and Maass [19].

The Kaiserslautern constructions (see, e.g., [32, 39, 43, 47]) are intrinsically based on the specific properties concerning the theory of spherical harmonics. Wavelet regularization and multiresolution techniques are applied to 'downward continuation' of spaceborne (satellite) observations (see, e.g., [31, 37, 38, 48] and the references therein). Multiscale signal-to noise ratio modeling is done by signal and noise decorrelation Freeden and Maier [33, 34]. Freeden and Schreiner [42] are interested in a compromise connecting zonal function expressions and structured grids on the sphere to obtain fast algorithms. Freeden et al. [49] and Freeden and Gerhards [31] generate locally supported wavelets by regularizing fundamental solutions to pseudodifferential operators.

Finally, much of the material presented in this paper within a spherical framework can be readily formulated for non-spherical reference surfaces, even for vector and tensor data. Nevertheless, work remains to be done for more realistic geometries such as (the actual) Earth's surface, real satellite orbits, etc. These are challenges for future research.

4.4. Regularized functional matching pursuit and its variants

The Regularized Functional Matching Pursuit (RFMP) and its variants were developed by Fischer [23], Fischer and Michel [24], Michel [69], Michel and Telschow [72, 73], and Telschow [95]. They are based on the Matching Pursuit (MP) and its enhancements as described by Mallat and Zhang [65] and Vincent and Bengio [98], where the problem consisted of finding a greedy algorithm for the approximation of an unknown signal F based on given samples $F(x_j)$, $j = 1, \ldots, N$, usually on Euclidean domains.

For the RFMP, matching pursuit had to be extended to the inverse problem

$$\mathcal{F}F = y$$

for a linear and continuous operator $\mathcal{F} : \mathcal{H}(D) \to \mathbb{R}^l$, a Hilbert space $\mathcal{H}(D)$ of (some) functions on $D \subset \mathbb{R}^d$ (e.g., $\mathrm{L}^2(D)$ or, more generally, a Sobolev space), a given data vector $y \in \mathbb{R}^l$ and an unknown function $F \in \mathcal{H}(D)$. Many inverse problems of this kind, such as the downward continuation (F is the gravitational potential at the surface $D(= \Omega)$ and y is a vector of samples at satellite height) or the inverse gravimetric problem (F is a volume or a surface mass distribution and y is a vector of samples of the gravitational potential), are ill posed. For this reason, a regularization technique also had to be included into the RFMP.

The different algorithms are summarized here starting with a short introduction of the MP. All algorithms have in common that a set of possibly useful trial functions, the 'dictionary' $\mathcal{D} \subset \mathcal{H}(D)$, is chosen in advance. These trial functions need not originate from one single basis system: $\mathcal{D}$ may be (and is often chosen on purpose as) overcomplete. If $\mathcal{D}$ is heterogeneous, it may contain different kinds of basis systems (in particular, with different frequency and space localization). For instance, in several numerical applications of the RFMP, an approximate solution

F was combined from spherical harmonics (for a coarse global approximation) and radial basis functions with different levels of localization (locally improving the result). Without loss of generality, one can assume that $\|d\|_{\mathcal{H}(D)} = 1$ for all $d \in \mathcal{D}$.

Matching pursuit. Assume that a function (signal) $F \in \mathcal{H}(D)$ is to be approximated by m elements of $\mathcal{D}$. In this context, the expression of the best-m-term approximation (see, e.g., [96]) occurs. It means that one looks for m elements $d_1, \ldots, d_m \in \mathcal{D}$ and associated coefficients $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that the approximation error

$$\left\| F - \sum_{k=1}^{m} \alpha_k d_k \right\|_{\mathcal{H}(D)}$$

becomes minimal in comparison to all other choices of d_k and α_k. In formal language, the objective is

$$\sigma_m(F, \mathcal{D}) = \inf_{d_j \in \mathcal{D},\, \alpha_j \in \mathbb{R};\; j=1,\ldots,m} \left\| F - \sum_{k=1}^{m} \alpha_k d_k \right\|_{\mathcal{H}(D)} .$$

For large m, it is often numerically too expensive to find an exact minimizer. However, this concept can be a guideline for the construction of a less expensive algorithm with still 'good' results.

The first idea is to construct an iterative algorithm, i.e., to find the pairs $(\alpha_1, d_1), \ldots, (\alpha_m, d_m)$ consecutively. The initial problem is to find $\alpha_1 \in \mathbb{R}$ and $d_1 \in \mathcal{D}$ such that

$$J(\alpha_1, d_1) = \|F - \alpha_1 d_1\|^2_{\mathcal{H}(D)} = \|F\|^2_{\mathcal{H}(D)} - 2\alpha_1 \langle F, d_1 \rangle_{\mathcal{H}(D)} + \alpha_1^2 \tag{159}$$

is minimal. With $\frac{\partial}{\partial \alpha_1} J(\alpha_1, d_1) = 0$, one obtains

$$-2 \langle F, d_1 \rangle_{\mathcal{H}(D)} + 2\alpha_1 = 0, \qquad \text{i.e.,} \qquad \alpha_1 = \langle F, d_1 \rangle_{\mathcal{H}(D)} .$$

Inserting this result in (159), one gets

$$J(\alpha_1, d_1) = \|F\|^2_{\mathcal{H}(D)} - \langle F, d_1 \rangle^2_{\mathcal{H}(D)} .$$

Consequently, this dictionary element $d_1 \in \mathcal{D}$ for which F has the largest projection, i.e., the dictionary element which is most collinear to F, is the optimal choice in the first step. The first approximation is, therefore,

$$F_1 = \langle F, d_1 \rangle_{\mathcal{H}(D)} \, d_1,$$

where $d_1 \in \mathcal{D}$ is a maximizer of $\langle F, d_1 \rangle^2_{\mathcal{H}(D)}$, i.e.,

$$d_1 = \arg\max_{d \in \mathcal{D}} \langle F, d \rangle^2_{\mathcal{H}(D)} .$$

With the residual $R^1 = F - F_1$, one can analogously proceed. In general, if R^n is given, then one has to find $d_{n+1} \in \mathcal{D}$ such that $\langle R^n, d_{n+1} \rangle^2_{\mathcal{H}(D)}$ is maximal and then sets

$$F_{n+1} = F_n + \langle R^n, d_{n+1} \rangle_{\mathcal{H}(D)} \, d_{n+1}.$$

Functional matching pursuit. In the case of an inverse problem $\mathcal{F}F = y$, one minimizes the data misfit

$$\begin{aligned} J(\alpha_1, d_1) &= \|y - \mathcal{F}(\alpha_1 d_1)\|^2_{\mathbb{R}^l} \\ &= \|y\|^2_{\mathbb{R}^l} - 2\alpha_1 \langle y, \mathcal{F}d_1\rangle_{\mathbb{R}^l} + \alpha_1^2 \|\mathcal{F}d_1\|^2_{\mathbb{R}^l}, \end{aligned}$$

which implies that, again by assuming that $\frac{\partial}{\partial\alpha_1} J(\alpha_1, d_1) = 0$,

$$\alpha_1 = \frac{\langle y, \mathcal{F}d_1\rangle_{\mathbb{R}^l}}{\|\mathcal{F}d_1\|^2_{\mathbb{R}^l}}.$$

Consequently,

$$J(\alpha_1, d_1) = \|y\|^2_{\mathbb{R}^l} - \frac{\langle y, \mathcal{F}d_1\rangle^2_{\mathbb{R}^l}}{\|\mathcal{F}d_1\|^2_{\mathbb{R}^l}}$$

shows that d_1 has to be chosen such that

$$\frac{\langle y, \mathcal{F}d_1\rangle^2_{\mathbb{R}^l}}{\|\mathcal{F}d_1\|^2_{\mathbb{R}^l}}$$

is maximal. Then,

$$F_1 = \frac{\langle y, \mathcal{F}d_1\rangle_{\mathbb{R}^l}}{\|\mathcal{F}d_1\|^2_{\mathbb{R}^l}} d_1$$

is the first approximation. With the residual $R^1 = y - \mathcal{F}F_1$, one proceeds again analogously. Hence, for a given residual R^n, one chooses d_{n+1} such that

$$\frac{\langle R^n, \mathcal{F}d_{n+1}\rangle^2_{\mathbb{R}^l}}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l}}$$

is maximal and we set

$$F_{n+1} = F_n + \frac{\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l}}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l}} d_{n+1}.$$

Regularized functional matching pursuit. For the handling of ill-posed inverse problems, the Regularized Functional Matching Pursuit (RFMP) includes a Tikhonov-type regularization term

$$\lambda \|F_n\|^2_{\mathcal{H}(D)},$$

where $\lambda \in \mathbb{R}^+$ is a regularization parameter. Note that the choice of the (Sobolev) space $\mathcal{H}(D)$ influences the obtained result by requiring a particular kind of 'smoothness'. For instance, the Sobolev space $\mathrm{H}^{(2)}(\Omega)$ yields a regularization term which is not equal but similar to the norm which occurs in the minimum principle of spherical spline interpolation (see Theorem 4).

In analogy to the above, let $F_n \in \mathcal{H}(D)$ be the approximation after iteration n and $R^n = y - \mathcal{F}F_n$ be the residual, the error on the right-hand side of the inverse problem $\mathcal{F}F = y$. We find $d_{n+1} \in \mathcal{D}$ and $\alpha_{n+1} \in \mathbb{R}$ such that

$$J_\lambda(\alpha_{n+1}, d_{n+1}) = \|R^n - \mathcal{F}(\alpha_{n+1} d_{n+1})\|^2_{\mathbb{R}^l} + \lambda \|F_n + \alpha_{n+1} d_{n+1}\|^2_{\mathcal{H}(D)}$$

is minimal. Treating the functional J_λ like J above, one obtains

$$J_\lambda(\alpha_{n+1}, d_{n+1}) = \|R^n\|^2_{\mathbb{R}^l} - 2\alpha_{n+1}\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l} + \alpha^2_{n+1}\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l} + \lambda\left(\|F_n\|^2_{\mathcal{H}(D)} + 2\alpha_{n+1}\langle F_n, d_{n+1}\rangle_{\mathcal{H}(D)} + \alpha^2_{n+1}\right), \quad (160)$$

where the necessary condition $\frac{\partial}{\partial\alpha_{n+1}} J_\lambda(\alpha_{n+1}, d_{n+1}) = 0$ yields

$$\alpha_{n+1} = \frac{\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l} - \lambda\langle F_n, d_{n+1}\rangle_{\mathcal{H}(D)}}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l} + \lambda}. \quad (161)$$

If one inserts (161) into (160), one gets

$$J_\lambda(\alpha_{n+1}, d_{n+1}) = \|R^n\|^2_{\mathbb{R}^l} + \lambda\|F_n\|^2_{\mathcal{H}(D)} - \frac{\left(\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l} - \lambda\langle F_n, d_{n+1}\rangle_{\mathcal{H}(D)}\right)^2}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l} + \lambda}$$

such that $d_{n+1} \in \mathcal{D}$ has to be chosen as a maximizer of

$$\frac{\left(\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l} - \lambda\langle F_n, d_{n+1}\rangle_{\mathcal{H}(D)}\right)^2}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l} + \lambda}.$$

This yields the following algorithm (where the Functional Matching Pursuit is a particular case for $\lambda = 0$).

Algorithm 6 (RFMP). *Let a data vector $y \in \mathbb{R}^l$, a linear and continuous operator $\mathcal{F} : \mathcal{H}(D) \to \mathbb{R}^l$, a dictionary $\mathcal{D} \subset \{d \in \mathcal{H}(D) \,|\, \|d\|_{\mathcal{H}(D)} = 1\}$ and an initial approximation $F_0 \in \mathcal{H}(D)$ be given.*

(i) *Initialize the iteration with $n = 0$ and $R^0 = y - \mathcal{F}F_0$ and select a stopping criterion (data-misfit-based, i.e., choose $\varepsilon > 0$ to require $\|R^{n+1}\| < \varepsilon$, or iteration-based, i.e., choose $N \in \mathbb{N}$ to require $n + 1 \leq N$) as well as a regularization parameter $\lambda \in \mathbb{R}^+_0$.*

(ii) *Determine*

$$d_{n+1} = \arg\max_{d\in\mathcal{D}} \frac{\left(\langle R^n, \mathcal{F}d\rangle_{\mathbb{R}^l} - \lambda\langle F_n, d\rangle_{\mathcal{H}(D)}\right)^2}{\|\mathcal{F}d\|^2_{\mathbb{R}^l} + \lambda}, \quad (162)$$

$$\alpha_{n+1} = \frac{\langle R^n, \mathcal{F}d_{n+1}\rangle_{\mathbb{R}^l} - \lambda\langle F_n, d_{n+1}\rangle_{\mathcal{H}(D)}}{\|\mathcal{F}d_{n+1}\|^2_{\mathbb{R}^l} + \lambda} \quad (163)$$

and set $F_{n+1} = F_n + \alpha_{n+1}d_{n+1}$ and $R^{n+1} = R^n - \alpha_{n+1}\mathcal{F}d_{n+1}$.

(iii) *If the stopping criterion is satisfied, then use F_{n+1} as an approximate solution to $\mathcal{F}F = y$. Otherwise, increase n by 1 and go to step* (ii).

The algorithm is accelerated if one implements the following procedures.

- *Normalize the dictionary:* use the assumption above and choose all $d \in \mathcal{D}$ such that $\|d\|_{\mathcal{H}(D)} = 1$, otherwise the norm of the dictionary elements occurs in (162) and (163) (see, e.g., [24]).

- *Move as much as possible to the preprocessing:* calculate $\|\mathcal{F}d\|_{\mathbb{R}^l}$ for all $d \in \mathcal{D}$ and the (symmetric) matrices with the components $\langle d, \tilde{d} \rangle_{\mathcal{H}(D)}$ and $\langle \mathcal{F}d, \mathcal{F}\tilde{d} \rangle_{\mathbb{R}^l}$, respectively, (with $d, \tilde{d} \in \mathcal{D}$) once and store them.
- *Use preprocessing for finding d_{n+1} and α_{n+1}:* note, in particular, that

$$\langle R^n, \mathcal{F}d \rangle_{\mathbb{R}^l} = \langle R^{n-1}, \mathcal{F}d \rangle_{\mathbb{R}^l} - \alpha_n \langle \mathcal{F}d_n, \mathcal{F}d \rangle_{\mathbb{R}^l} ,$$
$$\langle F_n, d \rangle_{\mathcal{H}(D)} = \langle F_{n-1}, d \rangle_{\mathcal{H}(D)} + \alpha_n \langle d_n, d \rangle_{\mathcal{H}(D)} ,$$

  where, in both cases, the first summands on the right-hand side are already known from the previous iteration step (i.e., step $n-1$).

It should be mentioned that, in (162), the maximizer need not be uniquely determined. In this case, no particular strategy for choosing between several maximizers has been applied yet.

One essential result is the following convergence theorem.

Theorem 7 (Convergence Theorem). *Let the dictionary $\mathcal{D}$ satisfy:*

(i) *'semi-frame condition': There exist a constant $c > 0$ and an integer N such that, for all expansions $H = \sum_{k=1}^{\infty} \beta_k d_k$ with $\beta_k \in \mathbb{R}$ and $d_k \in \mathcal{D}$, where the d_k are not necessarily pairwise distinct but $\{j \in \mathbb{N} \,|\, d_j = d_k\}$ is a finite set with at most N elements for each $k \in \mathbb{N}$,*

$$c\|H\|^2_{\mathcal{H}(D)} \leq \sum\nolimits_{k=1}^{\infty} \beta_k^2 .$$

(ii) *$\|d\|_{\mathcal{H}(D)} = 1$ for all $d \in \mathcal{D}$ and, if $\lambda = 0$, then $\inf_{d \in \mathcal{D}} \|\mathcal{F}d\|_{\mathbb{R}^l} > 0$ is required additionally.*

If the sequence $(F_n)_n$ is produced by the RFMP and no dictionary element is chosen more than N times, then $(F_n)_n$ converges in $\mathcal{H}(D)$ to $F_\infty = F_0 + \sum_{n=1}^{\infty} \alpha_n d_n \in \mathcal{H}(D)$. Moreover, the following holds true:

(a) *If $\overline{\operatorname{span} \mathcal{D}}^{\|\cdot\|_{\mathcal{H}(D)}} = \mathcal{H}(D)$ and $\lambda \in \mathbb{R}_0^+$ is an arbitrary parameter, then F_∞ solves*

$$(\mathcal{F}^*\mathcal{F} + \lambda \mathcal{I}) F_\infty = \mathcal{F}^* y ,$$

where $\mathcal{F}^*$ is the adjoint operator corresponding to $\mathcal{F}$ and $\mathcal{I}$ is the identity operator on $\mathcal{H}(D)$. In other words,

$$\|y - \mathcal{F}F_\infty\|^2_{\mathbb{R}^l} + \lambda \|F_\infty\|^2_{\mathcal{H}(D)} = \min_{F \in \mathcal{H}(D)} \left(\|y - \mathcal{F}F\|^2_{\mathbb{R}^l} + \lambda \|F\|^2_{\mathcal{H}(D)} \right) ,$$

where the minimizer is unique, if $\lambda > 0$.

(b) *If $\operatorname{span} \{\mathcal{F}d \,|\, d \in \mathcal{D}\} = \mathbb{R}^l$ and $\lambda = 0$, then F_∞ solves $\mathcal{F}F_\infty = y$.*

Note that the semi-frame condition has been changed (including the requirement on repeated choices of dictionary elements) in comparison to earlier publications on the RFMP by Michel [69] and Michel and Telschow [72], since an unlimited number of equally chosen dictionary elements would allow a counterexample for which the semi-frame condition could not be achieved, as it was pointed out in [73]. For a proof of the convergence theorem and additional properties, see [71]. For numerical examples of RFMP applied to geodetic problems, see [23–26, 72].

Regularized orthogonal functional matching pursuit. Numerical experiments show that the RFMP chooses some dictionary elements several times, which actually means that some of the previously calculated coefficients $\alpha_1, \ldots, \alpha_n$ are corrected. The reason for this phenomenon is that the dictionary elements (or their images in the data space) are typically non-orthogonal. In the case of the Matching Pursuit (MP), this effect is compensated for by introducing a particular orthogonal projection procedure in the Orthogonal Matching Pursuit (OMP, see [78]) and by using 'prefitting' (see [98]). However, the OMP requires that the data and the solution are in the same space for performing the projection and it also does not contain a regularization.

In [95] and [73], the idea behind OMP and 'prefitting' was used to enhance RFMP to Regularized Orthogonal Functional Matching Pursuit (ROFMP). It is now possible to update the coefficients α_i in every iteration. For this reason, the approximation after step n is represented by

$$F_n = \sum_{i=1}^{n} \alpha_i^{(n)} d_i.$$

If one measures the quality of an approximate solution in the data space, i.e., in the sense of the data misfit, then the best approximation (without a regularization) in terms of (fixed) $d_1, \ldots, d_n$ would be given by requiring that $\mathcal{F}F_n$ equals the orthogonal projection of y onto

$$\mathcal{V}_n = \text{span}\,\{\mathcal{F}d_1, \ldots, \mathcal{F}d_n\},$$

i.e., $\mathcal{F}F_n = \mathcal{P}_{\mathcal{V}_n} y$. This is equivalent to requiring that the residual $R^n = y - \mathcal{F}F_n$ is orthogonal to $\mathcal{V}_n$. Geometrically speaking, $\mathcal{F}F_n$ is the projection of y onto the hyperplane $\mathcal{V}_n$ and R^n is the associated plumbline, see Figure 16.

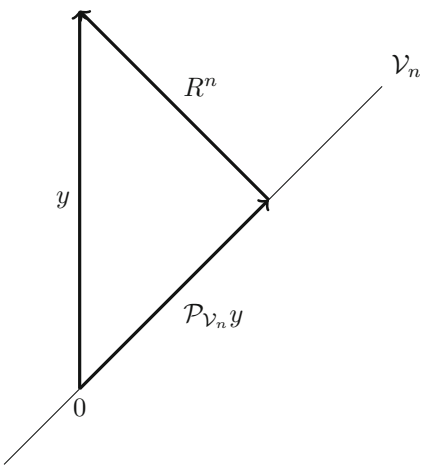


FIGURE 16. Illustration of the orthogonal projection $\mathcal{P}_{\mathcal{V}_n} y$ in $\mathbb{R}^l$.

Consequently, the next summand $\alpha_{n+1}d_{n+1}$ should complement the previous approximation F_n such that $\mathcal{F}F_{n+1} = \mathcal{P}_{\mathcal{V}_{n+1}}y$. However, in general, $\mathcal{P}_{\mathcal{V}_n}(\mathcal{F}d_{n+1}) \neq 0$. This projection would, however, deteriorate the previously exact approximation of $\mathcal{P}_{\mathcal{V}_n}y$ by $\mathcal{F}F_n$. For this reason, this redundant part is subtracted, i.e., one is interested in

$$\mathcal{F}F_{n+1} = \mathcal{F}F_n + \alpha_{n+1}\left[\mathcal{F}d_{n+1} - \mathcal{P}_{\mathcal{V}_n}\left(\mathcal{F}d_{n+1}\right)\right].$$

If one sets $\mathcal{P}_{\mathcal{V}_n}(\mathcal{F}d) = \sum_{i=1}^{n}\beta_i^{(n)}(d)\,\mathcal{F}d_i$, then

$$\begin{aligned}\mathcal{F}F_{n+1} &= \sum_{i=1}^{n}\alpha_i^{(n)}\,\mathcal{F}d_i - \alpha_{n+1}\sum_{i=1}^{n}\beta_i^{(n)}\left(d_{n+1}\right)\,\mathcal{F}d_i + \alpha_{n+1}\mathcal{F}d_{n+1}\\ &= \sum_{i=1}^{n}\left(\alpha_i^{(n)} - \alpha_{n+1}\beta_i^{(n)}\left(d_{n+1}\right)\right)\mathcal{F}d_i + \alpha_{n+1}\mathcal{F}d_{n+1}.\end{aligned}$$

Hence, the task is now (in step $n+1$) to find $\alpha \in \mathbb{R}$ and $d \in \mathcal{D}$ such that

$$\left\| y - \sum_{i=1}^{n}\left(\alpha_i^{(n)} - \alpha\beta_i^{(n)}\left(d\right)\right)\mathcal{F}d_i - \alpha\mathcal{F}d\right\|_{\mathbb{R}^l}$$

is minimized. As an approximation at step $n+1$, one uses then

$$F_{n+1} = \sum_{i=1}^{n}\left(\alpha_i^{(n)} - \alpha_{n+1}\beta_i^{(n)}\left(d_{n+1}\right)\right)d_i + \alpha_{n+1}d_{n+1},$$

$$\alpha_i^{(n+1)} = \alpha_i^{(n)} - \alpha_{n+1}\beta_i^{(n)}\left(d_{n+1}\right) \quad \text{for } i = 1,\ldots,n$$

and $\alpha_{n+1}^{(n+1)} = \alpha_{n+1}$. With the regularization, the functional to minimize is

$$\begin{aligned}&\left\| y - \sum_{i=1}^{n}\left(\alpha_i^{(n)} - \alpha\beta_i^{(n)}\left(d\right)\right)\mathcal{F}d_i - \alpha\mathcal{F}d\right\|_{\mathbb{R}^l}^2\\ &+ \lambda\left\|\sum_{i=1}^{n}\left(\alpha_i^{(n)} - \alpha\beta_i^{(n)}\left(d\right)\right)d_i + \alpha d\right\|_{\mathcal{H}(D)}^2.\end{aligned}$$

This is the principle of the ROFMP. We now introduce some abbreviations.

- The orthogonal complement of $\mathcal{V}_n$ in $\mathbb{R}^l$ is denoted by $\mathcal{W}_n$, i.e., $\mathcal{V}_n \oplus \mathcal{W}_n = \mathbb{R}^l$, and the projection of $\mathcal{F}d$ onto $\mathcal{W}_n$ is

$$\mathcal{P}_{\mathcal{W}_n}(\mathcal{F}d) = \mathcal{F}d - \sum_{i=1}^{n}\beta_i^{(n)}\left(d\right)\,\mathcal{F}d_i.$$

- The function associated to $\mathcal{P}_{\mathcal{V}_n}(\mathcal{F}d)$ in $\mathcal{H}(D)$ is denoted by

$$B_n(d) = \sum_{i=1}^{n}\beta_i^{(n)}\left(d\right)\,d_i.$$

Similar derivations as in the cases above finally yield the following algorithm.

Algorithm 8 (ROFMP). *Let a data vector $y \in \mathbb{R}^l$, a linear and continuous operator $\mathcal{F} : \mathcal{H}(D) \to \mathbb{R}^l$ and a dictionary $\mathcal{D} \subset \mathcal{H}(D) \setminus \{0\}$ be given.*

(i) *Initialize the iteration with $n = 0$, $F_0 = 0$ and $R^0 = y$ and select a stopping criterion (data-misfit-based, i.e., choose $\varepsilon > 0$ to require $\|R^{n+1}\| < \varepsilon$, or iteration-based, i.e., choose $N \in \mathbb{N}$ to require $n + 1 \leq N$) as well as a regularization parameter $\lambda \in \mathbb{R}_0^+$.*

(ii) *Determine*

$$d_{n+1} = \arg\max_{d \in \mathcal{D}} \frac{\left(\langle R^n, \mathcal{P}_{\mathcal{W}_n}(\mathcal{F}d) \rangle_{\mathbb{R}^l} + \lambda \langle F_n, B_n(d) - d \rangle_{\mathcal{H}(D)} \right)^2}{\|\mathcal{P}_{\mathcal{W}_n}(\mathcal{F}d)\|_{\mathbb{R}^l}^2 + \lambda \|B_n(d) - d\|_{\mathcal{H}(D)}^2},$$

$$\alpha_{n+1} = \frac{\langle R^n, \mathcal{P}_{\mathcal{W}_n}(\mathcal{F}d_{n+1}) \rangle_{\mathbb{R}^l} + \lambda \langle F_n, B_n(d_{n+1}) - d_{n+1} \rangle_{\mathcal{H}(D)}}{\|\mathcal{P}_{\mathcal{W}_n}(\mathcal{F}d_{n+1})\|_{\mathbb{R}^l}^2 + \lambda \|B_n(d_{n+1}) - d_{n+1}\|_{\mathcal{H}(D)}^2}.$$

(iii) *Update the coefficients as follows:*

$$\alpha_i^{(n+1)} = \alpha_i^{(n)} - \alpha_{n+1} \beta_i^{(n)}(d_{n+1}) \quad \text{for } i = 1, \ldots, n,$$
$$\alpha_{n+1}^{(n+1)} = \alpha_{n+1}$$

and set $F_{n+1} = \sum_{i=1}^{n+1} \alpha_i^{(n+1)} d_i$ as well as $R^{n+1} = y - \mathcal{F}F_{n+1}$.

(iv) *If the stopping criterion is satisfied, then use F_{n+1} as an approximate solution to $\mathcal{F}F = y$. Otherwise, increase n by 1 and go to step* (ii).

Obviously, a normalization of the dictionary elements to $\|d\|_{\mathcal{H}(D)} = 1$ does not yield an improvement for the implementation in the case of the ROFMP.

Note that the orthogonal projection becomes more and more expensive with an increasing number n. For this reason, it is advisable to restart the algorithm after a certain number of steps N by using $y - \mathcal{F}F_N$ as the new data vector to be approximated and recounting from $n = 0$. Due to the linearity of $\mathcal{F}$, the consecutively produced approximations can be summed up in the end to obtain an approximation of the solution F of $\mathcal{F}F = y$. It turned out to be useful to keep, after each restart, the previous approximation $F_N = \tilde{F}$ in the regularization term and to regularize with $\|\tilde{F} + F_n\|_{\mathcal{H}(D)}^2$, where (F_n) is the approximating sequence after the restart.

For details of the implementation, see [95]. For numerical experiments and theoretical results, see [73, 95]. Note that, in the non-regularized case ($\lambda = 0$), the algorithm is able to produce an exact solution of $\mathcal{F}F = y$ in at most l steps, where $y \in \mathbb{R}^l$.

5. Conclusion

For the last decades, the possibilities and challenges which have presented themselves to geodesists have changed dramatically. Due to tremendously increased precisions in measurement technologies and the availability of satellite missions, huge

amounts of highly accurate data related to the Earth have become available. This has opened previously unexpected options for observing, analyzing and predicting the processes of the Earth system. Such progresses can be seen in manifold ways, for example when the ocean dynamics can be understood better, when the mass transports due to climate change or seasonal climatic phenomena can be better quantified and localized, when static and dynamic models of the Earth's interior can be validated and improved by a more precise model or when unprecedented ways of determining heights become available to geodesists.

Since mathematics plays a central role in the processes of, e.g., denoising, analyzing or inverting geoscientific data, the changes in the data situation can be mapped to changes in the requirements on the methodologies in mathematical geodesy (see also [30]). In this paper, we focussed on the uncertainty principle of spherical signal analysis which tells us that precise localization in space and in frequency/momentum are mutually exclusive. Moreover, we can interpret the uncertainty principle as a fundamental property of a spectrum ranging from ideal frequency localization (i.e., no space localization) to ideal space localization (i.e., no frequency localization). The former is associated to the use of spherical harmonics, which have been a common choice as basis system in geodesy. Away from this extremal case, in order that trial functions possess a space localization, they need to be sums of several spherical harmonics. The closer we come to the latter end of the spectrum with ideal space localization, the more spherical harmonics degrees have to be summed up in a trial function leading, as a limit, finally to the (only as a theoretical concept existing) Dirac functional which includes all degrees.

The aforementioned new challenges due to today's data situation can be reflected in this spectrum. In former days, when only a few data were available which allowed a very coarse global modeling only, spherical harmonics were the ultimate and reasonable choice. Today, the demands on highly accurate models which are, in particular, provided with a very high resolution in space define the limits of the use of spherical harmonics. These models can be better constructed with trial functions which combine certain extents of space and frequency localization.

As we have shown, there are many facets of localized trial functions which can be positioned in the spectrum of space and momentum localization. They include basis functions generated from (reproducing) kernels of particular function spaces. Such tools have successfully been used for spline and wavelet approximations in the geosciences. They leave sufficient degrees of freedom to control their variance in space and momentum. Furthermore, also Slepian functions provide another equally valuable tool for regionally approximating or analyzing a signal. They provide us with an orthonormal basis which is, in contrast to spherical harmonics, spaceconcentrated (to a region which can be arbitrarily chosen). Moreover, the Slepian functions are also orthogonal in the L^2-space of the chosen region, which is essentially useful for the modeling of a signal which is only regionally available. Furthermore, Slepian functions can also be not only spaceconcentrated but even spacelimited with the price (due to the uncertainty principle) that they become nonbandlimited, i.e., they sum up an infinite number of spherical harmonics degrees.

Certainly, there exist many other systems of trial functions on the sphere but also the ball, which have their own characteristics regarding space- and frequency localization. We added some references to other methods in appropriate paragraphs but do not claim to have provided a complete overview. In general, a wide range of special functions systems is available for the analysis of geoscientific data. However, it appears that, still, the main focus of (too) many research projects in geodesy and other disciplines of Earth sciences lies on the data alone but not on the choice of the methodology for their handling.

In this paper, we have tried to break new synoptical ground in dealing with spherical harmonics based special function systems and their role in constructive approximation methods of mathematical geodesy. We have presented a short insight and guide for the zoo of spherical trial functions to encourage geoscientists to question the mathematical basis functions which they use for their models and not to use mathematical tools as 'blackboxes'. We have also summarized briefly the possibility that regularized functional matching pursuit and its variants yield as algorithms for generating a kind of a best basis out of a selection of different basis systems.

Further research on finding the 'optimal' basis system for particular problems in mathematical geodesy has to be done. However, the present state-of-the-art shows that there is a high potential in improving (not only) geodetic models by using sophisticated mathematical methodologies. Obviously, our work as presented here is selective, but not only with respect to the choice of discussed basis functions. Also, not all details on the treated topics could be discussed up to an appropriate extent. For example, most of the proofs have been left out completely, so that the interested reader is referred to the attached list of literature. Nonetheless, we believe that we have provided a deeper insight on how geoscientific and, particularly, geodetic problems can be attacked in a mathematically systematic and rigorous way.

References

[1] Amirbekyan, A. (2007) The Application of Reproducing Kernel Based Spline Approximation to Seismic Surface and Body Wave Tomography: Theoretical Aspects and Numerical Results. Ph.D.-thesis, Geomathematics Group, Dept. Mathematics, University of Kaiserslautern

[2] Amirbekyan, A., Michel, V. (2008) Splines on the Three-dimensional Ball and Their Application to Seismic Body Wave Tomography, Inverse Probl., 24, 015022 (25pp)

[3] Amirbekyan, A., Michel, V., Simons, F.J. (2008) Parameterizing Surface-wave Tomographic Models with Harmonic Spherical Splines, Geophys. J. Int., 174, pp. 617–628

[4] Antoine, J.P., Demanet, L., Jaques, L., Vandergheynst, P. (2002) Wavelets on the Sphere: Implementations and Approximations, Appl. Comput. Harmon. Anal., 13, pp. 177–200

[5] Antoine, J.P., Vandergheynst, P. (1999) Wavelets on the 2-Sphere: A Group-theoretic Approach, Appl. Comput. Harmon. Anal., 7, pp. 1–30

[6] Backus, G.E. (1966) Potentials for Tangent Tensor Fields on Spheroids, Arch. Ration. Mech. Anal., 22, pp. 210–252

[7] Backus, G.E. (1967) Converting Vector and Tensor Equations to Scalar Equations in Spherical Coordinates, Geophys. J.R. Astron. Soc., 13, pp. 61–101

[8] Backus, G.E. (1986) Poloidal and Toroidal Fields in Geomagnetic Field Modelling, Reviews of Geophysics, 24, pp. 75–109

[9] Ballani, L., Engels, J., Grafarend, E.W. (1993) Global Base Functions for the Mass Density in the Interior of a Massive Body (Earth), Manuscr. Geod., 18, pp. 99–114

[10] Berkel, P. (2009) Multiscale Methods for the Combined Inversion of Normal Mode and Gravity Variations. Ph.D.-thesis, Geomathematics Group, Dept. Mathematics, University of Kaiserslautern, Shaker, Aachen

[11] Berkel, P., Michel, V. (2010) On Mathematical Aspects of a Combined Inversion of Gravity and Normal Mode Variations by a Spline Method, Math. Geosci., 42, pp. 795–816

[12] Berkel, P., Fischer, D., Michel, V. (2011) Spline Multiresolution and Numerical Results for Joint Gravitation and Normal Mode Inversion with an Outlook on Sparse Regularisation, Int. J. Geomath., 1, pp. 167–204

[13] Blick, C., Freeden, W. (2011) Spherical Spline Application to Radio Occultation Data, J. Geodetic Science, 1, 379–396

[14] Bluhm, R., Kostelecký, V.A., Tudose, B. (1995) Elliptical Squeezed States and Rydberg Wave Packets, Phys. Rev. A, 52, pp. 2234–2244

[15] Carruthers, P., Nieto, M.M. (1968) Phase and Angle Variables in Quantum Mechanics, Rev. Mod. Phys., 40, pp. 411–440

[16] Clebsch, R.F.A. (1862) Ueber eine Eigenschaft der Kugelfunctionen. J. Reine Angew. Math., 1862(60), pp. 343–350

[17] Dahlen, F.A., Simons, F.J. (2008) Spectral Estimation on a Sphere in Geophysics and Cosmology, Geoph. J. Int., 174, pp. 774–807

[18] Dahlke, S., Dahmen, W., Schmitt, W., Weinreich, I. (1995) Multiresolution Analysis and Wavelets on S^2 and S^3, Numer. Funct. Anal. Optimiz., 16, pp. 19–41

[19] Dahlke, S., Maass, P. (1996) Continuous Wavelet Transforms with Applications to Analyzing Functions on Spheres, J. Fourier Anal. Appl., 2, pp. 379–396

[20] Daubechies, I. (1992) Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, 61, SIAM, Philadelphia

[21] Eshagh, M. (2009) Spatially Restricted Integrals in Gradiometric Boundary Value Problems, Artif. Sat., 44, pp. 131–148

[22] Fengler, M., Michel, D., Michel, V. (2006) Harmonic Spline-wavelets on the 3-dimensional Ball and Their Application to the Reconstruction of the Earth's Density Distribution From Gravitational Data at Arbitrarily Shaped Satellite Orbits, ZAMM-Z. Angew. Math. Me., 86, pp. 856–873

[23] Fischer, D. (2011) Sparse Regularization of a Joint Inversion of Gravitational Data and Normal Mode Anomalies. Ph.D.-thesis, Geomathematics Group, Dept. Mathematics, University of Siegen, Verlag Dr. Hut, Munich

[24] Fischer, D., Michel, V. (2012) Sparse Regularization of Inverse Gravimetry – Case Study: Spatial and Temporal Mass Variations in South America. Inverse Probl., 28, 065012 (34pp.)

[25] Fischer, D., Michel, V. (2013a) Automatic Best-basis Selection for Geophysical Tomographic Inverse Problems. Geophys. J. Int., 193, pp. 1291–1299

[26] Fischer, D., Michel, V. (2013b) Inverting GRACE Gravity Data for Local Climate Effects, J. Geodetic Science, 3, pp. 151–162

[27] Fokas, A.S., Hauk, O., Michel, V. (2012) Electro-Magneto-Encephalography for the Three-shell Model: Numerical Implementation via Splines for Distributed Current in Spherical Geometry, Inverse Probl., 28, 035009 (28pp.)

[28] Freeden, W. (1981) On Spherical Spline Interpolation and Approximation, Math. Meth. in the Appl. Sci., 3, pp. 551–575

[29] Freeden, W. (1999) Multiscale Modelling of Spaceborne Geodata, B.G. Teubner, Leipzig

[30] Freeden, W. (2015) Geomathematics: Its Role, Its Aim, and Its Potential, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 2nd ed., Springer, Berlin, Heidelberg, pp. 3–79

[31] Freeden, W., Gerhards, C. (2013) Geomathematically Oriented Potential Theory. Chapman and Hall, CRC Press, Boca Raton, London, New York

[32] Freeden, W., Gutting, M. (2013) Special Functions of Mathematical (Geo)Physics. Birkhäuser, Basel

[33] Freeden, W., Maier, T. (2002) On Multiscale Denoising of Spherical Functions: Basic Theory and Numerical Aspects, Electr. Transact. on Numer. Anal. (ETNA), 14, pp. 40–62

[34] Freeden, W., Maier, T. (2003) Spectral and Multiscale Signal-to-Noise Thresholding of Spherical Vector Fields, Comput. Geosciences, 7, pp. 215–250

[35] Freeden, W., Michel, V. (2000) Least-Squares Geopotential Approximation by Windowed Fourier Transform and Wavelet Transform, in: Wavelets in the Geosciences (R. Klees and R. Haagmans, Eds.), Lecture Notes in Earth Sciences, Springer, 90, pp. 189–241

[36] Freeden, W., Michel, V. (2004) Multiscale Potential Theory (with Applications to Geoscience), Birkhäuser Verlag, Boston, Basel, Berlin

[37] Freeden, W., Nutz, H. (2011) Satellite Gravity Gradiometry as Tensorial Inverse Problem, Int. J. Geomath., 2, pp. 177–218

[38] Freeden, W., Schneider, F. (1998) Regularization Wavelets and Multiresolution, Inverse Probl., 14, pp. 225–243

[39] Freeden, W., Schreiner, M. (1995) Non-Orthogonal Expansions on the Sphere, Math. Meth. Appl. Sci., 18, pp. 83–120

[40] Freeden, W., Schreiner, M. (1998) Orthogonal and Non-Orthogonal Multiresolution Analysis, Scale Discrete and Exact Fully Discrete Wavelet Transform on the Sphere. Constr. Approx., 14, pp. 493–515

[41] Freeden, W., Schreiner, M. (2006), Local Multiscale Modelling of Geoid Undulations from Deflections of the Vertical, J. Geodesy, 79, 641–651

[42] Freeden, W., Schreiner, M. (2007) Biorthogonal Locally Supported Wavelets on the Sphere Based on Zonal Kernel Functions, J. Fourier Anal. Appl., 13, pp. 693–709

[43] Freeden, W., Schreiner, M. (2009) Spherical Functions of Mathematical Geosciences – A Scalar, Vectorial, and Tensorial Setup, Springer, Berlin

[44] Freeden, W., Windheuser, U. (1996) Spherical Wavelet Transform and its Discretization, Adv. Comput. Math., 5, pp. 51–94

[45] Freeden, W., Windheuser, U. (1997) Combined Spherical Harmonic and Wavelet Expansion. Appl. Comput. Harmon. Anal., 4, pp. 1–37

[46] Freeden, W., Witte, B. (1982) A Combined (Spline-) Interpolation and Smoothing Method for the Determination of the External Gravitational Potential from Heterogeneous Data, Bull. Geod., 56, pp. 53–62

[47] Freeden, W., Gervens, T., Schreiner, M. (1998) Constructive Approximation on the Sphere (with Applications to Geomathematics), Oxford Science Publications, Clarendon, Oxford

[48] Freeden, W., Michel, V., Nutz, H. (2002) Satellite-to-Satellite Tracking and Satellite Gravity Gradiometry (Advanced Techniques for High-Resolution Geopotential Field Determination), J. Eng. Math., 43, pp. 19–56

[49] Freeden, W., Fehlinger, T., Klug, M., Mathar, D., and Wolf, K. (2009) Classical Globally Reflected Gravity Field Determination in Modern Locally Oriented Multiscale Framework, J. Geodesy, 83, pp. 1171–1191

[50] Gauß, C.F. (1838) Allgemeine Theorie des Erdmagnetismus, Resultate aus den Beobachtungen des magnetischen Vereins, Göttingen

[51] Grünbaum, F.A., Longhi, L., Perlstadt, M. (1982) Differential Operators Commuting with Finite Convolution Integral Operators: Some Non-Abelian Examples, SIAM J. Appl. Math., 42, pp. 941–955

[52] Guilloux, F., Faÿ, G., Cardoso, J.F. (2009) Practical Wavelet Design on the Sphere, Appl. Comput. Harmon. Anal., 26, pp. 143–160

[53] Haar, A. (1910) Zur Theorie der orthogonalen Funktionensysteme, Math. Ann., 69, pp. 331–371

[54] Heine, E. (1878) Handbuch der Kugelfunktionen, Reimer, Berlin

[55] Holschneider M (1996) Continuous Wavelet Transforms on the Sphere, J. Math. Phys., 37, pp. 4156–4165

[56] Kennedy, R.A. and Sadeghi, P. (2013) Hilbert Space Methods in Signal Processing, Cambridge University Press, Cambridge, UK

[57] Khalid, Z., Durrani, S., Sadeghi, P., Kennedy, R.A. (2012) Spatio-spectral Analysis on the Sphere Using Spatially Localized Spherical Harmonics Transform, IEEE Trans. Signal Process., 60, pp. 1487–1492

[58] Kowalski, K., Rembieliński, J. (2000) Quantum Mechanics on a Sphere and Coherent States, J. Phys. A: Math. Gen., 33, pp. 6035–6048

[59] Laín Fernández, N. (2003) Polynomial Bases on the Sphere, Ph.D.-thesis, University of Lübeck

[60] Landau, H.J. and Pollak, H.O. (1961) Prolate spheroidal wave functions, Fourier analysis and uncertainty – II, Bell Syst. Tech. J., 40, pp. 65–84

[61] Laplace, P.S. de (1785) Théorie des Attractions des Sphéroides et de la Figure des Planètes, Mém. de l'Acad., Paris

[62] Legendre, A.M. (1785) Recherches sur l'Attraction des Sphéroides Homogènes, Mém. math. phys. prés. à l'Acad. Aci. par divers savants, 10, pp. 411–434

[63] Leistedt, B., McEwen, J.D. (2012) Exact Wavelets on the Ball, IEEE Trans. Signal Process., 60, pp. 6257–6269

[64] Lyche, T., Schumaker, L. (2000) A Multiresolution Tensor Spline Method for Fitting Functions on the Sphere, SIAM J. Sci. Comput., 22, pp. 724–746

[65] Mallat, S.G., Zhang, Z. (1993) Matching Pursuits with Time-Frequency Dictionaries, IEEE Trans. Signal Process., 41, pp. 3397–3415

[66] Maxwell, J.C. (1891) A Treatise on Electricity and Magnetism (1873, 1881, 1891) Bd. 1 und 2 Ungekürzter Nachdruck der letzten Auflage (1954), Dover, New York

[67] Michel, V. (2013) Lectures on Constructive Approximation – Fourier, Spline, and Wavelet Methods on the Real Line, the Sphere, and the Ball, Birkhäuser, Boston

[68] Michel, V. (2015a) Tomography – Problems and Multiscale Solutions, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 2nd ed., Springer, Berlin, Heidelberg, pp. 2087–2119

[69] Michel, V. (2015b) RFMP – An Iterative Best Basis Algorithm for Inverse Problems in the Geosciences, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 2nd ed., Springer, Berlin, Heidelberg, pp. 2121–2147

[70] Michel, V., Orzlowski, S. (2016) On the Null Space of a Class of Fredholm Integral Equations of the First Kind, J. Inverse and Ill-Posed Problems, 24, pp. 687–710

[71] Michel, V., Orzlowski, S. (2017) On the Convergence Theorem for the Regularized Functional Matching Pursuit (RFMP) Algorithm, Int. J. Geomath., 8, pp. 183–190

[72] Michel, V., Telschow, R. (2014) A Non-linear Approximation Method on the Sphere, Int. J. Geomath., 5, pp. 195–224

[73] Michel, V., Telschow, R. (2016) The Regularized Orthogonal Functional Matching Pursuit for Ill-posed Inverse Problems, SIAM J. Numer. Anal., 54, pp. 262–287

[74] Michel, V., Wolf, K. (2008) Numerical Aspects of a Spline-based Multiresolution Recovery of the Harmonic Mass Density out of Gravity Functionals, Geophys. J. Int., 173, pp. 1–16

[75] Morse, P.M., Feshbach, H. (1953) Methods of Theoretical Physics, McGraw-Hill, New York

[76] Narcowich, F.J., Ward, J.D. (1996) Nonstationary Wavelets on the m-Sphere for Scattered Data, Appl. Comput. Harmon. Anal., 3, pp. 324–336

[77] Neumann, F. (1887) Vorlesungen über die Theorie des Potentials und der Kugelfunktionen, Teubner, Leipzig, pp. 135–154

[78] Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S. (1993) Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition, in: Asilomar Conference on Signals, Systems and Computers, IEEE Conf. Pub., 1, pp. 40–44

[79] Plattner, A., Simons, F.J. (2014) Spatiospectral Concentration of Vector Fields on a Sphere, Appl. Comput. Harmon. Anal., 36, pp. 1–22

[80] Plattner, A., Simons, F.J. (2015a) Potential-Field Estimation Using Scalar and Vector Slepian Functions at Satellite Altitude, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 2nd ed., Springer, Berlin, Heidelberg, pp. 2003–2055

[81] Plattner, A., Simons, F.J. (2015b) High-Resolution Local Magnetic Field Models for the Martian South Pole From Mars Global Surveyor Data, J. Geophys. Res., pp. 1543–1566

[82] Potts, D., Tasche, M. (1995) Interpolatory Wavelets on the Sphere, in: Approximation Theory VIII (C.K. Chui, L.L. Schumaker, Eds.), World Scientific, Singapore, pp. 335–342

[83] Schröder, P., Sweldens, W. (1995) Spherical Wavelets: Efficiently Representing Functions on the Sphere, in: Computer Graphics Proceedings (SIGGRAPH95), pp. 161–175

[84] Simons, F.J. (2010) Slepian Functions and Their Use in Signal Estimation and Spectral Analysis, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 1st ed., Springer, Heidelberg, pp. 891–923

[85] Simons, F.J., Dahlen, F.A. (2006) Spherical Slepian Functions and the Polar Gap in Geodesy, Geoph. J. Int., 166, pp. 1039–1061

[86] Simons, F.J., Plattner, A. (2015) Scalar and Vector Slepian Functions, Spherical Signal Estimation and Spectral Analysis, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 2nd ed., Springer, Berlin, Heidelberg, pp. 2563–2608

[87] Simons, F.J., Dahlen, F.A., Wieczorek, M. (2006) Spatiospectral Concentration on a Sphere, SIAM Rev., 48, pp. 504–536

[88] Simons, F.J., Hawthorne, J.C., Beggan, C.D. (2009) Efficient Analysis and Representation of Geophysical Processes Using Localized Spherical Basis Functions, in: Wavelets XIII, Proc. SPIE, (V. Goyal, M. Papadakis and D. Van de Ville, Eds.), 7446(74460G)

[89] Simons, F.J., Loris, I., Brevdo, E., Daubechies, I.C. (2011) Wavelets and Wavelet-like Transforms on the Sphere and Their Application to Geophysical Data Inversion, in: Wavelets and Sparsity XIV, Proc. SPIE, (M. Papadakis, V. Goyal, and D. Van de Ville, Eds.), 8138(81380X)

[90] Slepian, D. (1983) Some Comments on Fourier Analysis, Uncertainty and Modeling, SIAM Rev., 25, pp. 379–393

[91] Slobbe, D.C., Simons, F.J., Klees, R. (2012) The Spherical Slepian Basis as a Means to Obtain Spectral Consistency Between Mean Sea Level and the Geoid, J. Geod., 86, pp. 609–628

[92] Svensson, S.L. (1983) Pseudodifferential Operators – A New Approach to the Boundary Value Problems of Physical Geodesy, Manuscr. Geod., 8, pp. 1–40

[93] Sylvester, T. (1876) Note on Spherical Harmonics, Phil. Mag., II, 291 and 400

[94] Tegmark, M. (1996) A Method for Extracting Maximum Resolution Power Spectra from Microwave Sky Maps, Mon. Not. Roy. Astron. Soc., 280, pp. 299–308

[95] Telschow, R. (2015) An Orthogonal Matching Pursuit for the Regularization of Spherical Inverse Problems, Ph.D.-thesis, Geomathematics Group, Dept. Mathematics, University of Siegen, Verlag Dr. Hut, Munich

[96] Temlyakov, V.N. (2003) Nonlinear Methods of Approximation, Found. Comput. Math. 3, pp. 33–107

[97] Tscherning, C.C. (1996) Isotropic Reproducing Kernels for the Inner of a Sphere or Spherical Shell and Their Use as Density Covariance Functions. Math. Geol. 28, pp. 161–168

[98] Vincent, P., Bengio, Y. (2002) Kernel Matching Pursuit, Mach. Learn., 48, pp. 169–191

[99] Wahba, G. (1981) Spline Interpolation and Smoothing on the Sphere, SIAM J. Sci. Stat. Comput., 2, pp. 5–16, (also errata: SIAM J. Sci. Stat. Comput., 3, pp. 385–386)

[100] Wahba, G. (1990) Spline Models for Observational Data, CBMS-NSF Regional Conference Series in Applied Mathematics, 59, SIAM, Philadelphia

[101] Weinreich, I. (2001) A Construction of $C^{(1)}$-Wavelets on the Two-dimensional Sphere, Appl. Comput. Harmon. Anal., 10, pp. 1–26

[102] Wicht, J., Stellmach, S., Harder, H. (2010) Numerical Dynamo Simulations: From Basic Concepts to Realistic Models, in: Handbook of Geomathematics (W. Freeden, M.Z. Nashed, and T. Sonar, Eds.), 1st ed., pp. 459–502, Springer, Heidelberg

[103] Wieczorek, M.A. (2015) The Gravity and Topography of the Terrestrial Planets, in: Treatise on Geophysics (T. Spohn, ed.), 2nd ed., vol. 10, pp. 153–2193, Elsevier, Amsterdam

[104] Wieczorek, M.A., Simons, F.J. (2005) Localized Spectral Analysis on the Sphere, Geophys. J. Int., 162, pp. 655–675

Willi Freeden
Geomathematics Group
University of Kaiserslautern
MPI-Gebäude, Paul-Ehrlich-Str. 26
D-67663 Kaiserslautern, Germany

Volker Michel
Geomathematics Group
University of Siegen
Walter-Flex-Str. 3
D-57068 Siegen, Germany

Frederik J. Simons
Princeton University
Department of Geosciences
Princeton, NJ, USA