Spectral and Multiscale Signal-To-Noise Thresholding of Spherical Scalar Fields

W. Freeden, T. Maier

University of Kaiserslautern
Geomathematics Group
67653 Kaiserslautern
P.O. Box 3049
Germany

phone: ++49 631 205-3867
fax: ++49 631 29081
email: freeden@mathematik.uni-kl.de
tmaier@mathematik.uni-kl.de
www: http://www.mathematik.uni-kl.de/~wwwgeo

Abstract. The basic concepts of selective multiscale reconstruction of functions on the sphere from error–affected data is outlined for scalar functions. The selective reconstruction mechanism is based on the premise that multiscale approximation can be well–represented in terms of only a relatively small number of expansion coefficients at various resolution levels. A new pyramid scheme is presented to efficiently remove the noise at different scales using a priori statistical information.

Correspondence to W. Freeden

1. Introduction. While standard Fourier methods in terms of spherical harmonics are very successful at picking out frequencies from a spherical signal, they are utterly incapable of dealing properly with data changing on small spatial scales. This fact has been well–known for years. To improve the applicability of Fourier analysis, various methods such as 'windowed Fourier transform' have been developed on the sphere to modify the usual Fourier procedure to allow analysis of the frequency content of a signal at each position (cf. W. Freeden et al. (1998), W. Freeden, V. Michel (1999)). However, the amount of localization in space and in frequency is not completely satisfactory in Fourier transform and its windowed variant. For example, geopotentials refer to a certain combination of frequencies, and the frequencies themselves are spatially changing. This space evolution of the frequencies is not reflected in the Fourier transform in terms of non-space-localizing spherical harmonics. Even the windowed Fourier transform contains information about frequencies over a certain area of positions instead of showing how the frequencies vary in space. With spherical wavelets, the amount of localization in space and in frequency is automatically adapted, in that only a narrow space–window is needed to examine high–frequency content, but a wide space–window is allowed when investigating low–frequency phenomena. The basic framework of this approach has been provided by the spherical wavelet theory developed by the Geomathematics Group at the University of Kaiserslautern during the last years (see http://www.mathematik.uni-kl.de/~wwwgeo/pub1.html)

When dealing with real geophysically relevant data it should be kept in mind that each measurement does not really give the value of the observable under consideration but that – at least to some extend – the data are contaminated with noise. That is, in order to succesfully improve geomathematical field modeling, one main aspect is to extract the true portion of the observable from the actual signal. In consequence, a particular emphasis lies on the subject of denoising. This endeavor is precisely the goal in statistical function estimation. Here, the interest is to 'smooth' the noisy data in order to obtain an estimate of the underlying function. In Euclidean theory of wavelets signal processors now have new, fast tools that are well–suited for denoising signals (for a survey the reader is e.g. referred to T.R. Ogden (1997) and the references therein).

The objective of this article is to introduce multiscale signal-to-noise thresholding, providing the wavelet oriented basis of denoising spherical data sets. First, we develop the corresponding theory of denoising spherical functions (cf. W. Freeden et al. (2000)). With the basic introduction at hand, selective thresholding within a pyramid scheme is presented. The thresholding scheme is designed to distinguish between coefficients which contribute signficantly to the signal, and those which are negligible. It should be noted that our approach is essentially influenced by the concept of sparse wavelet representations in Euclidean spaces (cf. J.B. Weaver et al. (1991), D.L. Donoho (1994) and I.M Johnstone (1995)) and the stochastic model used in satellite geodesy (see e.g. R. Rummel (1997)). Using a multiscale approach we are thus able to include detail information of small spatial extend while suppressing the noise in the approximation appropriately. A simple example of denoising geomagnetic field data will be given as an illustration.

2. Preliminaries. $\mathbb{R}^3$ denotes three–dimensional Euclidean space. For $x, y \in \mathbb{R}^3$, $x = (x_1, x_2, x_3)^T, y = (y_1, y_2, y_3)^T$ the inner product is defined as usual by

$$x \cdot y = x^T y = \sum_{i=1}^{3} x_i y_i$$

For all elements $x \in \mathbb{R}^3$, $x = (x_1, x_2, x_3)^T$, different from the origin, we have

$$x = r\xi, \quad r = |x| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + x_3^2},$$

where $\xi = (\xi_1, \xi_2, \xi_3)^T$ is the uniquely determined directional unit vector of x. The unit sphere in $\mathbb{R}^3$ is denoted by Ω. If the vectors $\varepsilon^1, \varepsilon^2, \varepsilon^3$ form the canonical orthonormal basis in $\mathbb{R}^3$, the points $\xi \in \Omega$ may be represented in polar coordinates by

$$\begin{aligned} &\xi = t\varepsilon^3 + \sqrt{1 - t^2}\left(\cos\varphi\varepsilon^1 + \sin\varphi\varepsilon^2\right), \\ &t = \cos\vartheta, \vartheta \in [0, \pi], \varphi \in [0, 2\pi) \ . \end{aligned}$$

3. Spherical Harmonics. The *spherical harmonics* Y_n of degree n are defined as the everywhere on Ω infinitely differentiable eigenfunctions of the Beltrami operator Δ^* corresponding to the eigenvalues $(\Delta^*)^\wedge(n) = -n(n+1)$, $n = 0, 1, \ldots$, where the Beltrami-operator is the angular part of the Laplace-operator Δ in $\mathbb{R}^3$. As it is well–known, the functions $H_n : \mathbb{R}^3 \to \mathbb{R}$ defined by $H_n(x) = r^n Y_n(\xi)$, $x = r\xi$, $\xi \in \Omega$, are homogeneous polynomials in rectangular coordinates which satisfy the Laplace-equation $\Delta_x H_n(x) = 0$, $x \in \mathbb{R}^3$. Conversely, every homogeneous harmonic polynomial of degree n when restricted to Ω is a spherical harmonic of degree n. The *Legendre polynomials* $P_n : [-1, +1] \to [-1, +1]$ are the only everywhere on $[-1, +1]$ infinitely differentiable eigenfunctions of the Legendre-operator $(1-t^2)(d/dt)^2 - 2t(d/dt)$, which satisfy $P_n(1) = 1$. Apart from a multiplicative constant, the 'ε^3-Legendre function' $P_n(\varepsilon^3 \cdot) : \Omega \to [-1, +1]$, $\xi \mapsto P_n(\varepsilon^3 \cdot \xi)$, $\xi \in \Omega$, is the only spherical harmonic of degree n which is invariant under orthogonal transformations leaving ε^3 fixed. The linear space $Harm_n$ of all spherical harmonics of degree n is of dimension $\dim(Harm_n) = 2n + 1$. Thus, there exist $2n + 1$ linearly independent spherical harmonics $Y_{n,1}, \ldots, Y_{n,2n+1}$ in $Harm_n$. Throughout the remainder of this paper we assume this system to be orthonormal in the sense of the $\mathcal{L}^2(\Omega)$–inner product

$$(Y_{n,j}, Y_{m,k})_{\mathcal{L}^2(\Omega)} = \int_\Omega Y_{n,j}(\eta) Y_{m,k}(\eta) \ d\omega(\eta) = \delta_{n,m}\delta_{j,k}$$

($d\omega$ denotes the surface element). An outstanding result of the theory of spherical harmonics is the *addition theorem*

$$\sum_{k=1}^{2n+1} Y_{n,k}(\xi) Y_{n,k}(\eta) = \frac{2n+1}{4\pi} P_n(\xi \cdot \eta), \quad (\xi, \eta) \in \Omega \times \Omega \ .$$

The close connection between the orthogonal invariance and the addition theorem is established by the *Funk–Hecke formula*

$$\int_\Omega H(\xi \cdot \eta) P_n(\zeta \cdot \eta) \ d\omega(\eta) = (H(\xi \cdot), P_n(\zeta \cdot))_{\mathcal{L}^2(\Omega)} = H^\wedge(n) P_n(\xi \cdot \zeta),$$

$H \in \mathcal{L}^1[-1,+1]$, $\xi, \zeta \in \Omega$, where the *Legendre transform* $LT : H \to (LT)(H), H \in \mathcal{L}^1[-1,1]$, is given by

$$(LT)(H)(n) = H^\wedge(n) = 2\pi \int_{-1}^{+1} H(t)P_n(t)\ dt, \quad n = 0, 1, \dots \ .$$

The sequence $\{H^\wedge(n)\}_{n\in\mathbb{N}_0}$ is called the *symbol* of H. For more details about the theory of spherical harmonics the reader is referred, for example, to C. Müller (1966) and W. Freeden et al.(1998).

We let

$$Harm_{0,\dots,m} = \operatorname*{span}_{\substack{n=0,\dots,m \\ k=1,\dots,2n+1}} (Y_{n,k}) \, .$$

Of course,

$$Harm_{0,\dots,m} = \bigoplus_{n=0}^{\infty} Harm_n$$

so that

$$dim(Harm_{0,\dots,m}) = \sum_{n=0}^{m} (2n+1) = (m+1)^2 .$$

As it is well known, $K_{Harm_{0,\dots,m}} : \Omega \times \Omega \to \mathbb{R}$ defined by

$$K_{Harm_{0,\dots,m}}(\xi,\eta) = \sum_{n=0}^{m} \sum_{k=1}^{2n+1} Y_{n,k}(\xi)Y_{n,k}(\eta) = \sum_{n=0}^{m} \frac{2n+1}{4\pi} P_n(\xi\cdot\eta) \tag{3.1}$$

is the reproducing kernel of the space $Harm_{0,\dots,m}$ with respect to $(\cdot,\cdot)_{\mathcal{L}^2(\Omega)}$. Moreover it is worth mentioning that

$$\begin{aligned} &\int_\Omega K_{Harm_{0,\dots,m+1}}(\xi,\eta)Y(\eta)\ d\omega(\eta) \\ =&\int_\Omega K_{Harm_{0,\dots,m}}(\xi,\eta)Y(\eta)\ d\omega(\eta) \\ =&\, Y(\xi) \end{aligned} \tag{3.2}$$

for all $\xi \in \Omega$ and all $Y \in Harm_{0,\dots,m}$.

In what follows we are mainly interested in results for the Hilbert space $\mathcal{L}^2(\Omega)$ equipped with the inner product $(\cdot,\cdot)_{\mathcal{L}^2(\Omega)}$. Any function of class $\mathcal{L}^2(\Omega)$ of the form $H_\xi : \Omega \to \mathbb{R}$, $\eta \mapsto H_\xi(\eta) = H(\xi\cdot\eta)$, $\eta \in \Omega$, is called a *ξ–zonal function* on Ω. Zonal functions are constant on the sets of all $\eta \in \Omega$, with $\xi\cdot\eta = h$, $h \in [-1,+1]$. The set of all ξ–zonal functions is isomorphic to the set of functions $H : [-1,+1] \to \mathbb{R}$. This gives rise to interpret the space $\mathcal{L}^2[-1,+1]$ with norm defined correspondingly by

$$\|H\|_{\mathcal{L}^2[-1,+1]} = \left(2\pi \int_{-1}^{+1} |H(t)|^2\ dt \right)^{1/2} = \|H(\varepsilon^3\cdot)\|_{\mathcal{L}^2(\Omega)}, \quad H \in \mathcal{L}^2[-1,+1] \ .$$

as subspace of $\mathcal{L}^2(\Omega)$.

The *spherical Fourier transform* $H \mapsto (FT)(H)$, $H \in \mathcal{L}^2(\Omega)$, is given by

$$((FT)(H))\,(n,k) = H^\wedge(n,k) = (H, Y_{n,k})_{\mathcal{L}^2(\Omega)}, \quad n = 0,1,\dots; k = 1,\dots,2n+1.$$

FT forms a mapping from $\mathcal{L}^2(\Omega)$ onto the space $l^2(\mathcal{N})$ of all sequences $\{W_{n,k}\}_{(n,k)\in\mathcal{N}}$ satisfying

$$\sum_{(n,k)\in\mathcal{N}} W_{n,k}^2 = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} W_{n,k}^2 < \infty,$$

where we have used the abbreviation

$$\mathcal{N} = \{(n,k) | n = 0,1,\dots; k = 1,\dots,2n+1\} \quad .$$

The series

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

is called the *spherical Fourier expansion* of F (with Fourier coefficients $F^\wedge(n,k)$, $(n,k) \in \mathcal{N}$). For all $F \in \mathcal{L}^2(\Omega)$ we have

$$\lim_{N\to\infty} \left\| F - \sum_{n=0}^{N}\sum_{k=1}^{2n+1} F^\wedge(n,k) Y_{n,k} \right\|_{\mathcal{L}^2(\Omega)} = 0.$$

4. Convolutions. A kernel $H : \Omega \times \Omega \to \mathbb{R}$ is called a *square-summable product kernel* if H is of the form

$$H(\xi,\eta) = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} H^\wedge(n,k) Y_{n,k}(\xi) Y_{n,k}(\eta)$$

such that

$$\int_\Omega (H(\xi,\eta))^2 \, d\omega(\eta) \le \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \sup_{k=1,\dots,2n+1} \left(H^\wedge(n,k)\right)^2 < \infty. \tag{4.1}$$

In the case of rotation invariance of the kernel H, i.e. $H^\wedge(n,k) = H^\wedge(n)$ for $n = 0,1,\dots, k = 1,\dots,2n+1$, the last condition is equivalent to the $l^2(\mathcal{N})$-summability (cf. W. Freeden et al. 1998), i.e.

$$\sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \left(H^\wedge(n)\right)^2 < \infty.$$

Assume that H is a square-summable product kernel and $F \in \mathcal{L}^2(\Omega)$. Then the *convolution of H against F* is defined by

$$H * F = \int_\Omega H(\cdot,\eta)\, F(\eta) \; d\omega(\eta).$$

Two important properties of spherical convolutions should be listed: (i) If $F \in \mathcal{L}^2(\Omega)$ and H is a square summable product kernel, then $H * F$ is of class $\mathcal{L}^2(\Omega)$. (ii) If

H_1, H_2 are square-summable product kernels, then the convolution of H_1, H_2 defined by

$$(H_1 * H_2)(\xi \cdot \zeta) = \int_\Omega H_1(\xi \cdot \eta)\, H_2(\eta \cdot \zeta)\ d\omega(\eta)$$

is a square-summable product kernel with

$$(H_1 * H_2)^\wedge(n,k) = H_1^\wedge(n,k)\ H_2^\wedge(n,k), \quad (n,k) \in \mathcal{N}\ .$$

We usually write $H^{(2)} = H * H$ to indicate the convolution of H with itself. $H^{(2)}$ is said to be the *(second) iterated kernel of* H. More general, $H^{(p)} = H^{(p-1)} * H$ for $p = 2, 3, \ldots$ and $H^{(1)} = H$. Obviously,

$$\left(H^{(p)}\right)^\wedge(n,k) = \left(H^\wedge(n,k)\right)^p, \quad (n,k) \in \mathcal{N}, \quad p \in \mathbb{N}\ .$$

5. Multiscale Approximation. Next we consider a strict monotonically decreasing sequence $\{\rho_j\}_{j\in\mathbb{Z}}$ of real numbers satisfying

$$\lim_{j\to\infty} \rho_j = 0$$

and

$$\lim_{j\to-\infty} \rho_j = \infty$$

(for example, $\rho_j = 2^{-j}$, $j \in \mathbb{Z}$). The sequence $\{\rho_j\}_{j\in\mathbb{Z}}$ can be understood as a subdivision of the 'scale interval' $(0, \infty)$ into a countable, strict monotonically decreasing sequence.

Let $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ be a family of square-summable product kernels satisfying the condition $\Phi^\wedge_{\rho_j}(0,1) = 1$ for all $j \in \mathbb{Z}$. Then, the family $\{I_{\rho_j}\}_{j\in\mathbb{Z}}$ of operators I_{ρ_j}, defined by $I_{\rho_j}(F) = \Phi_{\rho_j} * F$, $\quad F \in \mathcal{L}^2(\Omega)$, is called a *singular integral in* $\mathcal{L}^2(\Omega)$. $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is called *kernel of the singular integral.* If $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is a kernel of a singular integral satisfying the conditions:

(i) $\lim\limits_{j\to\infty} \left(\Phi^\wedge_{\rho_j}(n,k)\right)^2 = 1$ for all $(n,k) \in \mathcal{N}$,

(ii) $\left(\Phi^\wedge_{\rho_{j+1}}(n,k)\right)^2 \geq \left(\Phi^\wedge_{\rho_j}(n,k)\right)^2$ for all $j \in \mathbb{Z}$ and $(n,k) \in \mathcal{N}$,

(iii) $\lim\limits_{j\to-\infty} \left(\Phi^\wedge_{\rho_j}(n,k)\right)^2 = 0$ for all $(n,k) \in \mathcal{N}$,

then the corresponding singular integral $\left\{I^{(2)}_{\rho_j}\right\}_{j\in\mathbb{Z}}$ with

$$I^{(2)}_{\rho_j} = \Phi^{(2)}_{\rho_j} * F, \quad j \in \mathbb{Z}$$

is called an *approximate identity in* $\mathcal{L}^2(\Omega)$. It is known (see e.g. W. Freeden et al. (1998)) that

$$\begin{aligned}
&\lim_{j\to\infty} \|I^{(2)}_{\rho_j}(F) - F\|_{\mathcal{L}^2(\Omega)} \\
&= \lim_{j\to\infty} \left(\sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \left(F^\wedge(n,k)\right)^2 \left(1 - \Phi^\wedge_{\rho_j}(n,k)\right)^4 \right)^{1/2} \\
&= 0\ ,
\end{aligned}$$

provided that $\left\{I^{(2)}_{\rho_j}\right\}_{j\in\mathbb{Z}}$ is an approximate identity.

Our results immediately lead us to the following statement:

LEMMA 5.1. *Assume that* $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ *is a kernel constituting an approximate identity in* $\mathcal{L}^2(\Omega)$. *Then the limit relation*

$$\begin{aligned}
&\lim_{j\to\infty}\left\|\int_\Omega \Phi^{(2)}_{\rho_j}(\cdot,\eta)F(\eta)d\omega(\eta)-F\right\|_{\mathcal{L}^2(\Omega)}\\
&=\lim_{j\to\infty}\left\|\int_\Omega\int_\Omega \Phi_{\rho_j}(\eta,\zeta)F(\zeta)d\omega(\zeta)\Phi_{\rho_j}(\cdot,\eta)d\omega(\eta)-F\right\|_{\mathcal{L}^2(\Omega)}=0
\end{aligned}$$

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

For $J\in\mathbb{Z}$ we set

$$F_J=\Phi^{(2)}_{\rho_J}*F=\int_\Omega \Phi^{(2)}_{\rho_J}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)\;\;.$$

Consider a kernel $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ constituting an approximate identity in $\mathcal{L}^2(\Omega)$. Assume that F is of class $\mathcal{L}^2(\Omega)$. Then a simple calculation shows us that for all $N\in\mathbb{N}$ and $J\in\mathbb{Z}$,

$$\begin{aligned}
\int_\Omega \Phi^{(2)}_{\rho_{J+N}}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)&=\int_\Omega \Phi^{(2)}_{\rho_J}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)\\
&+\sum_{j=J}^{J+N-1}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)F(\eta)\;\;d\omega(\eta),
\end{aligned}\tag{5.1}$$

where we have introduced the family $\{\Psi_{\rho_j}\}_{j\in\mathbb{Z}}$ by the *spectral refinement condition*

$$\Psi^\wedge_{\rho_j}(n,k)=\left(\left(\Phi^\wedge_{\rho_{j+1}}(n,k)\right)^2-\left(\Phi^\wedge_{\rho_j}(n,k)\right)^2\right)^{1/2},\quad (n,k)\in\mathcal{N}.\tag{5.2}$$

In other words,

$$\Psi^{(2)}_{\rho_j}(\xi,\eta)=\Phi^{(2)}_{\rho_{j+1}}(\xi,\eta)-\Phi^{(2)}_{\rho_j}(\xi,\eta),\tag{5.3}$$

$j\in\mathbb{Z},(\xi,\eta)\in\Omega\times\Omega$. Hence, letting N tend to infinity we get the following *multiscale reconstruction formula*

$$F=F_J+\sum_{j=J}^{\infty}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)\tag{5.4}$$

for every $J\in\mathbb{Z}$ (in the sense of $\|\cdot\|_{\mathcal{L}^2(\Omega)}$). Moreover, we find

$$\int_\Omega \Phi^{(2)}_{\rho_J}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)=F_{J-N}+\sum_{j=J-N}^{J-1}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)F(\eta)\;\;d\omega(\eta),$$

hence,

$$\int_\Omega \Phi^{(2)}_{\rho_J}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)=\sum_{j=-\infty}^{J-1}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)F(\eta)\;\;d\omega(\eta)\;\;.\tag{5.5}$$

Combining (5.4) and (5.5) we finally obtain the following result:

LEMMA 5.2. *The multiscale representation of* $F \in \mathcal{L}^2(\Omega)$

$$F = \sum_{j=-\infty}^{\infty} \int_{\Omega} \Psi^{(2)}_{\rho_j}(\cdot,\eta) F(\eta) \ d\omega(\eta),$$

holds in the sense of $\|\cdot\|_{\mathcal{L}^2(\Omega)}$ *provided that the so–called 'scaling function'* $\{\Phi^{(2)}_{\rho_j}\}_{j\in\mathbb{Z}}$ *forms an approximate identity in* $\mathcal{L}^2(\Omega)$ *and the 'wavelet'* $\{\Psi^{(2)}_{\rho_j}\}_{j\in\mathbb{Z}}$ *satisfies the difference equation (5.3).*

By construction, the wavelet theory leads to a partition of unity as follows

$$\sum_{j=-\infty}^{\infty} \left(\Psi^{(2)}_{\rho_j}\right)^{\wedge}(n,k) = \left(\Phi^{(2)}_{\rho_J}\right)^{\wedge}(n,k) + \sum_{j=J}^{\infty} \left(\Psi^{(2)}_{\rho_j}\right)^{\wedge}(n,k) = 1$$

for all $(n,k) \in \mathcal{N}$. The class $\mathcal{V}_{\rho_j}$ of all functions $P \in \mathcal{L}^2(\Omega)$ of the form

$$P = \Phi^{(2)}_{\rho_j} * F, \quad F \in \mathcal{L}^2(\Omega),$$

is called the *scale space of level* j (with respect to the scaling function $\{\Phi^{(2)}_{\rho_j}\}_{j\in\mathbb{Z}}$), whereas the class $\mathcal{W}_{\rho_j}$ of all functions $Q \in \mathcal{L}^2(\Omega)$ of the representation

$$Q = \Psi^{(2)}_{\rho_j} * F, \quad F \in \mathcal{L}^2(\Omega),$$

is called the *detail space of level* j (with respect to the scaling function $\{\Phi^{(2)}_{\rho_j}\}_{j\in\mathbb{Z}}$). It is easily seen from (5.1) that

$$\mathcal{V}_{\rho_{j+1}} = \mathcal{V}_{\rho_j} + \mathcal{W}_{\rho_j} \tag{5.6}$$

for all $j \in \mathbb{Z}$. But it should be remarked that the sum (5.6) generally is neither direct nor orthogonal (note that an orthogonal decomposition is given by the Shannon scaling function). The equation (5.6) can be interpreted in the following way: The set $\mathcal{V}_{\rho_j}$ contains a filtered ('smoothed') version of a function belonging to $\mathcal{L}^2(\Omega)$. The lower the scale, the stronger the intensity of smoothing. By adding 'details' contained in the detail space $\mathcal{W}_{\rho_j}$ the space $\mathcal{V}_{\rho_{j+1}}$ is created, which consists of a filtered ('smoothed') version at resolution $j+1$ (see W. FREEDEN et al. (1998) for more details of spherical theory, W. FREEDEN (1999) for the application harmonic theory and V. MICHEL (1999) for application in gravimetry).

Finally, it is worth mentioning that the scale spaces satisfy the following properties:

(i) $\mathcal{V}_{\rho_j} \subset \mathcal{V}_{\rho_{j'}} \subset \ldots \subset \mathcal{L}^2(\Omega), \quad j \leq j'$

(ii) $\overline{\bigcup_{j=-\infty}^{\infty} \mathcal{V}_{\rho_j}}^{\|\cdot\|_{\mathcal{L}^2(\Omega)}} = \mathcal{L}^2(\Omega).$

A collection of subspaces of $\mathcal{L}^2(\Omega)$ satisfying (i) and (ii) is called a *multiresolution analysis* of $\mathcal{L}^2(\Omega)$.

6. Examples. Singular integrals on the sphere are of basic interest in geomathematical applications. We essentially distinguish two types, namely bandlimited and non–bandlimited singular integrals.

6.1. Bandlimited Singular Integrals.

Shannon Singular Integral. The family $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is defined by

$$\Phi^{\wedge}_{\rho_j}(n,k) = \Phi^{\wedge}_{\rho_j}(n) = \begin{cases} 1 & \text{for} \quad n \in [0,\rho_j^{-1}), k = 1,\dots,2n+1 \\ 0 & \text{for} \quad n \in [\rho_j^{-1},\infty), k = 1,\dots,2n+1 \end{cases}$$

with a strict monotonically decreasing sequence of integers $\{\rho_j\}_{j\in\mathbb{Z}}$ satisfying

$$\lim_{j\to-\infty} \rho_j = \infty \quad \text{ and } \lim_{j\to\infty} \rho_j = 0$$

(for example: $\rho_j = 2^{-j}$).

Smoothed Shannon Singular Integral. The family $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is given by

$$\Phi^{\wedge}_{\rho_j}(n,k) = \Phi^{\wedge}_{\rho_j}(n) = \begin{cases} 1 & \text{for} \quad n \in [0,\sigma_j^{-1}), k = 1,\dots,2n+1 \\ \tau_j(n) & \text{for} \quad n \in [\sigma_j^{-1},\rho_j^{-1}), k = 1,\dots,2n+1 \\ 0 & \text{for} \quad n \in [\rho_j^{-1},\infty), k = 1,\dots,2n+1. \end{cases}$$

where $\{\rho_j\}_{j\in\mathbb{Z}}$ is defined as in the Shannon case and $\{\sigma_j\}_{j\in\mathbb{Z}}$ is a strict monotonically decreasing sequence of integers satisfying

$$\lim_{j\to-\infty} \sigma_j = \infty \ , \ \lim_{j\to\infty} \sigma_j = 0,$$
$$\sigma_j > \rho_j,$$

and τ_j is a strict monotonically decreasing and continuous function of class $C[\sigma_j^{-1},\rho_j^{-1}]$, $j \geq 0$, such that

$$\tau_j(\sigma_j^{-1}) = 1, \ \tau_j(\rho_j^{-1}) = 0,$$

for example $\tau_j(t) = 2 - 2^{-j}t$ with $\rho_j = 2^{-j-1}$ and $\sigma_j = 2^{-j}$.

6.2. Non–bandlimited Singular Integrals.

Abel–Poisson Singular Integral. The family $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is given by

$$\Phi^{\wedge}_{\rho_j}(n,k) = \Phi^{\wedge}_{\rho_j}(n) = e^{-n\rho_j}, \quad (n,k) \in \mathcal{N}, j \in \mathbb{Z} \ .$$

Tikhonov Singular Integral. The family $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ is given by

$$\Phi^{\wedge}_{\rho_j}(n,k) = \frac{\sigma^2_{n,k}}{\sigma^2_{n,k} + \rho^2_j}, \qquad (n,k) \in \mathcal{N}, j \in \mathbb{Z}, \tag{6.1}$$

where $\{\sigma_{n,k}\}_{(n,k)\in\mathcal{N}}$ is a sequence satisfying the following conditions:

(i) $\sigma_{n,k} \neq 0$ for all $(n,k) \in \mathcal{N}$,

(ii) $\{\sigma_{n,k}\}_{(n,k)\in\mathcal{N}}$ is $l^2(\mathcal{N})$-summable, i.e.

$$\sum_{(n,k)\in\mathcal{N}} \sigma^2_{n,k} = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \sigma^2_{n,k} < \infty \ .$$

7. Spectral Signal–to–Noise Response. Geoscientists mostly think of their measurements (after possible linearization) as a linear operator on an 'input signal' F producing an 'output signal' G

$$\Lambda F = G, \tag{7.1}$$

where Λ is an operator mapping the space $\mathcal{L}^2(\Omega)$ into itself such that

$$\Lambda Y_{n,k} = \Lambda^\wedge(n,k) Y_{n,k}, \quad (n,k) \in \mathcal{N},$$

where the so–called *symbol* $\{\Lambda^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$ is the sequence of the real numbers $\Lambda^\wedge(n,k)$. Different linear operators Λ, of course, are characterized by different sequences $\{\Lambda^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$. The 'amplitude spectrum' $\{G^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$ of the response of Λ is described in terms of the amplitude spectrum of functions (signals) by a simple multiplication by the 'transfer' $\Lambda^\wedge(n,k)$. For a large number of problems in geophysics and geodesy Λ is a rotation-invariant operator, i.e. $\Lambda^\wedge(n,k) = \Lambda^\wedge(n)$ for all $(n,k) \in \mathcal{N}$.

7.1. Noise Model. Thus far only a (deterministic) function model has been discussed. If a comparison of the 'output function' with the actual value were done, discrepancies would be observed. A mathematical description of these discrepancies has to follow the laws of probability theory in a stochastic model (see e.g. R.T. OGDON (1997)).

Usually the observations are not looked upon as a time series, but rather a function $\tilde{G}$ on the sphere Ω ('$\sim$' for stochastic). According to this approach we assume that, we have

$$\tilde{G} = G + \tilde{\varepsilon},$$

where $\tilde{\varepsilon}$ is the *observation noise*. Moreover, in our approach motivated by information in satellite technology, we suppose the *covariance* to be known

$$Cov\left[\tilde{G}(\xi), \tilde{G}(\eta)\right] = E\left[\tilde{\varepsilon}(\xi), \tilde{\varepsilon}(\eta)\right] = K(\xi,\eta), \qquad (\xi,\eta) \in \Omega \times \Omega,$$

where the following conditions are imposed on the symbol $\{K^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$ of the kernel function $K : \Omega \times \Omega \to \mathbb{R}$:

(C1) $K^\wedge(n,k) \geq 0$ for all $(n,k) \in \mathcal{N}$,

(C2) $\sum\limits_{n=0}^{\infty} \frac{2n+1}{4\pi} \sup\limits_{k=1,\ldots,2n+1} \left(K^\wedge(n,k)\right)^2 < \infty.$

Condition (C2) , indeed, implies in the case of rotation-invariance, i.e.

$$K^\wedge(n,k) = K^\wedge(n), \ \ n = 0,1,\ldots, \ \ k = 1,\ldots,2n+1,$$

the $l^{(2)}(\mathcal{N})$-summability of the symbol $\{K^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$, i.e.

$$\sum_{(n,k)\in\mathcal{N}} K^\wedge(n,k) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} (K^\wedge(n))^2 < \infty.$$

7.2. Degree Variances. As any 'output function' (output signal) can be expanded into an orthogonal series of surface spherical harmonics

$$\begin{aligned}\tilde{G} = \widetilde{\Lambda F} &= \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \Lambda^\wedge(n,k) \tilde{F}^\wedge(n,k) Y_{n,k} \\ &= \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} \tilde{G}^\wedge(n,k) Y_{n,k}\end{aligned}$$

in the sense of $\|\cdot\|_{\mathcal{L}^2(\Omega)}$, we get a spectral representation of the form

$$\tilde{G}^\wedge(n,k) = (\widetilde{\Lambda F})^\wedge(n,k) = \Lambda^\wedge(n,k)\tilde{F}^\wedge(n,k), \quad (n,k) \in \mathcal{N}.$$

The *signal degree and order variance* of $\tilde{G} = \widetilde{\Lambda F}$ is defined by

$$\begin{aligned} Var_{n,k}\left(\widetilde{\Lambda F}\right) &= \left(\left(\widetilde{\Lambda F}\right)^\wedge(n,k)\right)^2 \\ &= \int_\Omega \int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\left(\widetilde{\Lambda F}\right)(\eta) Y_{n,k}(\xi) Y_{n,k}(\eta)\, d\omega(\xi)\, d\omega(\eta) \ . \end{aligned}$$

Correspondingly, the *signal degree variances of* $\tilde{G} = \widetilde{\Lambda F}$ are given by

$$\begin{aligned} Var_n\left(\widetilde{\Lambda F}\right) &= \sum_{k=1}^{2n+1} Var_{n,k}\left(\widetilde{\Lambda F}\right) \\ &= \sum_{k=1}^{2n+1} \left(\left(\widetilde{\Lambda F}\right)^\wedge(n,k)\right)^2 \\ &= \frac{2n+1}{4\pi} \int_\Omega \int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\left(\widetilde{\Lambda F}\right)(\eta) P_n(\xi\cdot\eta)\, d\omega(\xi)\, d\omega(\eta), \end{aligned}$$

$n = 0, 1, \ldots$. According to Parseval's identity we clearly have

$$\|\widetilde{\Lambda F}\|^2_{\mathcal{L}^2(\Omega)} = \sum_{n=0}^{\infty} Var_n\left(\widetilde{\Lambda F}\right) = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} Var_{n,k}\left(\widetilde{\Lambda F}\right).$$

Physical devices do not transmit spherical harmonics of arbitrarily high frequency without severe attenuation. The 'transfer' $\Lambda^\wedge(n,k)$ usually tends to zero with increasing n. It follows that the amplitude spectra of the responses (observations) to functions (signals) of finite $\mathcal{L}^2(\Omega)$–energy are negligibly small beyond some finite frequency. Thus, both because of the frequency limiting nature of the used devices and because of the nature of the 'transmitted signals', the geoscientist is soon led to consider bandlimited functions. These are the functions $\tilde{G} \in \mathcal{L}^2(\Omega)$, whose 'amplitude spectra' vanish for all $n > N$ ($N \in \mathbb{N}_0$, fixed). In other words, $Var_n(\tilde{G}) = 0$, for all $n > N$.

7.3. Degree Error Covariances. The error spectral theory is based on the *degree and order error covariance*

$$Cov_{n,k}(K) = \int_\Omega \int_\Omega K(\xi,\eta) Y_{n,k}(\xi) Y_{n,k}(\eta)\, d\omega(\xi)\, d\omega(\eta), \quad (n,k) \in \mathcal{N},$$

and the *spectral degree error covariance*

$$Cov_n(K) = \sum_{k=1}^{2n+1} \int_\Omega \int_\Omega K(\xi,\eta) Y_{n,k}(\xi) Y_{n,k}(\eta)\, d\omega(\xi)\, d\omega(\eta), \quad n \in \mathbb{N}_0 \ .$$

Obviously,

$$Cov_{n,k}(K) = K^\wedge(n,k) \ .$$

In other words, the spectral degree and order error covariance is simply the orthogonal coefficient of the kernel K.

7.4. Examples of Spectral Error Covariances. To make the preceding considerations more concrete we would like to list two particularly important examples:

(1) *Bandlimited white noise.* Suppose that for some $n_K \in \mathbb{N}_0$

$$K^\wedge(n,k) = K^\wedge(n) = \begin{cases} \frac{\sigma^2}{(n_K+1)^2} & , \quad n \le n_K, k = 1,\dots,2n+1 \\ 0 & , \quad n > n_K, k = 1,\dots,2n+1, \end{cases}$$

where $\tilde{\varepsilon}$ is assumed to be $N(0,\sigma^2)$-distributed.The kernel reads as follows:

$$K(\xi,\eta) = \frac{\sigma^2}{(n_K+1)^2} \sum_{n=0}^{n_K} \frac{2n+1}{4\pi} P_n(\xi \cdot \eta) \ .$$

Note that this sum, apart from a multiplicative constant, may be understood as a truncated Dirac δ-functional. It is known (see e.g. N.N. Lebedew (1973)) that for $(\xi,\eta) \in \Omega \times \Omega$

$$((\xi \cdot \eta) - 1)\, K(\xi,\eta) = \frac{\sigma^2}{4\pi(n_K+1)} \left(P_{n_K+1}(\xi \cdot \eta) - P_{n_K}(\xi \cdot \eta)\right).$$

(2) *Non–bandlimited colored noise.* Assume that $K : \Omega \times \Omega \to \mathbb{R}$ is given in such a way that $K^\wedge(n,k) = K^\wedge(n) > 0$ for an infinite number of pairs $(n,k) \in \mathcal{N}$, the integral $\int_{-1}^{\delta} K(t)dt$ is sufficiently small (for some $\delta \in (1-\varepsilon, 1)$ for some $\varepsilon > 0$), and $K(\xi,\xi)$ coincides with σ^2 for all $\xi \in \Omega$.

Geophysically relevant examples are the following kernels:

(i) $K(\xi,\eta) = \frac{\sigma^2}{exp(-c)} \exp(-c(\xi \cdot \eta)), \qquad (\xi,\eta) \in \Omega \times \Omega,$

where c is to be understood as the inverse spherical correlation length *(first degree Gauß–Markov model).*

(ii) $K(\xi \cdot \eta) = \frac{\sigma^2}{(L^{(s)}_{\rho_{J^*}})^{(2)}(1)} (L^{(s)}_{\rho_{J^*}})^{(2)}(\xi \cdot \eta), \qquad (\xi,\eta) \in \Omega \times \Omega,$

for some sufficiently large $J^* \in \mathbb{N}$ *(model of small correlation length).* The family of *locally supported singular integrals* $\{L^{(s)}_{\rho_j}\}_{j\in\mathbb{Z}} \subset \mathcal{L}^2[-1,+1]$ is given by

$$\left(L^{(s)}_{\rho_j}\right)^\wedge(n,k) = \left(L^{(s)}_{\rho_j}\right)^\wedge(n) = 2\pi \int_{-1}^{+1} L^{(s)}_{\rho_j}(t) P_n(t)\, dt, \quad (n,k) \in \mathcal{N},$$

where

$$L^{(s)}_{\rho_j}(t) = \begin{cases} 0 & \text{for} \quad -1 \le t \le 1-\rho_j \\ \frac{1}{2\pi} \frac{s+1}{\rho_j^{s+1}} (t - 1 + \rho_j)^s & \text{for} \quad 1-\rho_j < t \le 1. \end{cases}$$

For the case $k = 0$ this example is known as the *Haar singular integral* (more details about Haar wavelets can be found in W. Freeden, K. Hesse (2000)).

7.5. Spectral Estimation. Now we are in position to compare the signal spectrum with that of the noise.

Signal and noise spectrum 'intersect' at the so–called *degree and order resolution set* $\mathcal{N}_{\mathrm{res}}$ (with $\mathcal{N}_{\mathrm{res}} \subset \mathcal{N}$). We distinguish the following cases:

(i) *signal dominates noise*

$$Var_{n,k}(\widetilde{\Lambda F}) \geq Cov_{n,k}(K), \quad (n,k) \in \mathcal{N}_{\mathrm{res}},$$

(ii) *noise dominates signal*

$$Var_{n,k}(\widetilde{\Lambda F}) < Cov_{n,k}(K), \quad (n,k) \notin \mathcal{N}_{\mathrm{res}} \ .$$

Filtering is achieved by convolving a square-summable product kernel H with the 'symbol' $\{H^\wedge(n,k)\}_{(n,k)\in\mathcal{N}}$ against $\widetilde{\Lambda F}$:

$$\widehat{\Lambda F} = \int_\Omega H(\cdot,\eta)\widetilde{\Lambda F}(\eta) \ \ d\omega(\eta)$$

('$\wedge$' denotes 'estimated'). In spectral language this reads

$$\widehat{\Lambda F}(n,k) = H^\wedge(n,k)\widetilde{\Lambda F}(n,k), \quad (n,k) \in \mathcal{N}. \tag{7.2}$$

Two important types of filtering are as follows:

(i) *Spectral thresholding*

$$\widehat{\Lambda F} = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} I_{\mathcal{N}_{\mathrm{res}}}(n,k)H^\wedge(n,k)\left(\widetilde{\Lambda F}\right)^\wedge(n,k)Y_{n,k}, \tag{7.3}$$

where I_A denotes the *indicator function of the set* A. This approach represents a 'keep or kill' filtering, where the signal dominated coefficients are filtered by a square-summable product kernel, and the noise dominated coefficients are set to zero. This thresholding can be thought of as a non–linear operator on the set of coefficients, resulting in a set of estimated coefficients.

As a special filter we mention the *ideal low–pass (Shannon) filter* H of the form

$$H^\wedge(n,k) = H^\wedge(n) = \left\{ \begin{array}{lcl} 1 & , & (n,k) \in \mathcal{N}_{res} \\ 0 & , & (n,k) \notin \mathcal{N}_{res}, \end{array} \right. \tag{7.4}$$

In that case all 'frequencies' $(n,k) \in \mathcal{N}_{res}$ are allowed to pass, whereas all other frequencies are completely eliminated.

(ii) *Wiener–Kolmogorov filtering.* Now we choose

$$\widehat{\Lambda F} = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} H^\wedge(n)\left(\widetilde{\Lambda F}\right)^\wedge(n,k)Y_{n,k} \tag{7.5}$$

with

$$H^\wedge(n) = \frac{Var_n(\widetilde{\Lambda F})}{Var_n(\widetilde{\Lambda F}) + Cov_n(K)}, \quad n \in \mathbb{N}_0 \ . \tag{7.6}$$

This filter produces an optimal weighting between signal and noise (provided that complete independence of signal and noise is assumed). Note the similarity to the rotation-invariant Tikhonov singular integral in (6.1).

8. Multiscale Signal–to–Noise Response. Consider a sequence $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ of square-summable product kernels constituting an approximate identity in $\mathcal{L}^2(\Omega)$. Then we have verified that an 'output signal' $\tilde{G} \in \mathcal{L}^2(\Omega)$ of an operator Λ can be represented in multiscale approximation as follows

$$\tilde{G} = \sum_{j=-\infty}^{+\infty} \int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)\tilde{G}(\eta) \; d\omega(\eta), \tag{8.1}$$

where the equality is understood in $\|\cdot\|_{\mathcal{L}^2(\Omega)}$–sense. The identity (8.1) is equivalent to the identity

$$\lim_{N\to\infty} \left\| \widetilde{\Lambda F} - \left(\left(\widetilde{\Lambda F}\right)_{J_0} + \sum_{j=J_0}^{N} \int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\eta)(\widetilde{\Lambda F})(\eta) \; d\omega(\eta) \right) \right\|_{\mathcal{L}^2(\Omega)} = 0$$

for every $J_0 \in \mathbb{Z}$.

8.1. Scale and Position Variances. Denote by $\mathcal{L}^2(\mathbb{Z}\times\Omega)$ the space of functions $H : \mathbb{Z}\times\Omega \to \mathbb{R}$ satisfying

$$\sum_{j=-\infty}^{\infty} \int_\Omega (H(j;\eta))^2 \, d\omega(\eta) < \infty \; .$$

$\mathcal{L}^2(\mathbb{Z}\times\Omega)$ is a Hilbert space equipped with the inner product

$$(H_1,H_2)_{\mathcal{L}^2(\mathbb{Z}\times\Omega)} = \sum_{j=-\infty}^{+\infty} \int_\Omega H_1(j;\eta)H_2(j;\eta) \; d\omega(\eta)$$

corresponding to the norm

$$\|H\|_{\mathcal{L}^2(\mathbb{Z}\times\Omega)} = \left(\sum_{j=-\infty}^{+\infty} \int_\Omega (H(j;\eta))^2 \, d\omega(\eta) \right)^{1/2} .$$

Consider a family of square-summable product kernels $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ constituting an approximate identity in $\mathcal{L}^2(\Omega)$. From the multiscale formulation of an 'output function' $\tilde{G} = \widetilde{\Lambda F} \in \mathcal{L}^2(\Omega)$ we immediately obtain (cd. W. Freeden et al. (2000))

$$\begin{aligned}
&\left(\widetilde{\Lambda F},\widetilde{\Lambda F}\right)_{\mathcal{L}^2(\Omega)} \\
&= \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} \left(\left(\widetilde{\Lambda F}\right)^\wedge(n,k)\right)^2 \sum_{j=-\infty}^{+\infty} \left(\left(\Psi_{\rho_j}\right)^\wedge(n,k)\right)^2 \\
&= \sum_{j=-\infty}^{+\infty} \int_\Omega \left(\int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\Psi_{\rho_j}(\eta,\xi) d\omega(\xi) \right)^2 d\omega(\eta) \\
&= \sum_{j=-\infty}^{\infty} \int_\Omega\int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\left(\widetilde{\Lambda F}\right)(\zeta)\Psi^{(2)}_{\rho_j}(\xi,\zeta) \; d\omega(\xi) \; d\omega(\zeta) \\
&= \sum_{j=-\infty}^{+\infty} \int_\Omega \left(\int_\Omega\int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\left(\widetilde{\Lambda F}\right)(\zeta)\Psi_{\rho_j}(\xi,\eta)\Psi_{\rho_j}(\zeta,\eta) \; d\omega(\xi) \; d\omega(\zeta) \right) d\omega(\eta) \; .
\end{aligned}$$

The *signal scale and space variance of* $\widetilde{\Lambda F}$ *at position* $\eta \in \Omega$ *and scale* $j \in \mathbb{Z}$ is defined by

$$Var_{j;\eta}\left(\widetilde{\Lambda F}\right) = \int_\Omega \int_\Omega \left(\widetilde{\Lambda F}\right)(\xi)\left(\widetilde{\Lambda F}\right)(\zeta)\Psi_{\rho_j}(\xi,\eta)\Psi_{\rho_j}(\zeta,\eta)d\omega(\xi)d\omega(\zeta) \ .$$

The *signal scale variance of* $\widetilde{\Lambda F}$ is defined by

$$Var_j(\widetilde{\Lambda F}) = \int_\Omega Var_{j;\eta}(\widetilde{\Lambda F})d\omega(\eta) \ .$$

Obviously, we have

$$\begin{aligned}\left\|\widetilde{\Lambda F}\right\|^2_{\mathcal{L}^2(\Omega)} &= \sum_{j=-\infty}^{+\infty} Var_j(\widetilde{\Lambda F}) \\ &= \sum_{j=-\infty}^{+\infty} \int_\Omega Var_{j;\eta}\left(\widetilde{\Lambda F}\right)d\omega(\eta) \\ &= \left\|(Var_{\cdot;\cdot}\left(\widetilde{\Lambda F}\right))^{1/2}\right\|^2_{\mathcal{L}^2(\mathbb{Z}\times\Omega)} \ .\end{aligned}$$

Expressed in the spectral language of spherical harmonics we get

$$Var_j\left(\widetilde{\Lambda F}\right) = \int_\Omega Var_{j;\eta}\left(\widetilde{\Lambda F}\right)d\omega(\eta) = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1}\left(\Psi^\wedge_{\rho_j}(n,k)\right)^2\left(\left(\widetilde{\Lambda F}\right)^\wedge(n,k)\right)^2 \ .$$

With the convention $\mathcal{Z} = \mathbb{Z}\times\Omega$ we are formally able to write

$$\left\|\widetilde{\Lambda F}\right\|^2_{\mathcal{L}^2(\Omega)} = \left\|\left(Var_{\cdot,\cdot}(\widetilde{\Lambda F})\right)^{1/2}\right\|_{\mathcal{L}^2(\mathcal{Z})} \ . \tag{8.2}$$

We mention that the Beppo-Levi Theorem justifies to interchange integration and summation. Note that all integrations are understood in the Lebesgue-sense.

8.2. Noise Model. Let $K : (\xi,\eta) \mapsto K(\xi,\eta)$, $(\xi,\eta) \in \Omega\times\Omega$, satisfy the conditions (C1) and (C2) stated in Section 7.1. The error theory is based on the *scale and space error covariance* at $\eta \in \Omega$

$$Cov_{j;\eta}(K) = \int_\Omega\int_\Omega K(\xi,\zeta)\Psi_{\rho_j}(\xi,\eta)\Psi_{\rho_j}(\zeta,\eta)d\omega(\xi)d\omega(\zeta), \qquad \eta\in\Omega \ .$$

The *scale error covariance* is defined by

$$Cov_j(K) = \int_\Omega Cov_{j;\eta}(K)d\omega(\eta) \ .$$

We obviously have in spectral language

$$Cov_{j;\eta}(K) = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} K^\wedge(n,k)\left(\Psi^\wedge_{\rho_j}(n,k)\right)^2 \ .$$

It is clear from our stochastic model, i.e. from the special representation of the covariance as a product kernel, that the scale error covariance cannot be dependent on the position $\eta \in \Omega$. This is also indicated by the spectral formula

$$Cov_{j;\eta}(K) = \frac{1}{4\pi} \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} Cov_n(K) \left(\Psi_{\rho_j}^{\wedge}(n,k) \right)^2 \quad .$$

Our error model is particularly useful for the proper handling of the satellite data in Earth's gravitational or magnetic potential determination (see W.Freeden (1999) and the references therein).

8.3. Scale and Space Estimation. Signal and noise scale 'intersect' at the so–called *scale and space resolution set* $\mathcal{Z}_{\mathrm{res}}$ *with* $\mathcal{Z}_{\mathrm{res}} \subset \mathcal{Z}$. We distinguish the following cases:

(i) *signal dominates noise*

$$Var_{j;\eta}\left(\widetilde{\Lambda F}\right) \geq Cov_{j;\eta}(K), \quad (j;\eta) \in \mathcal{Z}_{\mathrm{res}} \quad .$$

(ii) *noise dominates signal*

$$Var_{j;\eta}\left(\widetilde{\Lambda F}\right) < Cov_{j;\eta}(K), \quad (j;\eta) \notin \mathcal{Z}_{\mathrm{res}} \quad .$$

Via the multiscale reconstruction formula the (filtered) J–level approximation of the error–affected function $\widetilde{\Lambda F}$ reads as follows

$$(\widetilde{\Lambda F})_J = \sum_{j=-\infty}^{J} \int_{\Omega} \Psi_{\rho_j}^{(2)}(\cdot,\eta) \left(\widetilde{\Lambda F}\right)(\eta) d\omega(\eta) \quad .$$

For J sufficiently large, $\widetilde{\Lambda F}$ is well–represented by $(\widetilde{\Lambda F})_J$. In other words, all the higher–level coefficients are regarded as being negligible, i.e. $(\widetilde{\Lambda F})_J \backsimeq \widetilde{\Lambda F}$.

9. Selective Multiscale Reconstruction. Similar to what is known in taking Fourier approximation, we are able to take multiscale approximation by replacing the (unknown) error–free function ΛF of the representation

$$\begin{aligned}(\Lambda F)_J &= \int_\Omega \Phi^{(2)}_{\rho_{J_0}}(\cdot,\zeta)(\Lambda F)(\zeta)d\omega(\zeta)\\ &\quad + \sum_{j=J_0}^{J-1}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\zeta)(\Lambda F)(\zeta)d\omega(\zeta)\end{aligned}$$

by (an estimate from) the error–affected function $\widetilde{\Lambda F}$ such as

$$\begin{aligned}(\widetilde{\Lambda F})_J &= \int_\Omega \Phi^{(2)}_{\rho_{J_0}}(\cdot,\zeta)(\widetilde{\Lambda F})(\zeta)d\omega(\zeta)\\ &\quad + \sum_{j=J_0}^{J-1}\int_\Omega \Psi^{(2)}_{\rho_j}(\cdot,\zeta)(\widetilde{\Lambda F})(\zeta)d\omega(\zeta),\end{aligned}$$

$J > J_0$. Computing the following coefficients at position $\eta \in \Omega$

$$\begin{aligned}V_{J_0;\eta} &= \int_\Omega \Phi^{(2)}_{\rho_{J_0}}(\eta,\zeta)(\Lambda F)(\zeta)d\omega(\zeta)\\ W_{j;\eta} &= \int_\Omega \Psi^{(2)}_{\rho_j}(\eta,\zeta)(\Lambda F)(\zeta)d\omega(\zeta), \quad j = J_0,\ldots,J-1,\end{aligned}$$

and

$$\begin{aligned}\tilde{V}_{J_0;\eta} &= \int_\Omega \Phi^{(2)}_{\rho_{J_0}}(\eta,\zeta)(\widetilde{\Lambda F})(\zeta)d\omega(\zeta)\\ \tilde{W}_{j;\eta} &= \int_\Omega \Psi^{(2)}_{\rho_j}(\eta,\zeta)(\widetilde{\Lambda F})(\zeta)d\omega(\zeta), \quad j = J_0,\ldots,J-1\end{aligned}$$

will, of course, require adequate methods of numerical integration on the sphere.

9.1. Numerical Integration on the Sphere. Many integration techniques are known from the literature (for a survey on approximate integration on the sphere see, for example, W. Freeden et al. (1998) and the references therein). In what follows we base integration on the approximate formulae associated to known weights $w_i^{N_j} \in \mathbb{R}$ and knots $\eta_i^{N_j} \in \Omega$

$$\begin{aligned}\tilde{V}_{J_0;\eta} &\backsimeq \sum_{i=1}^{N_{J_0}} w_i^{N_{J_0}}\Phi^{(2)}_{\rho_{J_0}}(\eta,\eta_i^{N_{J_0}})(\widetilde{\Lambda F})(\eta_i^{N_{J_0}}),\\ \tilde{W}_{j;\eta} &\backsimeq \sum_{i=1}^{N_j} w_i^{N_j}\Psi^{(2)}_{\rho_j}(\eta,\eta_i^{N_j})(\widetilde{\Lambda F})(\eta_i^{N_j}), \quad j = J_0,\ldots,J-1\end{aligned}$$

('$\backsimeq$' always means that the error is assumed to be negligible). An example (cf. W. Freeden et al. (2000)) is equidistribution (i.e. $w_i^{N_j} = \frac{4\pi}{N_j}, i = 1,\ldots,N_j$).

9.2. A Pyramid Scheme. Next we deal with some aspects of scientific computing. We are interested in a pyramid scheme for the (approximate) recursive computation of the integrals $\tilde{V}_{J_0;\eta}$, $\tilde{W}_{j;\eta}$ for $j = J_0,\ldots,J-1$.

What we are going to realize is a *tree algorithm (pyramid scheme)* with the following ingredients: Starting from a sufficiently large J such that

$$\widetilde{\Lambda F}(\eta) \simeq \Phi^{(2)}_{\rho_J}(\cdot,\eta) * \widetilde{\Lambda F} \simeq \sum_{i=1}^{N_J} \Phi^{(2)}_{\rho_J}(\eta,\eta_i^{N_J})\tilde{a}_i^{N_J}, \quad \eta \in \Omega, \tag{9.1}$$

we want to show that the coefficient vectors $\tilde{a}^{N_j} = \left(\tilde{a}_1^{N_j},\dots,\tilde{a}_{N_j}^{N_j}\right)^T \in \mathbb{R}^{N_j}$ $j = J_0,\dots,J-1$, (being, of course, dependent on the function $\widetilde{\Lambda F}$ under consideration) can be calculated such that the following statements hold true:

(i) The vectors $\tilde{a}^{N_j}$, $j = J_0,\dots,J-1$, are obtainable by recursion from the values $\tilde{a}_i^{N_J}$.
(ii) For $j = J_0,\dots,J$

$$\Phi^{(2)}_{\rho_j}(\cdot,\eta) * \widetilde{\Lambda F} \simeq \sum_{i=1}^{N_j} \Phi^{(2)}_{\rho_j}(\eta,\eta_i^{N_j})\tilde{a}_i^{N_j}.$$

For $j = J_0,\dots,J-1$

$$\Psi^{(2)}_{\rho_j}(\cdot,\eta) * \widetilde{\Lambda F} \simeq \sum_{i=1}^{N_j} \Psi^{(2)}_{\rho_j}(\eta,\eta_i^{N_j})\tilde{a}_i^{N_j} \quad .$$

Our considerations are divided into two parts, viz. the initial step concerning the scale level J and the pyramid step establishing the recursion relation:

The Initial Step. For a suitably large integer J, $\Phi^{(2)}_{\rho_J}(\cdot,\eta) * \widetilde{\Lambda F}$ is sufficiently close to $(\widetilde{\Lambda F})(\eta)$ for all $\eta \in \Omega$. Formally spoken, the kernel $\Phi^{(2)}_{\rho_J}$ replaces the Dirac–functional δ as follows:

$$\Phi^{(2)}_{\rho_J}(\cdot,\eta) * \widetilde{\Lambda F} \simeq \widetilde{\Lambda F}(\eta) = \left(\delta * \widetilde{\Lambda F}\right)(\eta) = \delta_\eta * \widetilde{\Lambda F} \quad ,$$

where

$$\delta(\xi,\eta) = \delta_\xi(\eta) = \sum_{n=0}^{\infty}\sum_{k=1}^{2n+1} Y_{n,k}(\xi)Y_{n,k}(\eta)$$

and the series has to be understood in distributional sense. The formulae

$$\Phi^{(2)}_{\rho_J}(\cdot,\eta_i^{N_J}) * \widetilde{\Lambda F} \simeq \widetilde{\Lambda F}(\eta_i^{N_J}), \;\; i = 1,\dots,N_J$$

are the reason why the coefficients for the initial step, i.e. $\tilde{a}^{N_J} = (\tilde{a}_1^{N_J},\dots,\tilde{a}_{N_J}^{N_J})^T \in \mathbb{R}^{N_J}$, are assumed to be simply given in the form

$$\tilde{a}_i^{N_J} = w_i^{N_J}\left(\widetilde{\Lambda F}\right)\left(\eta_i^{N_J}\right) \quad , i = 1,\dots,N_J \tag{9.2}$$

The Pyramid Step. The essential idea for the development of a pyramid scheme is the existence of kernel functions $\Xi_j : \Omega\times\Omega \to \mathbb{R}$ such that

$$\Phi^{(2)}_{\rho_j} \simeq \Xi_j * \Phi^{(2)}_{\rho_j} \tag{9.3}$$

and

$$\Xi_j \simeq \Xi_{j+1} * \Xi_j \tag{9.4}$$

for $j = J_0, \ldots, J$.
Note that for bandlimited scaling functions the kernels Ξ_j, $j = J_0, \ldots, J$, may be chosen to be the reproducing kernels of the finite-dimensional scale spaces $\mathcal{V}_{\rho_j}$ (cf. (3.1)), whereas in the non-bandlimited case Ξ_j, $j = J_0, \ldots, J$, may be chosen such that $\Xi_j = \delta \simeq \Phi^{(2)}_{\rho_J}$.
Observing our approximate integration formulae we obtain in connection with relation (9.3)

$$\Phi^{(2)}_{\rho_j} * \widetilde{\Lambda F} \simeq \Phi^{(2)}_{\rho_j} * \Xi_j * \widetilde{\Lambda F} \simeq \sum_{i=1}^{N_j} \Phi^{(2)}_{\rho_j}(\cdot, \eta_i^{N_j}) \tilde{a}_i^{N_j}, \tag{9.5}$$

where

$$\tilde{a}_i^{N_j} = w_i^{N_j} \big(\Xi_j * \widetilde{\Lambda F}\big)(\eta_i^{N_j}), \ \ j = J_0, \ldots, J-1.$$

Now it follows by use of our approximate integration formulae and the assumption (9.4) that

$$\begin{aligned}
\tilde{a}_i^{N_j} &= w_i^{N_j} \big(\Xi_j * \widetilde{\Lambda F}\big)(\eta_i^{N_j}) \\
&\simeq w_i^{N_j} \big(\Xi_j * \Xi_{j+1} * \widetilde{\Lambda F}\big)(\eta_i^{N_j}) \\
&\simeq w_i^{N_j} \sum_{l=1}^{N_{j+1}} w_l^{N_{j+1}} \Xi_j(\eta_i^{N_j}, \eta_l^{N_{j+1}}) \big(\Xi_{j+1} * \widetilde{\Lambda F}\big)(\eta_l^{N_{j+1}}) \\
&= w_i^{N_j} \sum_{l=1}^{N_{j+1}} \Xi_j(\eta_i^{N_j}, \eta_l^{N_{j+1}}) \tilde{a}_i^{N_{j+1}}.
\end{aligned}$$

In other words, the coefficients $\tilde{a}_i^{N_{J-1}}$ can be calculated recursively starting from the data $\tilde{a}_i^{N_J}$ for the initial level J, $\tilde{a}_i^{N_{J-2}}$ can be deduced recursively from $\tilde{a}_i^{N_{J-1}}$, etc. Moreover, the coefficients are independent of the special choice of the kernel (Observe that (9.5) is equivalent to $\big(\widetilde{\Lambda F}\big)^\wedge(n,k) \simeq \sum_{i=1}^{N_j} \tilde{a}_i^{N_j} Y(\eta_i^{N_j})$ for $n = 0, 1, \ldots,$ $k = 1, \ldots, 2n+1$). This finally leads us to the formulae

$$\Phi^{(2)}_{\rho_j}(\cdot, \eta) * \widetilde{\Lambda F} \backsimeq \sum_{i=1}^{N_j} \Phi^{(2)}_{\rho_j}(\eta, \eta_i^{N_j}) \tilde{a}_i^{N_j}, \quad j = J_0, \ldots, J,$$

and

$$\Psi^{(2)}_{\rho_j}(\cdot, \eta) * \widetilde{\Lambda F} \backsimeq \sum_{i=1}^{N_j} \Psi^{(2)}_{\rho_j}(\eta, \eta_i^{N_j}) \tilde{a}_i^{N_j}, \quad j = J_0, \ldots, J-1,$$

with coefficients $\tilde{a}_i^{N_j}$ given by (9.2) and (9.5). In the bandlimited case (with Ξ_j chosen as indicated above) the sign " $\simeq$ " can be replaced by " $=$ " provided that spherical harmonic exact integration formulae of suitable degree are used (cf. W. Freeden (1999)).

This recursion procedure leads us to the following *decomposition scheme*:

$$\begin{array}{ccccccccc}
\widetilde{\Lambda F} & \rightarrow & \tilde{a}^{N_J} & \rightarrow & \tilde{a}^{N_{J-1}} & \rightarrow & \ldots & \rightarrow & \tilde{a}^{N_{J_0}} \\
 & & \downarrow & & \downarrow & & & & \downarrow \\
 & & \tilde{W}_{J;\eta} & & \tilde{W}_{J-1;\eta} & & & & \tilde{W}_{J_0;\eta} \ .
\end{array}$$

The coefficient vectors $\tilde{a}^{N_{J_0}}$, $\tilde{a}^{N_{J_0+1}}, \ldots$ allow the following *reconstruction scheme* of $\widetilde{\Lambda F}$:

$$\begin{array}{cccccccccc}
\tilde{a}^{N_{J_0}} & & & \tilde{a}^{N_{J_0+1}} & & & \tilde{a}^{N_{J_0+2}} & & & \\
\downarrow & & & \downarrow & & & \downarrow & & & \\
\Psi^{(2)}_{\rho_{J_0}} * \widetilde{\Lambda F} & & & \Psi^{(2)}_{\rho_{J_0+1}} * \widetilde{\Lambda F} & & & \Psi^{(2)}_{\rho_{J_0+2}} * \widetilde{\Lambda F} & & & \\
 & \searrow & & & \searrow & & & \searrow & & \\
\Phi^{(2)}_{\rho_{J_0}} * \widetilde{\Lambda F} & \rightarrow & + \rightarrow & \Phi^{(2)}_{\rho_{J_0+1}} * \widetilde{\Lambda F} & \rightarrow & + \rightarrow & \Phi^{(2)}_{\rho_{J_0+2}} * \widetilde{\Lambda F} & \rightarrow & + \rightarrow \ldots & .
\end{array}$$

Once again it is worth mentioning that the coefficient vectors $\tilde{a}^{N_j}$ do *not* depend on the special choice of the scaling function $\{\Phi^{(2)}_{\rho_j}\}_{j\in\mathbb{Z}}$ in $\mathcal{L}^2(\Omega)$. Moreover, the coefficients can be used to calculate the wavelet transforms $\Psi_{\rho_j}(\cdot,\eta) * \widetilde{\Lambda F}$ for $j = J_0, \ldots, J-1$ and all $\eta \in \Omega$.

10. Scale Thresholding. Since the large 'true' coefficients are the ones that should be included in a selective reconstruction, in estimating an unknown function it is natural to include only coefficients larger than some specified threshold value. In our context a 'larger' coefficient is taken to mean one that satisfies for $j = J_0, \ldots, J$ and $i = 1, \ldots, N_j$

$$\begin{aligned}
\left(\tilde{a}_i^{N_j}\right)^2 &= \left(w_i^{N_j}\left(\Xi_j * \widetilde{\Lambda F}\right)(\eta_i^{N_j})\right)^2 \\
&= (w_i^{N_j})^2 \int_\Omega\int_\Omega \widetilde{\Lambda F}(\xi)\widetilde{\Lambda F}(\zeta)\ \Xi_j(\xi,\eta_i^{N_j})\ \Xi_j(\zeta,\eta_i^{N_j})\ d\omega(\xi)\ d\omega(\zeta) \\
&\geq (w_i^{N_j})^2 \int_\Omega\int_\Omega K(\xi,\zeta)\ \Xi_j(\xi,\eta_i^{N_j})\ \Xi_j(\zeta,\eta_i^{N_j})\ d\omega(\xi)\ d\omega(\zeta) \\
&= (k_i^j)^2.
\end{aligned}$$

Remark 10.1. In particular for "bandlimited white noise" of the form

$$K(\eta,\xi) = K(\eta\cdot\xi) = \frac{\sigma^2}{4\pi}P_0(\eta\cdot\xi) = \frac{\sigma^2}{4\pi},$$

$(\eta,\xi) \in \Omega\times\Omega$ *and* $w_i^{N_j} = \frac{4\pi}{N_j}$ *(i.e. equidistributions), we find*

$$(k_i^j)^2 = \frac{2\sqrt{\pi}}{N_j}\sigma\left(\Xi_j^\wedge(0,1)\right)^2, \quad j = J_0,\ldots,J\ ,i = 1,\ldots,N_j.$$

For the given threshold values k_i^j such an estimator can be written in explicit form:

$$\begin{aligned}
\left(\widehat{\Lambda F}\right)_J &= \sum_{i=1}^{N_{J_0}} I_{\{(\tilde{a}_i^{N_{J_0}})^2 \geq (k_i^{J_0})^2\}} \Phi^{(2)}_{\rho_{J_0}}(\cdot,\eta_i^{N_{J_0}})\tilde{a}_i^{N_{J_0}} \\
&\quad + \sum_{j=J_0}^{J-1}\sum_{i=1}^{N_j} I_{\{(\tilde{a}_i^{N_j})^2 \geq (k_i^j)^2\}} \Psi^{(2)}_{\rho_j}\left(\cdot,\eta_i^{N_j}\right)\tilde{a}_i^{N_j} \quad .
\end{aligned}$$

In other words, the large coefficients (relative to the threshold $k_i^j, i = 1, \ldots, N_j, j = J_0, \ldots, J-1$) are kept intact and the small coefficients are set to zero. Motivated by our former results the thresholding will be performed on $\tilde{V}_{J_0;\eta}$ and $\tilde{W}_{j;\eta}$, $j = J_0, \ldots, J-1$. The *thresholding estimators* of the true coefficients $V_{J_0;\eta}$, $W_{j;\eta}$ can thus be written in the form

$$\hat{V}_{J_0;\eta} = \sum_{i=1}^{N_{J_0}} \delta^{\text{hard}}_{(k_i^{J_0})^2} \left((\tilde{a}_i^{N_{J_0}})^2 \right) \Phi^{(2)}_{\rho_{J_0}} \left(\eta, \eta_i^{N_{J_0}} \right) \tilde{a}_i^{N_{J_0}},$$

$$\hat{W}_{j;\eta} = \sum_{i=1}^{N_j} \delta^{\text{hard}}_{(k_i^{j})^2} \left((\tilde{a}_i^{N_j})^2 \right) \Psi^{(2)}_{\rho_j} \left(\eta, \eta_i^{N_j} \right) \tilde{a}_i^{N_j}, \tag{10.1}$$

where the function $\delta^{\text{hard}}_\lambda$ is the *hard thresholding function*

$$\delta^{\text{hard}}_\lambda(x) = \begin{cases} 1 & \text{if } |x| \geq \lambda \\ 0 & \text{otherwise} \end{cases}.$$

The 'keep or kill' hard thresholding operation is not the only reasonable way of estimating the coefficients. Recognizing that each coefficient $\tilde{W}_{j;\eta}$ consists of both a signal portion and a noise portion, it might be desirable to attempt to isolate the signal contribution by removing the noisy part. This idea leads to the ***soft thresholding function*** (confer the considerations by D.L. Donoho, I.M. Johnstone (1994, 1995))

$$\delta^{\text{soft}}_\lambda(x) = \begin{cases} max\{0, 1 - \frac{\lambda}{|x|}\} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

which can also be used in the identities (10.1) stated above. When soft thresholding is applied to a set of empirical coefficients, only coefficients greater than the threshold (in absolute value) are included, but their values are 'shrunk' toward zero by an amount equal to the threshold λ.

Summarizing all our results we finally obtain the following *thresholding multiscale estimator*

$$\left(\widehat{\Lambda F}\right)_J = \sum_{i=1}^{N_{J_0}} \delta_{(k_i^{J_0})^2} \left((\tilde{a}_i^{N_{J_0}})^2 \right) \Phi^{(2)}_{\rho_{J_0}} (\cdot, \eta_i^{N_{J_0}}) \tilde{a}_i^{N_{J_0}}$$
$$+ \sum_{j=J_0}^{J-1} \sum_{i=1}^{N_j} \delta_{(k_i^j)^2} \left((\tilde{a}_i^{N_j})^2 \right) \Psi^{(2)}_{\rho_j} \left(\cdot, \eta_i^{N_j} \right) \tilde{a}_i^{N_j} \quad .$$

In doing so $(\widehat{\Lambda F})_J$ first is approximated by a thresholded $(\widetilde{\Lambda F})_{J_0}$, which represents the smooth components of the data. Then the coefficients at higher resolutions are thresholded, so that the noise is suppressed but the fine–scale details are included in the calculation.

11. Example. In order to illustrate the effectiveness of our multiscale denoising technique we present a simple example using synthetic geomagnetic data. For this purpose we introduce geomagnetic coordinates X, Y and Z. X denotes the so-called north-, Y the east- and Z the downward-component. Using spherical polar coordinates and identifying 0 degree longitude with Greenwich and 0 degree latitude with the equator we end up with the following correspondence:

$$\begin{aligned} X &\leftrightarrow \varepsilon^t, \\ Y &\leftrightarrow \varepsilon^\varphi, \\ Z &\leftrightarrow -\varepsilon^r, \end{aligned}$$

where $\varepsilon^t, \varepsilon^\phi$ and ε^r are the usual unit vectors in spherical polar coordinates (for explicit representations see e.g. W. Freeden et al. 1998). This means that X,Y and Z form a local triad with X always pointing towards the geographic northpole, Y pointing into the geographic east direction and Z always being directed towards the Earth's body.

From a bandlimited (up to degree and order 12) geomagnetic potential due to C.J Cain et al. 1984 we calculated the corresponding gradient field in geomagnetic coordinates, i.e. north (ε^t), east (ε^φ) and downward ($-\varepsilon^r$) components, which served as unnoised data. We then added some bandlimited white noise with variance σ and bandwidth N of approximately 0.9 and 60, respectively. This resulted in noise of the order of magnitude 10^0 [nT] in a field of the order of magnitude 10^4 [nT]. The noised signal has then been decomposed and reconstructed using Shannon wavelets up to scale 4. During the reconstruction process only those wavelet coefficients containing a predominant amount of the clear signal have been used in accordance to our considerations in section 8.3. Fig. 11.1 shows the $-\varepsilon^r$ component of the unnoised data, while Fig. 11.2 shows the absolute values of the added noise.

Figs. 11.3 and 11.4 show the denoised $-\varepsilon^r$ component and the corresponding absolute error with respect to the unnoised data. Using our multiscale denoising technique the root-mean-square error of the noised data (w.r.t. the clear data), $(\Delta\varepsilon^r_{\text{noised}})_{\text{rms}} = 1.13$ [nT], has been reduced to $(\Delta\varepsilon^r_{\text{denoised}})_{\text{rms}} = 0.35$ [nT], which is an improvement of about 60 per cent.

Comparing Figs. 11.2 and 11.4 it can be realized how the rough structure of the noise has been smoothed out by the denoising process and how the peaks have been reduced throughout the whole data set. This example obviously shows the functionality of our approach.

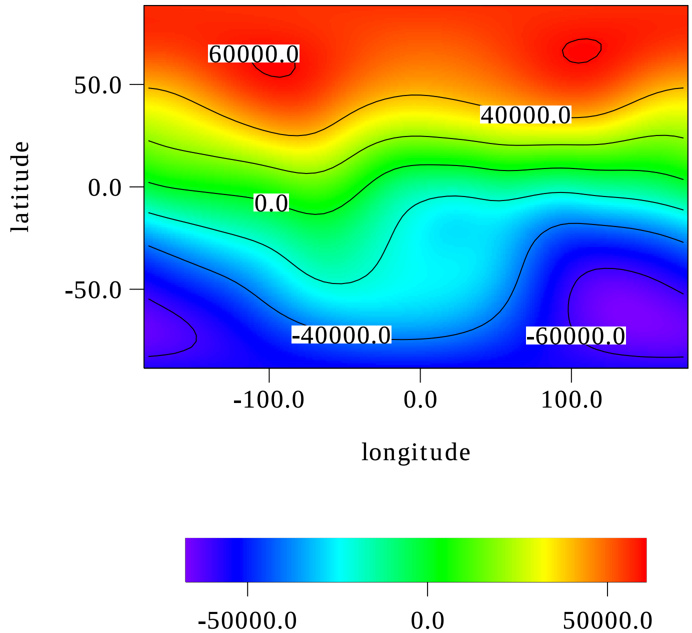

FIG. 11.1. *$-\varepsilon^r$ component of unnoised data [nT]*

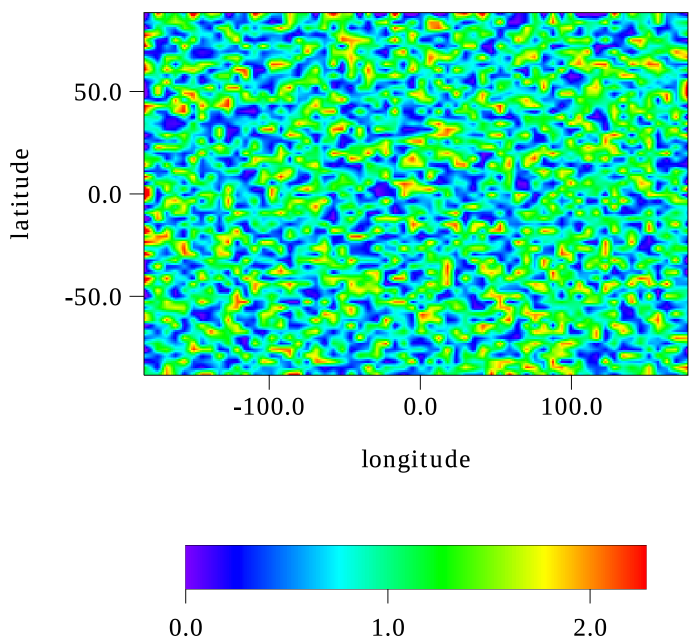

FIG. 11.2. *absolute value of added noise [nT]*

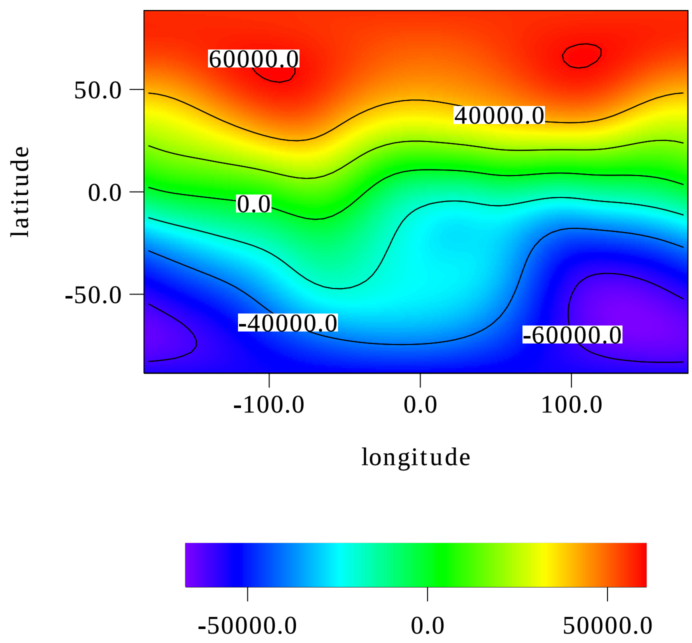

FIG. 11.3. *$-\varepsilon^r$ component of denoised data [nT]*

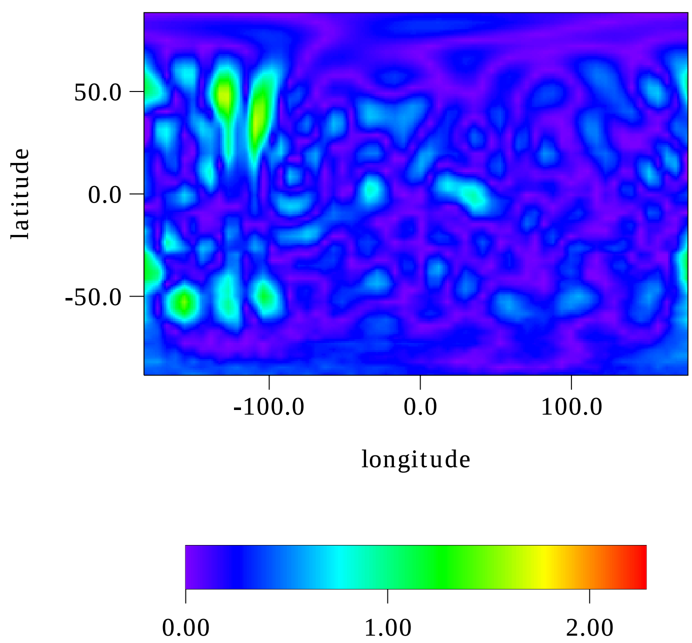

FIG. 11.4. *absolute error of denoised $-\varepsilon^r$ component [nT]*

Acknowledgements.

The support by German Research Foundation (DFG contract No. FR 761-10-1) is gratefully acknowledged.

References.

CAIN, C.J., SCHMITZ, D.R., MUTH, L. (1984) *Small-Scale Features in the Earth's Magnetic Field observed by MAGSAT.* Journal of Geophysical Research, Vol. 89, NO. B2, 1070-1076.

DONOHO, D.L., JOHNSTONE, I.M. (1994) *Ideal Spatial Adaptation by Wavelet Shrinkage.* Biometrika, **81**, 425-455.

DONOHO, D.L., JOHNSTONE, I.M. (1995) *Adapting to Unknown Smoothness Via Wavelet Shrinkage.* J. Amer. Statistical Association, **90**, 1200-1224.

FREEDEN, W. (1999): Multiscale Modelling of Spaceborne Geodata, B.G. Teubner, Stuttgart, Leipzig

FREEDEN, W., GERVENS, T., SCHREINER, M. (1998) Constructive Approximation on the Sphere (With Applications to Geomathematics). Oxford Science Publications, Clarendon.

FREEDEN, W., HESSE, K., (2000) *Locally Supported Difference Wavelets on the Sphere.* AGTM-Report 225, Laboratory of Technomathematics, Geomathematics Group, University of Kaiserslautern.

FREEDEN, W., MICHEL, V., STENGER, M., (2000) *Multiscale Signal–to–Noise Thresholding.* AGTM-Report 224, Laboratory of Technomathematics, Geomathematics Group, University of Kaiserslautern.

LEBEDEW N.N. (1973) Spezielle Funktionen und ihre Anwendung. BI–Wissenschaftsverlag, Mannheim, Wien, Zürich.

MICHEL, V. (1999) A Multiscale Method for the Gravimetry Problem — Theoretical and Numerical Aspects of Harmonic and Anharmonic Modelling. Doctoral Thesis, University of Kaiserslautern, Geomathematics Group, Shaker Verlag, Aachen.

MÜLLER, C. (1966) Spherical Harmonics. Lecture Notes in Mathematics, **17**, Springer, Berlin, Heidelberg.

OGDEN, R.T. (1997) Essential Wavelets for Statistical Applications and Data Analysis. Birkhäuser, Boston, Basel, Berlin.

RUMMEL, R. (1997) *Spherical Spectral Properties of the Earth's Gravitational Potential and its First and Second Derivatives.* In: Lecture Notes in Earth Sciences, **65**, 359–404.

WEAVER, J.B., YANSUN, X., HEALY, D.M.JR., CROMWELL L.D. (1991): *Filtering Noise From Images With Wavelet Transforms.* Magnetic Resonance in Medicine, **24** 288-295.

Spectral and Multiscale Signal–to–Noise Thresholding of Spherical Vector Fields

W. Freeden, T. Maier

University of Kaiserslautern
Geomathematics Group
67653 Kaiserslautern
P.O. Box 3049
Germany

phone:	++49 631 205-3867
fax:	++49 631 29081
email:	freeden@mathematik.uni-kl.de
	tmaier@mathematik.uni-kl.de
www:	http://www.mathematik.uni-kl.de/~wwwgeo

Abstract. The basic concepts of spectral and multiscale selective reconstruction of (geophysically relevant) vector fields on the sphere from error–affected data is outlined in detail. The reconstruction mechanism is formulated under the assumption that spectral as well as multiscale approximation is well–representable in terms of only a certain number of expansion coefficients at the various resolution levels. It is shown that spectral denoising by means of orthogonal expansions in terms of vector spherical harmonics reflects global a priori information of the noise (e.g. in form of a covariance tensor field), whereas multiscale signal–to–noise thresholding can be performed under locally dependent noise information within a multiresolution analysis in terms of spherical vector wavelets.

AMS Mathematics Subject Classification: 33C55, 42C40, 65T60, 86A25
Key words. Bilinear spherical vector wavelet theory, tensor–vector approximation, scalar–vector approximation, spectral and multiscale variance–covariance model, hard and soft thresholding.

Correspondence to W. Freeden

1. Introduction. Launching the multifunctional satellite CHAMP (Challenging Mini-Satellite Payload) on July 15, 2000 commenced the so-called *Decade of Geopotentials,* an era of enhanced and improved research in gravitational and geomagnetic field modelling. Only with satellites it is possible to cover the entire earth densely with measurements of uniform quality. In the research field of earth's gravitation the concept of high–low satellite–to–satellite tracking (hi–lo SST) realized by CHAMP seems to be capable of attaining strongly improved spatial refinement and accuracy. As far as the magnetic field is concerned, CHAMP is expected to provide global vector data of highest precision hopefully leading to unpredecented accuracy models of the main and crustal magnetic field and the space– and time–variability of these components.

While standard Fourier techniques for geopotential field determination – usually in terms of gradient fields of scalar spherical harmonics – are useful tools for modelling global trends, they are utterly incapable of dealing with geopotential data varying on small spatial scales. This, as well as the inherent vectorial character of gravitational and geomagnetic field approximation strongly point at using new, innovative mathematically and physically realistic approximation methods in order to ensure enhanced field models from the upcoming high quality satellite data.

When dealing with actual observational data it should be kept in mind that each measurement does not really give the value of the observable under consideration but that – at least to some extend – the data are contaminated with noise. In order to successfully improve geophysically relevant field modelling, one main aspect is to denoise the data, i.e. to extract the true portion of the observable from the actual signal. In the case of hi–lo SST different approaches to noisy measurements by specifying a certain variance–covariance model as usual in physical geodesy are viable. In the case of geomagnetism, denoising of vectorial functions means that the noise should be handled differently in each vectorial component. This is due to the configuration of the magnetic instruments on the one hand and the methods of attitude determination on the other hand. Considering the (nearly spherical) orbit of the CHAMP satellite, for example, one expects higher noise levels in the tangential components than in the spherically radial component.

The objective of this article is to introduce spectral as well as multiscale signal-to-noise thresholding, providing the vector spherical harmonic and wavelet oriented basis of denoising spherical vector fields such as the gravitational field (cf. W. Freeden et al. (1999)) as well as geomagnetic field (cf. M. Bayer et al. (2000)). First, we introduce the corresponding theory of spectral and multiscale approximation. Seen from mathematical point of view spherical vector wavelets have to reflect the (vectorial) isotropy (i.e. rotational invariance of vector fields) of geophysically relevant (pseudo)differential operators representing the observables of gravitational and geomagnetic data. Thus, vector spherical harmonics have to be used in an appropriate framework of orthogonal expansions, while the spherical wavelets have to be generated by a concept of multiresolution analysis using rotation invariant scaling functions, i.e. radial basis functions. A first proposal towards the idea of isotropic multiscale vector field modelling was already made by W. Freeden et al. (1998) based on a tensorial concept by use of tensor radial basis functions. Later on, in restriction to the bilinear theory, M. Bayer et al. (1998) were able to distinguish a decomposition and reconstruction step which offers non–tensorial (i.e. scalar–vector) alternatives of vectorial wavelet approximation. It will be shown in this article that both approaches (i.e. the linear tensor–vector as well as the bilinear scalar–vector variant) can be used to denoise error–affected data. Altogether, spherical vector wavelets provide a tool of automatically adapting space and frequency localization in a multiresolution framework (i.e. zoom–in procedure) for spherical vector field reconstruction and decomposition. The signal–to–noise thresholding scheme is designed to distinguish between coefficients which contribute significantly to the signal, and those which are negligible.

It should be noted that our thresholding approaches are influenced by the concept of sparse wavelet representations in Euclidean spaces (cf. J.B. Weaver et al. (1991), D.L. Donoho, I.M Johnstone (1994, 1995), R.T. Ogden (1997)) and the stochastic variance–covariance model used for satellite data in physical geodesy (see e.g. R. Rummel (1997) and the references therein). By virtue of the multiscale approach we are thus able to include detail information of small spatial extend while suppressing the noise in suitable approximation.

The paper is organized as follows: Chapter 2 gives basic material on spectral expansion of vector fields by means of vector spherical harmonics. Two variants are distinguishable, viz. tensor-vector and scalar-vector representation. Chapter 3 deals with the introduction of vector and tensor radial basis functions. Scaling functions and vector wavelets are outlined as the constituting ingredients of multiscale approximation of vector fields in Chapter 4 and Chapter 5, respectively. Examples of bandlimited and non-bandlimited type can be found in Chapter 6. Signal-to-noise thresholding in its spectral formulation in terms of vector spherical harmonics is discussed in Chapter 7. In accordance with our vector spherical approach two variants of spectral denoising are describable, namely the tensor based and vector based formulation. Finally Chapter 8 explains selective multiscale reconstruction of spherical vector fields. The paper ends with an example of multiscale denoising in geomagnetic field determination.

2. Preliminaries. We begin by introducing some basic notation that will be used throughout the paper.

2.1. Notation. Let us use $x, y, \ldots$ to represent the elements of Euclidean space $\mathbb{R}^3$. For all $x \in \mathbb{R}^3$, $x = (x_1, x_2, x_3)^T$, different from the origin, we have

$$x = r\xi, \quad r = |x| = \sqrt{x_1^2 + x_2^2 + x_3^2}, \tag{2.1}$$

where $\xi = (\xi_1, \xi_2, \xi_3)^T$ is the uniquely determined directional unit vector of $x \in \mathbb{R}^3$. The unit sphere in $\mathbb{R}^3$ will be denoted by Ω. If the vectors $\varepsilon^1, \varepsilon^2, \varepsilon^3$ form the canonical orthonormal basis in $\mathbb{R}^3$, we may represent the points $\xi \in \Omega$ in polar coordinates by

$$\begin{aligned} \xi \quad &= \quad t\varepsilon^3 + \sqrt{1-t^2}\left(\cos\varphi\varepsilon^1 + \sin\varphi\varepsilon^2\right), \\ & \quad\quad -1 \leq t \leq 1, \quad 0 \leq \varphi < 2\pi, \quad t = \cos\theta \ . \end{aligned} \tag{2.2}$$

Inner, vector, and dyadic (tensor) product of two vectors $x, y \in \mathbb{R}^3$, respectively, are defined by

$$x \cdot y = x^T y = \sum_{i=1}^{3} x_i y_i, \tag{2.3}$$

$$x \wedge y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)^T, \tag{2.4}$$

$$x \otimes y = xy^T = \begin{pmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \\ x_3y_1 & x_3y_2 & x_3y_3 \end{pmatrix} . \tag{2.5}$$

As usual (cf. e.g. Gurtin (1972)), a *second order tensor* $\mathbf{f} \in \mathbb{R}^{3\times 3}$ is understood to be a linear mapping that assigns to each $x \in \mathbb{R}^3$ a vector $y \in \mathbb{R}^3$. The (cartesian) components F_{ij} of $\mathbf{f}$ are defind by

$$F_{ij} = \varepsilon^i \cdot (\mathbf{f}\varepsilon^j) = (\varepsilon^i)^T(\mathbf{f}\varepsilon^j), \tag{2.6}$$

so that $y = \mathbf{f}x$ is equivalent to

$$y \cdot \varepsilon^i = \sum_{j=1}^{3} F_{ij}(x \cdot \varepsilon^j) \ . \tag{2.7}$$

We write $\mathbf{f}^T$ for the transpose of $\mathbf{f}$; it is the unique tensor satisfying $(\mathbf{f}y)\cdot x = y \cdot (\mathbf{f}^T x)$ for all $x, y \in \mathbb{R}^3$. Moreover, we write $\mathrm{tr}(\mathbf{f})$ for the trace and $\det(\mathbf{f})$ for the determinant of $\mathbf{f}$.

The dyadic (tensor) product $x \otimes y$ of two elements $x, y \in \mathbb{R}^3$ (cf. (2.5)) is the tensor that assigns to each $u \in \mathbb{R}^3$ the vector $(y \cdot u)x$. More explicitly,

$$(x \otimes y)u = (y \cdot u)x \tag{2.8}$$

for every $u \in \mathbb{R}^3$.

The inner product $\mathbf{f} \cdot \mathbf{g}$ of two second order tensors $\mathbf{f}, \mathbf{g} \in \mathbb{R}^{3\times 3}$ is defined by

$$\mathbf{f} \cdot \mathbf{g} = \mathrm{tr}(\mathbf{f}^T\mathbf{g}) = \sum_{i,j=1}^{3} F_{ij}G_{ij}, \tag{2.9}$$

while

$$|\mathbf{f}| = (\mathbf{f} \cdot \mathbf{f})^{1/2} \tag{2.10}$$

is called the norm of $\mathbf{f}$.

Given any tensor $\mathbf{f}$ and any pair $x,\, y \in \mathbb{R}^3$ we have

$$x \cdot (\mathbf{f}y) = \mathbf{f} \cdot (x \otimes y) \ . \tag{2.11}$$

Furthermore, $(x \otimes y)\mathbf{f} = x \otimes \mathbf{f}^T y$. Moreover, for $x, y, w, z \in \mathbb{R}^3$, we have

$$(x \otimes y)(w \otimes z) = (y \cdot w)(x \otimes z) \ . \tag{2.12}$$

It is easy to see that

$$(\varepsilon^i \otimes \varepsilon^j) \cdot (\varepsilon^k \otimes \varepsilon^l) = \delta_{ik}\delta_{jl}, \tag{2.13}$$

so that the nine tensors $\varepsilon^i \otimes \varepsilon^j$ are orthonormal (δ_{ik} denotes the Kronecker symbol). Moreover, it follows that

$$\sum_{i,j=1}^{3} \left(F_{ij}\varepsilon^i \otimes \varepsilon^j\right) x = \sum_{i,j=1}^{3} F_{ij}(x \cdot \varepsilon^j)\varepsilon^i = \mathbf{f}x \tag{2.14}$$

and thus $\mathbf{f} \in \mathbb{R}^{3\times 3}$ with

$$\mathbf{f} = \sum_{i,j=1}^{3} F_{ij}\varepsilon^i \otimes \varepsilon^j \ . \tag{2.15}$$

In particular, the identity tensor $\mathbf{i}$ is given by $\mathbf{i} = \sum_{i=1}^{3} \varepsilon^i \otimes \varepsilon^i$. Moreover it is easy to see that

$$\mathrm{tr}(x \otimes y) = x \cdot y, \quad x, y \in \mathbb{R}^3, \tag{2.16}$$

and

$$\mathbf{f} \cdot (\mathbf{g}\mathbf{h}) = (\mathbf{g}^T\mathbf{f}) \cdot \mathbf{h} = (\mathbf{f}\mathbf{h}^T) \cdot \mathbf{g}, \qquad \mathbf{f}, \mathbf{g}, \mathbf{h} \in \mathbb{R}^{3\times 3} \ . \tag{2.17}$$

Next we come to some operators, which are of particular importance in the vector spherical context. In terms of the polar coordinates (2.2) the gradient ∇ in $\mathbb{R}^3$ reads

$$\nabla_x = \xi \frac{\partial}{\partial r} + \frac{1}{r}\nabla^*_\xi, \tag{2.18}$$

where ∇^* is the *surface gradient* of the unit sphere $\Omega \subset \mathbb{R}^3$. Moreover, the *Laplace operator* $\Delta = \nabla \cdot \nabla$ in $\mathbb{R}^3$ has the representation

$$\Delta_x = \left(\frac{\partial}{\partial r}\right)^2 + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_\xi, \tag{2.19}$$

where $\Delta^* = \nabla^* \cdot \nabla^*$ is the *Beltrami operator* of the unit sphere Ω (for explicit representations in terms of polar coordinates see e.g. W. Freeden et al. 1998).

Throughout this paper scalar valued (resp. vector valued, tensor valued) functions are denoted by capital (resp. small, small bold) letters.

A function $F : \Omega \to \mathbb{R}$ (resp. $f : \Omega \to \mathbb{R}^3$) possessing k continuous derivatives on the unit sphere Ω is said to be of class $\mathcal{C}^{(k)}(\Omega)$ (resp. $c^{(k)}(\Omega)$). $\mathcal{C}^{(0)}(\Omega)$ (resp. $c^{(0)}(\Omega)$) is the class of real continuous scalar–valued (resp. vector–valued) functions on Ω.

For $F \in C^{(1)}(\Omega)$ we introduce the *surface curl gradient* L^* by

$$L^*_\xi F(\xi) = \xi \wedge \nabla^*_\xi F(\xi), \quad \xi \in \Omega, \tag{2.20}$$

while $\nabla^*_\xi \cdot f(\xi)$, $\xi \in \Omega$, and $L^*_\xi \cdot f(\xi)$, $\xi \in \Omega$, respectively, denote the *surface divergence* and *surface curl* of the vector field f at $\xi \in \Omega$.

It is worth mentioning that the operators ∇^*, L^*, Δ^* will be always used in coordinate–free representation throughout this paper (thereby avoiding any singularity at the poles). Nevertheless, for the convenience of the reader, we give a list of their expressions in local coordinates (2.2):

$$\begin{aligned}\nabla^*_\xi = {} & \frac{1}{\sqrt{1-t^2}}\left(-\sin\varphi\ \varepsilon^1 + \cos\varphi\ \varepsilon^2\right)\frac{\partial}{\partial\varphi}, \\ & + \sqrt{1-t^2}\left(-t\cos\varphi\ \varepsilon^1 - t\sin\varphi\ \varepsilon^2 + \sqrt{1-t^2}\varepsilon^3\right)\frac{\partial}{\partial t},\end{aligned} \tag{2.21}$$

$$\begin{aligned}L^*_\xi = {} & \sqrt{1-t^2}\left(\sin\varphi\ \varepsilon^1 - \cos\varphi\ \varepsilon^2\right)\frac{\partial}{\partial t} \\ & + \frac{1}{\sqrt{1-t^2}}\left(-t\cos\varphi\ \varepsilon^1 - t\sin\varphi\ \varepsilon^2 + \sqrt{1-t^2}\ \varepsilon^3\right)\frac{\partial}{\partial t},\end{aligned} \tag{2.22}$$

$$\Delta^*_\xi = \frac{\partial}{\partial t}(1-t^2)\frac{\partial}{\partial t} + \frac{1}{1-t^2}\left(\frac{\partial}{\partial\varphi}\right)^2 \quad . \tag{2.23}$$

The operators $o^{(i)} : C^{(1)}(\Omega) \to c(\Omega)$, $i = 1,2,3$, defined by

$$o^{(1)}_\xi F(\xi) = \xi F(\xi), \qquad \xi \in \Omega, \tag{2.24}$$

$$o^{(2)}_\xi F(\xi) = \nabla^*_\xi F(\xi), \qquad \xi \in \Omega, \tag{2.25}$$

$$o^{(3)}_\xi F(\xi) = L^*_\xi F(\xi), \qquad \xi \in \Omega, \tag{2.26}$$

are of particular importance for our considerations. Therefore we discuss some of their properties in more detail. For all $\xi \in \Omega$ we have

$$o^{(i)}_\xi F(\xi) \cdot o^{(j)}_\xi F(\xi) = 0 \tag{2.27}$$

whenever $j \neq i$, $i, j \in \{1,2,3\}$. Moreover, if $G \in C^{(1)}[-1,+1]$ and $(\xi,\eta) \in \Omega \times \Omega$ it is not hard to see that

$$o^{(2)}_\xi G(\xi\cdot\eta) = G'(\xi\cdot\eta)(\eta - (\xi\cdot\eta)\xi), \tag{2.28}$$

$$o^{(3)}_\xi G(\xi\cdot\eta) = G'(\xi\cdot\eta)(\xi \wedge \eta) \quad . \tag{2.29}$$

Furthermore, Green's identities show us that

$$\begin{aligned}F(\xi) = {} & \frac{1}{4\pi}\int_\Omega F(\eta)d\omega(\eta) \\ & - \int_\Omega \left(o^{(i)}_\eta G(\Delta^*;\xi,\eta)\right)\cdot\left(o^{(i)}_\eta F(\eta)\right)d\omega(\eta)\end{aligned} \tag{2.30}$$

holds for all $\xi \in \Omega$, $i \in \{2,3\}$ and $F \in C^{(1)}(\Omega)$, where $G(\Delta^*;\cdot,\cdot)$ is the Green function with respect to the Beltrami operator Δ^* (cf. W. Freeden (1980))

$$G(\Delta^*;\xi,\eta) = \frac{1}{4\pi}\ln(1-\xi\cdot\eta) + \frac{1}{4\pi} - \frac{1}{4\pi}\ln 2, \quad -1 \le \xi\cdot\eta < 1 \ . \tag{2.31}$$

The integral equations (cf. W. Freeden et al. (1998))

$$\int_\Omega f(\xi)\cdot\nabla^*_\xi F(\xi)\ d\omega(\eta) = -\int_\Omega F(\xi)\nabla^*_\xi\cdot f(\xi)\ d\omega(\eta), \tag{2.32}$$

$$\int_\Omega f(\xi)\cdot L^*_\xi F(\xi)\ d\omega(\eta) = -\int_\Omega F(\xi)L^*_\xi\cdot f(\xi)\ d\omega(\eta) \tag{2.33}$$

lead us to operators $O^{(i)} : c^{(1)}(\Omega) \to C^{(0)}(\Omega)$, $i = 1,2,3$, which are adjoint to $o^{(i)}$. More explicitly, for $f \in c^{(1)}(\Omega)$ and $F \in C^{(1)}(\Omega)$, we have

$$\int_\Omega o^{(i)}_\xi F(\xi)\cdot f(\xi)\ d\omega(\eta) = \int_\Omega F(\xi)\ O^{(i)}_\xi f(\xi)\ d\omega(\eta), \qquad i = 1,2,3, \tag{2.34}$$

where

$$O^{(1)}_\xi f(\xi) = \xi\cdot f(\xi), \qquad \xi\in\Omega, \tag{2.35}$$

$$O^{(2)}_\xi f(\xi) = -\nabla^*_\xi\cdot f(\xi), \qquad \xi\in\Omega, \tag{2.36}$$

$$O^{(3)}_\xi f(\xi) = -L^*_\xi\cdot f(\xi), \qquad \xi\in\Omega \ . \tag{2.37}$$

It can be easily seen that

$$O^{(i)}_\xi o^{(j)}_\xi F(\xi) = 0, \quad i \ne j, \qquad i,j\in\{1,2,3\} \tag{2.38}$$

and

$$O^{(i)}o^{(i)}F(\xi) = \begin{cases} F(\xi) & , \quad i = 1 \\ -\Delta^*_\xi F(\xi) & , \quad i = 2,3 \end{cases} \tag{2.39}$$

provided that $F \in C^{(2)}(\Omega)$. For more details the reader is referred to W. Freeden et al. (1998).

2.2. Vector Spherical Harmonics. Any function $f : \Omega \to \mathbb{R}^3$ is called *a spherical vector field*. By $l^2(\Omega)$ we denote the space of (Lebesgue) square–integrable vector fields on Ω, i.e.

$$l^2(\Omega) = \left\{ f : \Omega \to \mathbb{R}^3 \,\middle|\, \int_\Omega f(\xi)\cdot f(\xi)\ d\omega(\xi) < \infty \right\} \ . \tag{2.40}$$

$l^2(\Omega)$ is a Hilbert space equipped with the inner product

$$(f,g)_{l^2} = \left(\int_\Omega f(\xi)\cdot g(\xi)\ d\omega(\xi)\right)^{1/2} \ . \tag{2.41}$$

$c^{(k)}(\Omega)$ denotes the space of vector fields with k–times continuously differentiable components on Ω.

For a given vector field $f : \Omega \to \mathbb{R}^3$

$$f_{nor} : \xi \mapsto f_{nor}(\xi) = (\xi \cdot f(\xi))\, \xi, \quad \xi \in \Omega \tag{2.42}$$

is called the *normal part* of f, while

$$f_{tan} : \xi \mapsto f_{tan}(\xi) = f(\xi) - f_{nor}(\xi), \quad \xi \in \Omega \tag{2.43}$$

is called the *tangential part* of f. A vector field f is called *tangential* (resp. *normal*), if $f(\xi) = f_{tan}(\xi)$ (resp. $f(\xi) = f_{nor}(\xi)$) for all $\xi \in \Omega$.

The study of vector fields on the sphere can be greatly simplified by the *Helmholtz decomposition theorem* for continuously differentiable vector fields $f : \Omega \to \mathbb{R}^3$

$$f(\xi) = f_{nor}(\xi) + f_{tan}(\xi), \quad \xi \in \Omega\ . \tag{2.44}$$

To be more precise, any continuously differentiable vector field on the unit sphere $\Omega \subset \mathbb{R}^3$ (i.e.: $f \in c^{(1)}(\Omega)$) may be represented by a decomposition in terms of scalar functions $F^{(i)} \in C^{(2)}(\Omega)$, $i = 1, 2, 3$, such that

$$f_{nor}(\xi) = o_\xi^{(1)} F^{(1)}(\xi), \quad \xi \in \Omega \tag{2.45}$$

$$f_{tan}(\xi) = o_\xi^{(2)} F^{(2)}(\xi) + o_\xi^{(3)} F^{(3)}(\xi), \quad \xi \in \Omega, \tag{2.46}$$

where

$$F^{(1)}(\xi) = \xi \cdot f(\xi)\ , \qquad \xi \in \Omega, \tag{2.47}$$

$$F^{(2)}(\xi) = -\int_\Omega G(\Delta^*; \xi, \eta)\, O_\eta^{(2)} f(\eta)\, d\omega(\eta), \quad \xi \in \Omega, \tag{2.48}$$

$$F^{(3)}(\xi) = -\int_\Omega G(\Delta^*; \xi, \eta)\, O_\eta^{(3)} f(\eta)\, d\omega(\eta), \quad \xi \in \Omega\ . \tag{2.49}$$

These representations of spherical vector fields lay the groundwork for the main subject of this note. The explicit representations of the tangential as well as the normal field, in fact, are essential for the constructive approximation of spherical vector fields.

2.3. Spectral Approximation. First we describe (equivalent) ways of expanding a vector field in terms of vector spherical harmonics.

2.3.1. Tensor-Vector Representation. The orthogonal (Fourier) expansion of the functions $F^{(i)}$, $i \in \{1, 2, 3\}$, in terms of an $\mathcal{L}^2(\Omega)$-orthonormal system $\{Y_{n,k}\}_{\substack{n=0,1,\ldots \\ k=1,\ldots,2n+1}}$ of *scalar spherical harmonics* $Y_{n,k}$

$$F^{(i)} = \sum_{n=0}^{\infty} \sum_{k=1}^{2n+1} (F^{(i)})^\wedge(n,k) Y_{n,k}, \quad i = 1, 2, 3, \tag{2.50}$$

with orthogonal coefficients

$$(F^{(i)})^\wedge(n,k) = (F^{(i)}, Y_{n,k})_{\mathcal{L}^2(\Omega)} = \int_\Omega F^{(i)}(\eta) Y_{n,k}(\eta)\, d\omega(\eta) \tag{2.51}$$

leads to an orthogonal expansion of the vector field f in terms of *vector spherical harmonics* $\{y_{n,k}^{(i)}\}$ *of type* i given by

$$y_{n,k}^{(i)} = (\mu_n^{(i)})^{-1/2} o^{(i)} Y_{n,k}\ , \qquad n = 0_i, 0_i + 1, \ldots; k = 1, \ldots, 2n+1 \tag{2.52}$$

($0_1 = 0,\ 0_i = 1,\ i = 2,3$, and $\mu_n^{(1)} = 1,\ \mu_n^{(i)} = n(n+1),\ i = 2,3$) of the following form:

$$f = f^{(1)} + f^{(2)} + f^{(3)}, \tag{2.53}$$

where

$$f^{(i)} = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (f^{(i)})^\wedge(n,k)\ y_{n,k}^{(i)}, \qquad i = 1,2,3, \tag{2.54}$$

where $(f^{(i)})^\wedge(n,k)$ are the orthogonal coefficients in terms of vector spherical harmonics

$$(f^{(i)})^\wedge(n,k) = (f, y_{n,k}^{(i)})_{l^2(\Omega)} = \int_\Omega f(\eta) \cdot y_{n,k}^{(i)}(\eta)\ d\omega(\eta). \tag{2.55}$$

Then any $f \in l^{(2)}(\Omega)$ may be represented in the form

$$f = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (f^{(i)})^\wedge(n,k) y_{n,k}^{(i)} \tag{2.56}$$

where the equality is understoood in the sense of $\|\cdot\|_{l^2(\Omega)}$. Using the addition theorem for vector spherical harmonics (see W. Freeden, T. Gervens (1991)) the vectorial expansion (2.53), (2.54) may be rewritten in the form

$$f = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} \int_\Omega \mathbf{p}_n^{(i,i)}(\cdot,\eta) f^{(i)}(\eta) d\omega(\eta), \tag{2.57}$$

where the *i-th Legendre tensor function* $\mathbf{p}_n^{(i,i)} : \Omega \times \Omega \to \mathbb{R}^3 \otimes \mathbb{R}^3$ reads as follows:

$$\begin{aligned} &\sum_{k=1}^{2n+1} y_{n,k}^{(i)}(\xi) \otimes y_{n,k}^{(i)}(\eta) \\ &= \frac{2n+1}{4\pi} \mathbf{p}_n^{(i,i)}(\xi,\eta) = (\mu_n^{(i)})^{-1} \frac{2n+1}{4\pi} o_\xi^{(i)} o_\eta^{(i)} P_n(\xi \cdot \eta) \end{aligned} \tag{2.58}$$

with

$$\mathbf{p}_n^{(1,1)}(\xi,\eta) = P_n(\xi \cdot \eta) \xi \otimes \eta, \tag{2.59}$$

$$\mathbf{p}_n^{(2,2)}(\xi,\eta) = \frac{1}{n(n+1)} \left(P_n''(\xi \cdot \eta)(\eta - (\xi \cdot \eta)\xi) \otimes (\xi - (\xi \cdot \eta)\eta)\right. \tag{2.60}$$

$$\left. + P_n'(\xi \cdot \eta)(\mathbf{i} - \xi \otimes \xi - (\eta - (\xi \cdot \eta)\xi) \otimes \eta)\right), \tag{2.61}$$

$$\mathbf{p}_n^{(3,3)}(\xi,\eta) = \frac{1}{n(n+1)} \left(P_n''(\xi \cdot \eta) \xi \wedge \eta \otimes \eta \wedge \xi\right. \tag{2.62}$$

$$\left. + P_n'(\xi \cdot \eta)((\xi \cdot \eta)(\mathbf{i} - \xi \otimes \xi) - (\eta - (\xi \cdot \eta)\xi) \otimes \xi)\right) \tag{2.63}$$

for all $(\xi,\eta) \in \Omega \times \Omega$. P_n is the scalar Legendre polynomial of degree n given via the addition theorem by

$$P_n(\xi \cdot \eta) = \frac{4\pi}{2n+1} \sum_{k=1}^{2n+1} Y_{n,k}(\xi) Y_{n,k}(\eta), \quad (\xi,\eta) \in \Omega \times \Omega, \tag{2.64}$$

for any scalar $\mathcal{L}^2(\Omega)$-orthonormal system $\{Y_{n,k}\}_{k=1,\dots,2n+1}$. Once again observe that

$$\nabla^*_\xi P_n(\xi\cdot\eta) = P'_n(\xi\cdot\eta)(\eta-(\xi\cdot\eta)\xi), \tag{2.65}$$

$$L^*_\xi P_n(\xi\cdot\eta) = P'_n(\xi\cdot\eta)(\xi\wedge\eta) \tag{2.66}$$

for all $(\xi,\eta)\in\Omega^2$ and $n=1,2,\dots$.

Using an orthogonal expansion in terms of vector spherical harmonics as described above, we have three orthogonal vector fields $y^{(i)}_{n,k}$ that depend on one scalar polynomial, namely the scalar spherical harmonic $Y_{n,k}$ such that the vectorial system $\{y^{(i)}_{n,k}\}$ is able to generate both the normal field ($i=1$) and the tangential fields ($i=2,3$) on the sphere. In doing so, the Hilbert space $l^2(\Omega)$ of square–integrable vector fields admits a geophysically motivated orthogonal decomposition into three orthogonal subspaces $l^2_{(i)}(\Omega)$; the first subspace $l^2_{(1)}(\Omega)$ only consists of square-integrable normal fields, while the second subspace $l^2_{(2)}(\Omega)$ and third subspace $l^2_{(3)}(\Omega)$ consist of tangential vector fields that are curl-free and divergence-free, respectively:

$$l^2(\Omega) = l^2_{nor}(\Omega)\oplus l^2_{tan}(\Omega), \tag{2.67}$$

with

$$l^2_{nor}(\Omega) = l^2_{(1)}(\Omega), \qquad l^2_{tan}(\Omega) = l^2_{(2)}(\Omega)\oplus l^2_{(3)}(\Omega). \tag{2.68}$$

2.4. Scalar-Vector Representation. Tensor representations as proposed in Section 2.3.1 are difficult to handle in numerical computations. In what follows, therefore, we give an alternative approach of expanding vector fields in terms of spherical harmonics. Its key idea is based on the fact that the vectorial expansion (2.53), (2.54) may be written by use of the adjoint operators in the form

$$f(\xi) = \sum_{i=1}^{3}\sum_{n=0_i}^{\infty}\frac{2n+1}{4\pi}\left(\mu^{(i)}_n\right)^{-1}\int_\Omega o^{(i)}_\xi P_n(\xi\cdot\eta)\; O^{(i)}_\eta f^{(i)}(\eta)\; d\omega(\eta),\ \xi\in\Omega, \tag{2.69}$$

or equivalently

$$f(\xi) = \sum_{i=1}^{3}\sum_{n=0_i}^{\infty}\frac{2n+1}{4\pi}\left(\mu^{(i)}_n\right)^{-1/2}\int_\Omega p^{(i)}_n(\xi,\eta)\; O^{(i)}_\eta f^{(i)}(\eta)\; d\omega(\eta), \quad \xi\in\Omega, \tag{2.70}$$

where the *i-th Legendre vector* $p^{(i)}_n:\Omega\times\Omega\to\mathbb{R}^3$ is given by

$$\begin{aligned}\sum_{k=1}^{2n+1} & y^{(i)}_{n,k}(\xi)Y_{n,k}(\eta) \\ &= p^{(i)}_n(\xi,\eta) = \left(\mu^{(i)}_n\right)^{-1/2} o^{(i)}_\xi P_n(\xi\cdot\eta), \qquad (\xi,\eta)\in\Omega^2\ .\end{aligned} \tag{2.71}$$

The Legendre vector functions written out read as follows:

$$\begin{aligned}
p^{(1)}_n(\xi,\eta) &= \xi P_n(\xi\cdot\eta), \qquad n=0,1,\dots,\\
p^{(2)}_n(\xi,\eta) &= \frac{1}{\sqrt{n(n+1)}}(\eta-(\xi\cdot\eta)\,\xi)P'_n(\xi\cdot\eta), \qquad n=1,2,\dots,\\
p^{(3)}_n(\xi,\eta) &= \frac{1}{\sqrt{n(n+1)}}(\xi\wedge\eta)P'_n(\xi\cdot\eta), \qquad n=1,2,\dots
\end{aligned}$$

for $(\xi,\eta)\in\Omega^2$.

3. Vector and Tensor Radial Basis Functions. The relations between the different types of Legendre functions are as follows:

$$\begin{aligned} p_n^{(i)}(\xi,\eta) &= (\mu_n^{(i)})^{-1/2} O_\eta^{(i)} \mathbf{p}_n^{(i,i)}(\xi,\eta), \\ \mathbf{p}_n^{(i,i)}(\xi,\eta) &= (\mu_n^{(i)})^{-1/2} o_\eta^{(i)} p_n^{(i)}(\xi,\eta), \\ P_n(\xi\cdot\eta) &= (\mu_n^{(i)})^{-1/2} O_\xi^{(i)} p_n^{(i)}(\xi,\eta) \end{aligned}$$

for $(\xi,\eta) \in \Omega\times\Omega$ and $n = 0_i, 0_i+1, \ldots$. These identities lead us to the definition of *tensor and vector radial basis functions.*

We begin with the introduction of tensor radial basis functions (cf. W. Freeden et al. (1998), S. Beth (2000)).

Definition 3.1. *Any function* $\mathbf{k}^{(i)} : \Omega\times\Omega \to \mathbb{R}^3 \otimes \mathbb{R}^3$, $i \in \{1,2,3\}$, *of the form*

$$\mathbf{k}^{(i)}(\xi,\eta) = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (\mathbf{k}^{(i)})^\wedge(n) \mathbf{p}_n^{(i,i)}(\xi,\eta), \qquad (\xi,\eta) \in \Omega\times\Omega,$$

is called (square–summable) tensor radial basis function of type i if its symbol $\left\{(\mathbf{k}^{(i)})^\wedge(n)\right\}_{n=0_i,0_i+1,\ldots} \subset \mathbb{R}$ *satisfies the condition:*

$$\sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} \left((\mathbf{k}^{(i)})^\wedge(n)\right)^2 < \infty \ .$$

$\mathbf{k} = \sum_{i=1}^{3} \mathbf{k}^{(i)}$ *with* $\mathbf{k}^{(i)}$ *(square–summable) tensor radial basis functions of type i is called (square–summable) tensor radial basis function.*

A key property of a tensor radial basis function $\mathbf{k}$ is its invariance under orthogonal transformations $\mathbf{t}$, i.e.,

$$\mathbf{k}(\mathbf{t}\xi,\mathbf{t}\eta) = \mathbf{t}\mathbf{k}(\xi,\eta)\mathbf{t}^T, \qquad (\xi,\eta) \in \Omega\times\Omega \ .$$

This property falls back upon the Legendre tensors

$$\mathbf{p}_n^{(i,i)}(\mathbf{t}\xi,\mathbf{t}\eta) = \mathbf{t}\mathbf{p}_n^{(i,i)}(\xi,\eta)\mathbf{t}^T, \ \ (\xi,\eta) \in \Omega\times\Omega \ .$$

If a vector field $f : \Omega \to \mathbb{R}^3$ is invariant under orthogonal transformations, i.e.

$$f(\mathbf{t}\xi) = \mathbf{t}f(\xi), \quad \xi \in \Omega \tag{3.1}$$

for all orthogonal $\mathbf{t}$ which leave $\eta \in \Omega$ fixed, then the same applies for a vector field $g : \Omega \to \mathbb{R}^3$ given by

$$g(\xi) = \int_\Omega \mathbf{k}(\xi,\eta) f(\eta)\, d\omega(\eta) = \sum_{i=1}^{3} \int_\Omega \mathbf{k}^{(i)}(\xi,\eta) f(\eta)\, d\omega(\eta) \ .$$

Therefore, the definition of convolutions on the set of tensor radial basis functions makes sense.

Definition 3.2. *Let* $\mathbf{k}^{(i)}, \mathbf{h}^{(i)}$ *be (square–summable) tensor radial basis functions of type i. Suppose that f is of class $l^2(\Omega)$. Then $\mathbf{k}^{(i)} * f$ defined by*

$$\left(\mathbf{k}^{(i)} * f\right)(\xi) = \int_\Omega \mathbf{k}^{(i)}(\xi,\eta) f(\eta)\, d\omega(\eta), \quad \xi \in \Omega,$$

is called the convolution of $\mathbf{k}^{(i)}$ *against* f*. Furthermore,* $\mathbf{h}^{(i)} * \mathbf{k}^{(i)}$ *defined by*

$$(\mathbf{h}^{(i)} * \mathbf{k}^{(i)})(\xi,\eta) = \int_\Omega \mathbf{h}^{(i)}(\xi,\zeta)\mathbf{k}^{(i)}(\zeta,\eta)\ d\omega(\zeta), \quad (\xi,\eta) \in \Omega \times \Omega,$$

is said to be the '$*$' convolution of $\mathbf{h}^{(i)}$ *against* $\mathbf{k}^{(i)}$*. We let* $\mathbf{h} * \mathbf{k} = \sum_{i=1}^3 \mathbf{h}^{(i)} * \mathbf{k}^{(i)}$.

Obviously, we have

$$\mathbf{k}^{(i)} * f = \sum_{n=0_i}^{\infty} (\mathbf{k}^{(i)})^\wedge(n) \sum_{k=1}^{2n+1} (f^{(i)})^\wedge(n,k) y_{n,k}^{(i)}$$

and

$$\mathbf{h}^{(i)} * \mathbf{k}^{(i)} = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (\mathbf{h}^{(i)})^\wedge(n)(\mathbf{k}^{(i)})^\wedge(n)\mathbf{p}_n^{(i,i)} \ .$$

In particular, any $f \in l^2(\Omega)$ can be expressed as follows:

$$f = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi}\ \mathbf{p}_n^{(i,i)} * f \ .$$

Next we come to the discussion of vector radial basis functions, which will be based on vector Legendre functions (cf. M. BAYER et al. (1998), S. BETH (2000)).

DEFINITION 3.3. *Any function* $k^{(i)} : \Omega \times \Omega \to \mathbb{R}^3$, $i \in \{1,2,3\}$*, defined by*

$$k^{(i)}(\xi,\eta) = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (k^{(i)})^\wedge(n) p_n^{(i)}(\xi,\eta)$$

is called (square–summable) vector radial basis function of type i *if its symbol* $\{(k^{(i)})^\wedge(n)\}_{n=0_i,0_i+1,\ldots} \subset \mathbb{R}$ *satisfies the condition:*

$$\sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} \left((k^{(i)})^\wedge(n)\right)^2 < \infty \ .$$

$k = \sum_{i=1}^3 k^{(i)}$ *with* $k^{(i)}$ *(square–summable) vector radial basis functions of type* i *is called (square–summable) vector radial basis function.*

According to the occurence of scalar as well as vector spherical harmonics in the series expansion of vector radial basis functions two different types of convolutions are definable for vector radial basis functions.

DEFINITION 3.4. *Let* $h^{(i)}, k^{(i)}$ *be vector radial basis functions of type* i, $i \in \{1,2,3\}$*. Suppose that* f *is of class* $l^2(\Omega)$*. Then* $k^{(i)} * f$, $i \in \{1,2,3\}$*, defined by*

$$k^{(i)} * f = \int_\Omega k^{(i)}(\eta,\cdot) \cdot f(\eta)\ d\omega(\eta)$$

is called the convolution of $k^{(i)}$ *against* f*.*

Moreover, assume that F *is of class* $\mathcal{L}^2(\Omega)$*. Then* $k^{(i)} \star F$, $i \in \{1,2,3\}$*, given by*

$$k^{(i)} \star F = \int_\Omega k^{(i)}(\cdot,\eta) F(\eta)\ d\omega(\eta)$$

is called the convolution of $k^{(i)}$ against F. Furthermore, by definition, we let

$$h \star k * f = \sum_{i=1}^{3} h^{(i)} \star k^{(i)} * f$$

Looking at the types of functions involved in the aforementioned convolutions it becomes obvious which parts interact: '$*$' denotes a scalar product, whereas '$\star$' represents a scalar-vector multiplication.

The following theorem, yields within the framework of convolution integrals, an alternative in decomposing a tensor radial basis function either by tensor radial basis functions or by vector radial basis functions (cf. M. Bayer et al. (1998), S. Beth (2000)).

Theorem 3.5. *Let f be of class $l^2(\Omega)$. Assume that h, k are (square–summable) vector radial basis functions. Moreover, suppose that $\mathbf{h}, \mathbf{k}$ are (square–summable) tensor radial basis functions with*

$$(\mathbf{h}^{(i)})^\wedge(n) = (h^{(i)})^\wedge(n), \tag{3.2}$$

$$(\mathbf{k}^{(i)})^\wedge(n) = (k^{(i)})^\wedge(n), \tag{3.3}$$

for all $n = 0_i, 0_i + 1, \ldots$ and $i \in \{1, 2, 3\}$. Then

$$\mathbf{h} * \mathbf{k} * f = h \star k * f \ .$$

Proof. The completeness of the vector spherical harmonics (cf. W. Freeden et al. (1998)) enables us to verify that

$$\begin{aligned}
&\mathbf{h} * \mathbf{k} * f \\
&= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} (\mathbf{h}^{(i)})^\wedge(n) (\mathbf{k}^{(i)})^\wedge(n) \sum_{k=1}^{2n+1} (f^{(i)})^\wedge(n,k) y_{n,k}^{(i)} \\
&= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} (h^{(i)})^\wedge(n) (k^{(i)})^\wedge(n) \sum_{k=1}^{2n+1} (f^{(i)})^\wedge(n,k) y_{n,k}^{(i)} \\
&= h \star k * f \ .
\end{aligned}$$

This is the desired result. □

In other words, our considerations have shown that the different types of bilinear convolutions lead to equivalent results. The tensorial concept of convolutions as introduced above enabled W. Freeden, T. Gervens (1991), W. Freeden et al. (1998) to derive a variational theory of vector spherical spline approximation. Furthermore, the scalar spherical wavelet theory (see W. Freeden et al. (1998) and the references therein) admits a canonical generalization to the vector nomenclature. Vector radial basis functions have been discussed by M. Bayer et al. (1998), M. Bayer (2000) and S. Beth (2000). The vector radial basis functions show the advantage that their convolutions against scalar or vector fields are easier calculable. However, it should be mentioned that they have the disadvantage that they can be handled only in bilinear expression, whereas tensor radial basis functions can be used canonically in a linear as well as bilinear (wavelet) concept (cf. W. Freeden et al. (1998), S. Beth(2000)).

Theorem 3.5 gives rise to introduce the following definition.

Definition 3.6. *Let $h^{(i)}$, $k^{(i)}$, $i \in \{1,2,3\}$, be (square–summable) vector radial basis functions of type i. Then $h^{(i)} \star k^{(i)}$ defined by*

$$(h^{(i)} \star k^{(i)})(\xi,\eta) = \int_\Omega h^{(i)}(\xi,\zeta) \otimes k^{(i)}(\eta,\zeta)\ d\omega(\zeta)$$

is said to be the '$\star$'–convolution of $h^{(i)}$ against $k^{(i)}$. We let $h \star k = \sum_{i=1}^{3} h^{(i)} \star k^{(i)}$. Clearly, for $i = 1,2,3$, we find

$$h^{(i)} \star k^{(i)} = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (h^{(i)})^\wedge(n)(k^{(i)})^\wedge(n)\mathbf{p}_n^{(i,i)} = \mathbf{h}^{(i)} * \mathbf{k}^{(i)} \tag{3.4}$$

provided that (3.2), (3.3) hold true.

Finally it is worth mentioning the following lemma.

Lemma 3.7. *Let $\mathbf{k}$ be a (square–summable) tensor radial basis function such that*

$$\sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (n+1/2)^{2\ 0_i} \left((\mathbf{k}^{(i)})^\wedge(n) \right)^2 < \infty$$

for $i = 1,2,3$. Then $O_\eta^{(i)}\mathbf{k}(\cdot,\eta) \in l^2_{(i)}(\Omega)$, $\eta \in \Omega$, and $\varepsilon^q \cdot O_\eta^{(i)}\mathbf{k}(\cdot,\eta) \in \mathcal{L}^2(\Omega)$, $\eta \in \Omega$, with $i,q \in \{1,2,3\}$.

The proof follows easily from the definition of the $O^{(i)}$–symbols in connection with the properties of (square–summable) tensor radial basis functions.

4. Scaling Functions. Let $(\Phi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$, $i = 1,2,3$, be three functions satisfying the following conditions:

(i) $(\Phi_0^{(i)})^\wedge(n)$ is monotonically decreasing,

(ii) $(\Phi_0^{(i)})^\wedge(n)$ is piecewise continuous in $(0,\infty)$ and continuous at 0,

(iii) $(\Phi_0^{(i)})^\wedge(0) = 1$,

(iv) $\sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} |(\Phi_0^{(i)})^\wedge(n)|^2 < \infty$.

Then the vector function $(\Phi_0)^\wedge(n) = \Big((\Phi_0^{(1)})^\wedge(n), (\Phi_0^{(2)})^\wedge(n), (\Phi_0^{(2)})^\wedge(n)\Big)^T$ is called the *generator of the scale discrete scaling function* Φ_0 defined by

$$\mathbf{\Phi}_0(\xi,\eta) = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} (\Phi_0^{(i)})^\wedge(n)\, y_{n,l}^{(i)}(\xi) \otimes y_{n,l}^{(i)}(\eta), \qquad (\xi,\eta) \in \Omega^2 \quad . \tag{4.1}$$

Correspondingly the *scale discrete scaling function of type* i reads as follows:

$$\mathbf{\Phi}_0^{(i)}(\xi,\eta) = \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} (\Phi_0^{(i)})^\wedge(n)\, y_{n,l}^{(i)}(\xi) \otimes y_{n,l}^{(i)}(\eta), \qquad i = 1,2,3, \qquad (\xi,\eta) \in \Omega^2,$$

i.e.: $(\mathbf{\Phi}_0^{(i)})^\wedge(n) = (\Phi_0^{(i)})^\wedge(n)$, $n = 0_i, 0_i+1, \ldots$ and $i \in \{1,2,3\}$.

We introduce the *scale discrete dilation operator* D_j by

$$D_j(\Phi_0^{(i)})^\wedge(x) = (\Phi_j^{(i)})^\wedge(x) = (\Phi_0^{(i)})^\wedge(2^{-j}x), \; i = 1,2,3, \qquad j \in \mathbb{Z}. \tag{4.2}$$

We let

$$D_j(\Phi_0)^\wedge(x) = (\Phi_j)^\wedge(x) = \Big((\Phi_j^{(1)})^\wedge(x), (\Phi_j^{(2)})^\wedge(x), (\Phi_j^{(3)})^\wedge(x)\Big)^T, \; j \in \mathbb{Z}. \tag{4.3}$$

One important fact is that, by virtue of condition (iv), the dilates $(\Phi_j^{(i)})^\wedge(n)$, $i = 1,2,3$, satisfy

$$\sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} |(\Phi_j^{(i)})^\wedge(n)|^2 < \infty \tag{4.4}$$

for every $j \in \mathbb{Z}$, provided that the vector function $(\Phi_0)^\wedge(n)$ has this property (cf. W. Freeden, M. Schreiner (1997)). The definition of the dilated scaling function is straightforward by letting

$$D_j\mathbf{\Phi}_0(\xi,\eta) = \mathbf{\Phi}_j(\xi,\eta) = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} (\Phi_j^{(i)})^\wedge(n) y_{n,l}^{(i)}(\xi) \otimes y_{n,l}^{(i)}(\eta), \; (\xi,\eta) \in \Omega^2, \; j \in \mathbb{Z}.$$

We now come to the formulation of the *vectorial approximate identity* which is the central result in multiscale approximation.

Theorem 4.1. *Suppose that f is of class $l^2(\Omega)$. Let $\mathbf{\Phi}_0$ be a scale discrete scaling function generated by $(\Phi_0)^\wedge(n)$ (as described above). Then*

$$\lim_{J\to\infty} \|\mathbf{\Phi}_J^{(i)} * \mathbf{\Phi}_J^{(i)} * f - f^{(i)}\|_{l^2(\Omega)} = 0, \qquad i = 1,2,3, \tag{4.5}$$

and

$$\lim_{J\to\infty} \|\mathbf{\Phi}_J * \mathbf{\Phi}_J * f - f\|_{l^2(\Omega)} = 0. \tag{4.6}$$

Proof. For $i \in \{1,2,3\}$ and $J \in \mathbb{N}_0$ we have in the notation introduced before

$$\lim_{J\to\infty} \left\| \sum_{i=1}^{3} \mathbf{\Phi}_J^{(i)} * \mathbf{\Phi}_J^{(i)} * f - f \right\|_{l^2(\Omega)}^2 \tag{4.7}$$
$$= \lim_{J\to\infty} \int_\Omega \left| \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} \left(((\Phi_J^{(i)})^\wedge(n))^2 - 1 \right) (f^{(i)})^\wedge(n,l) y_{n,l}^{(i)}(\xi) \right|^2 d\omega(\xi).$$

According to our definition $(\Phi_J^{(i)})^\wedge(0) = 1$ and $(\Phi_J^{(i)})^\wedge(n)$ is monotonically decreasing. In conclusion, $((\Phi_J^{(i)})^\wedge(n))^2 - 1$ is smaller than one, and it is allowed to interchange sum and limit. But this yields the desired approximate identity. □

Consider the operators $p_j^{(i)} : l^2(\Omega) \to l^2_{(i)}(\Omega)$, $i = 1,2,3$, and $p_j : l^2(\Omega) \to l^2(\Omega)$ defined by

$$p_j^{(i)} = \mathbf{\Phi}_j^{(i)} * \mathbf{\Phi}_j^{(i)} * f, \qquad j \in \mathbb{Z}, \qquad i \in \{1,2,3\} \tag{4.8}$$

and

$$p_j = \mathbf{\Phi}_j * \mathbf{\Phi}_j * f, \qquad j \in \mathbb{Z}, \tag{4.9}$$

respectively. The corresponding scale spaces $s_j^{(i)}$, s_j are defined by

$$s_j^{(i)} = \{p_j^{(i)}(f) | f \in l^2(\Omega)\}, \qquad i \in \{1,2,3\} \tag{4.10}$$

and

$$s_j = \{p_j(f) | f \in l^2(\Omega)\}, \tag{4.11}$$

respectively. This leads to a *vector multiresolution analysis* in the following sense:

$$s_0^{(i)} \subset \ldots \subset s_l^{(i)} \subset s_{l+1}^{(i)} \subset \ldots \subset l^2_{(i)}(\Omega), \qquad i = 1,2,3,$$
$$s_0 \subset \ldots \subset s_l \subset s_{l+1} \subset \ldots \subset l^2(\Omega) \ .$$

A space corresponding to some resolution contains all the information about the space at lower resolution. Furthermore, we have

$$\overline{\bigcup_{j=0}^{\infty} s_j^{(i)}}^{\|\cdot\|_{l^2(\Omega)}} = l^2_{(i)}(\Omega), \qquad i = 1,2,3, \tag{4.12}$$

and

$$\overline{\bigcup_{j=0}^{\infty} s_j}^{\|\cdot\|_{l^2(\Omega)}} = l^2(\Omega), \tag{4.13}$$

i.e., as the resolution increases the approximation of the vector field converges to the original vector field and as the resolution decreases the approximated vector field contains less and less information. Correspondingly, the uncertainty principle (cf. F.J. Narcowich, J.D. Ward (1996) W. Freeden (1999), S. Beth (2000)) tells us that the space localization of the scaling function increases with increasing scale parameter j, while the frequency (momentum) localization of the scaling function decreases with increasing j.

Given the structure of a multiresolution analysis, the representation of a vector field in $s_j^{(i)}$ is given by $p_j^{(i)}(f)$. The multiresolution analysis framework is certainly not unique. Several multiresolution frameworks can be constructed depending upon the choice of the triple $(\Phi_0)^\wedge(n) = \left((\Phi_0^{(1)})^\wedge(n), (\Phi_0^{(2)})^\wedge(n), (\Phi_0^{(3)})^\wedge(n)\right)^T$.

Finally it should be noted that

$$p_j^{(i)}(f) = \Phi_j^{(i)} \star \Phi_j^{(i)} * f = \mathbf{\Phi}_j^{(i)} * \mathbf{\Phi}_j^{(i)} * f \tag{4.14}$$

provided that

$$\Phi_j^{(i)}(\xi,\eta) = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (\Phi_j^{(i)})^\wedge(n) p_n^{(i)}(\xi,\eta), \quad (\xi,\eta) \in \Omega^2 \tag{4.15}$$

for every $j \in \mathbb{Z}$. Thus it is clear from our construction that the scale spaces s_l introduced above are exactly the same like in the case of vector wavelets (cf. W. Freeden et al. (1998)) based on tensorial kernel representations. This observation is a keystone for efficient multiresolution techniques in vectorial theory on the sphere. Theorem 4.1 can be rewritten as follows.

Corollary 4.2. *Suppose that f is of class $l^2(\Omega)$. Then*

$$\lim_{J\to\infty} \|\Phi_J^{(i)} \star \Phi_J^{(i)} * f - f^{(i)}\|_{l^2(\Omega)} = 0, \qquad i = 1,2,3, \tag{4.16}$$

and

$$\lim_{J\to\infty} \|\Phi_J \star \Phi_J * f - f\|_{l^2(\Omega)} = 0 \ . \tag{4.17}$$

5. Vector Wavelets. Next we introduce scale discrete wavelets. Since most of the material is known from the scalar case (confer the approach due to W. Freeden, M. Schreiner (1997)) our treatment will be brief.

Let $\mathbf{\Phi}_0$ as defined by (4.1) be a scale discrete scaling function generated by $(\Phi_0)^\wedge(n) = ((\Phi_0^{(1)})^\wedge(n), (\Phi_0^{(2)})^\wedge(n), (\Phi_0^{(3)})^\wedge(n))^T$. Let $(\Psi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$, $i = 1,2,3$, be three functions satisfying the following conditions:

(i) $(\Psi_0^{(i)})^\wedge(n)$ are piecewise continuous in $[0,\infty)$,

(ii) $\sum\limits_{n=0_i}^{\infty} \frac{2n+1}{4\pi} |(\Psi_0^{(i)})^\wedge(n)|^2 < \infty$,

(iii) $((\Psi_0^{(i)})^\wedge(x))^2 = \left((\Phi_0^{(i)})^\wedge(\frac{x}{2})\right)^2 - \left((\Phi_0^{(i)})^\wedge(x)\right)^2$, $x \in [0,\infty)$.

Then $(\Psi_0)^\wedge(n) = ((\Psi_0^{(1)})^\wedge(n), (\Psi_0^{(2)})^\wedge(n), (\Psi_0^{(3)})^\wedge(n))^T$ defines the so-called *scale discrete mother wavelet* $\mathbf{\Psi_0}$ which is given by

$$\mathbf{\Psi}_0(\xi,\eta) = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} (\Psi_0^{(i)})^\wedge(n) y_{n,l}^{(i)}(\xi) \otimes y_{n,l}^{(i)}(\eta), \qquad (\xi,\eta) \in \Omega^2 \ . \tag{5.1}$$

For $j \in \mathbb{Z}$ the dilation operator D_j can be applied in the same way as before: $D_j(\Psi_0)^\wedge(x) = (\Psi_j)^\wedge(x) = (\Psi_0)^\wedge(2^{-j}x)$ for $x \in [0,\infty)$. In other words,

$$D_j\mathbf{\Psi}_0(\xi,\eta) = \mathbf{\Psi}_j(\xi,\eta) = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{l=1}^{2n+1} (\Psi_j^{(i)})^\wedge(n) y_{n,l}^{(i)}(\xi) \otimes y_{n,l}^{(i)}(\eta), \qquad \xi,\eta \in \Omega \ .$$

The wavelets $\mathbf{\Psi}_j^{(i)}$ of type i are defined as in the case of the scaling function. Scale discrete wavelets at all scales j and positions η as introduced above can be displayed as rotated and dilated versions of the mother wavelets:

$$\mathbf{\Psi}_j(\xi,\eta) = \mathbf{\Psi}_{j;\xi}(\eta) = (R_\xi D_j \mathbf{\Psi}_0)(\cdot,\eta), \qquad (\xi,\eta) \in \Omega^2 \ . \tag{5.2}$$

R_ξ is the ξ-rotation operator defined by $R_\xi \mathbf{\Psi}_0(\cdot,\eta) = \mathbf{\Psi}_0(\xi,\eta), \xi \in \Omega$.

For a vector field $f \in l^2(\Omega)$ we now introduce *the spherical vector wavelet transform of type* i and *the spherical vector wavelet transform*, respectively, as follows:

$$(WT)_{\Psi_0^{(i)}}(f)(j;\eta) = (\mathbf{\Psi}_j^{(i)} * f)(\eta) = \int_\Omega \mathbf{\Psi}_j^{(i)}(\xi,\eta) f(\xi)\, d\omega(\xi) \ , \quad j \in \mathbb{Z},\ i = 1,2,3,$$

$$(WT)_{\Psi_0}(f)(j;\eta) = (\mathbf{\Psi}_j * f)(\eta) = \int_\Omega \mathbf{\Psi}_j(\xi,\eta) f(\xi)\, d\omega(\xi) \ , \quad j \in \mathbb{Z},$$

$\eta \in \Omega$. Theorem 5.1 now tells us that any field f can be reconstructed by its wavelet transform.

Theorem 5.1. *(Reconstruction Formula). Let* $\mathbf{\Phi}_0$ *be a scale discrete scaling function and let* $\mathbf{\Psi}_0$, *be the corresponding mother wavelet. Then,*

$$f^{(i)} = \mathbf{\Phi}_0^{(i)} * \mathbf{\Phi}_0^{(i)} * f + \sum_{l=0}^{\infty} \mathbf{\Psi}_l^{(i)} * \mathbf{\Psi}_l^{(i)} * f, \qquad i = 1,2,3, \tag{5.3}$$

and

$$f = \mathbf{\Phi}_0 * \mathbf{\Phi}_0 * f + \sum_{l=0}^{\infty} \mathbf{\Psi}_l * \mathbf{\Psi}_l * f \tag{5.4}$$

holds for all vector fields $f \in l^2(\Omega)$ in $\|\cdot\|_{l^2(\Omega)}$-sense.

Proof. It is not difficult to see that

$$\begin{aligned}
&\sum_{i=1}^{3}\sum_{n=0_i}^{\infty}\sum_{k=1}^{2n+1}((\Phi_0^{(i)})^\wedge(n))^2(f^{(i)})^\wedge(n,k)y_{n,k}^{(i)}\\
&\quad+\sum_{l=0}^{J-1}\sum_{i=1}^{3}\sum_{n=0_i}^{\infty}\sum_{k=1}^{2n+1}\left[\left((\Phi_0^{(i)})^\wedge(2^{-l-1}n)\right)^2-\left((\Phi_0^{(i)})^\wedge(2^{-l}n)\right)^2\right](f^{(i)})^\wedge(n,k)y_{n,k}^{(i)}\\
&=\sum_{i=1}^{3}\sum_{n=0_i}^{\infty}\sum_{k=1}^{2n+1}((\Phi_J^{(i)})^\wedge(n))^2(f^{(i)})^\wedge(n,k)y_{n,k}^{(i)}\\
&=\mathbf{\Phi}_J*\mathbf{\Phi}_J*f
\end{aligned}$$

holds for all positive integers J. Letting J tend to infinity we obtain the required result in connection with Theorem 4.1. □

As a particular result of the previous proof the following identity becomes obvious which illustrates the relation between Theorem 4.1 and Theorem 5.1:

$$\mathbf{\Phi}_J*\mathbf{\Phi}_J*f=\mathbf{\Phi}_0*\mathbf{\Phi}_0*f+\sum_{l=0}^{J-1}\mathbf{\Psi}_l*\mathbf{\Psi}_l*f. \tag{5.5}$$

The formula (5.5) can be reformulated in the following way

$$\mathbf{\Phi}_J*\mathbf{\Phi}_J*f+\sum_{l=J}^{M}\mathbf{\Psi}_l*\mathbf{\Psi}_l*f=\mathbf{\Phi}_{M+1}*\mathbf{\Phi}_{M+1}*f \tag{5.6}$$

($M > J$). The scaling function at scale $M+1$ can be replaced by a scaling function at a smaller scale J together with the wavelets at all "intermediate" scales. This gives rise to introduce operators $r_j^{(i)}, r_j$ defined by

$$r_j^{(i)}(f)=\mathbf{\Psi}_j^{(i)}*\mathbf{\Psi}_j^{(i)}*f^{(i)},\qquad j\in\mathbb{Z},\qquad i\in\{1,2,3\}, \tag{5.7}$$

and

$$r_j(f)=\mathbf{\Psi}_j*\mathbf{\Psi}_j*f,\qquad j\in\mathbb{Z}, \tag{5.8}$$

$f \in l^2(\Omega)$, respectively. The operators r_j can be interpreted as *bandpass filters.*

The *detail space* d_j is then the image of $l^2(\Omega)$ under the operator r_j:

$$d_j=\{r_j(f)\ |\ f\in l^2(\Omega)\}. \tag{5.9}$$

Analogously, we understand the spaces $d_j^{(i)}$:

$$d_j^{(i)}=\{r_j^{(i)}(f)\ |\ f\in l^2_{(i)}(\Omega)\}. \tag{5.10}$$

The operators $p_j, p_j^{(i)}, r_j, r_j^{(i)}$ on the one hand and the spaces $s_j, s_j^{(i)}, d_j, d_j^{(i)}$ on the other hand allow the following relations:

- $p_j^{(i)} = p_{j-1}^{(i)} + r_{j-1}^{(i)}, \qquad p_j = p_{j-1} + r_{j-1}$

- $p_j^{(i)} = p_0^{(i)} + \sum\limits_{l=0}^{j-1} r_l^{(i)}, \qquad i = 1,2,3, \qquad p_j = p_0 + \sum\limits_{l=0}^{j-1} r_l$
- $s_j^{(i)} = s_{j-1}^{(i)} + d_{j-1}^{(i)}, \qquad i = 1,2,3, \qquad s_j = s_{j-1} + d_{j-1},$
- $s_j^{(i)} = s_0^{(i)} + \sum\limits_{l=0}^{j-1} d_l^{(i)}, \qquad i = 1,2,3, \qquad s_j = s_0 + \sum\limits_{l=0}^{j-1} d_l.$

Remember that $p_j(f)$ is understood as a smoothing of f at scale j. In order to improve this smoothing we have to add more and more "details" $r_j(f)$, each of which may be understood as difference of two smoothings.

It should be mentioned that

$$(p_j^{(i)}(f))^\wedge(n,l) = (f^{(i)})^\wedge(n,l)(\Phi_j^{(i)})^\wedge(n)(\Phi_j^{(i)})^\wedge(n),$$
$$(r_j^{(i)}(f))^\wedge(n,l) = (f^{(i)})^\wedge(n,l)(\Psi_j^{(i)})^\wedge(n)(\Psi_j^{(i)})^\wedge(n).$$

These formulas give wavelet representation an interpretation in terms of Fourier analysis by explaining how the frequency spectrum of a vector field $f^{(i)} \in s_j^{(i)}$ is divided up between the spaces $s_{j-1}^{(i)}$ and $d_{j-1}^{(i)}$, which enhances our understanding of what is meant by "smoothing" and "detail".

From Theorem 3.5 it is clear that $r_j^{(i)}(f)$, $i = 1,2,3$, can be represented equivalently in bilinear 'scalar–vector' nomenclature by

$$r_j^{(i)}(f) = \Psi_j^{(i)} \star \Psi_j^{(i)} * f, \tag{5.11}$$

where

$$\Psi_j^{(i)}(\xi,\eta) = \sum_{n=0_i}^{\infty} \frac{2n+1}{4\pi} (\Psi_j^{(i)})^\wedge(n) p_n^{(i)}(\xi,\eta). \tag{5.12}$$

This finally leads us to the following corollary of Theorem 5.1

COROLLARY 5.2. *Le f be of class $l^2(\Omega)$. Then*

$$f^{(i)} = \Phi_0^{(i)} \star \Phi_0^{(i)} * f + \sum_{l=0}^{\infty} \Psi_l^{(i)} \star \Psi_l^{(i)} * f, \quad i \in \{1,2,3\}, \tag{5.13}$$

and

$$f = \Phi_0 \star \Phi_0 * f + \sum_{l=0}^{\infty} \Psi_l \star \Psi_l * f \ . \tag{5.14}$$

6. Examples. Two types of wavelets may be distinguished with respect to their frequency band, viz. non-bandlimited and bandlimited wavelets.

6.1. Non-bandlimited Wavelets. All non-bandlimited wavelets share the property that their generators $(\Psi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$, $i = 1,2,3$, have a global support, i.e. supp $(\Psi_0^{(i)})^\wedge(n) = [0,\infty)$. Since there are only a few conditions for a function $(\Phi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$ to be a generator of a scaling function, many choices are at our disposal.

As particularly important examples of globally supported generators $(\Phi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$, $i = 1,2,3$, we mention (cf. W. Freeden, U. Windheuser (1996, 1997)):

(i) *rational wavelets*: $(\Phi_0^{(i)})^\wedge(x) = (1-x)^{-s}$, $x \in [0,\infty)$, $s > 1$,

(ii) *exponential wavelets*: $(\Phi_0^{(i)})^\wedge(x) = e^{-h(x)}$, $x \in [0,\infty)$, where $h : [0,\infty) \to \mathbb{R}$ is assumed to satisfy the conditions:
- $h \in C^{(\infty)}[0,\infty)$
- $h(0) = 0$, $h(x) > 0$ for $x > 0$
- $h(x) < h(x')$ whenever $0 < x < x'$,
- $\sum_{n=0}^{\infty} \frac{2n+1}{4\pi} e^{-h(n)} < \infty$.

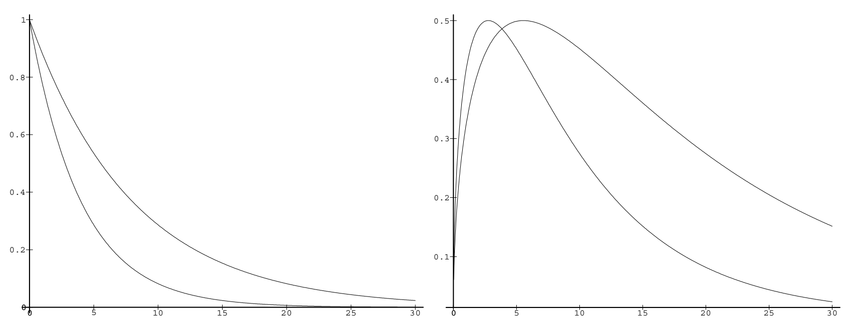

Fig. 6.1. *Generator of Abel-Poisson scaling function (left) and wavelet (right) for* $j = 2,3$

A proper choice is, for example, the *Abel–Poisson wavelet* with $h(x) = x$, $x \in [0,\infty)$ (see Figure 6.1 for the corresponding generators) or the *Gauss–Weierstrass wavelet* with $h(x) = x(x+1)$, $x \in [0,\infty)$.

In all these cases the detail spaces generally are not of finite dimension. Nevertheless, as we have seen before, a multiresolution analysis is still valid.

6.2. Bandlimited Wavelets. If all $(\Phi_0^{(i)})^\wedge(n) : [0,\infty) \to \mathbb{R}$, $i = 1,2,3$, are chosen to be compactly supported, then the scale spaces and consequently the detail spaces have finite dimensions. Moreover, for the purpose of reconstruction, the wavelet transform $(WT)_{\Psi_0^{(i)}}(f)(j;\cdot)$ has to be known for exact evaluation only on a finite set of points on the sphere Ω. This should be discussed in more detail.

We start with the simplest choice for a compactly supported generator.

6.2.1. Shannon Wavelet. We choose the scaling function as follows:

$$(\Phi_0^{(i)})^\wedge(x) = \begin{cases} 1, & x \in [0,1) \\ 0, & \text{else} \end{cases}$$

$i = 1, 2, 3$. The dilates and the wavelet generators are immediately at hand by

$$(\Phi_j^{(i)})^\wedge(x) = \begin{cases} 1, & x \in [0, 2^j) \\ 0, & \text{else}, \end{cases}$$

$$(\Psi_j^{(i)})^\wedge(x) = \begin{cases} 1, & x \in [2^j, 2^{j+1}) \\ 0, & \text{else}. \end{cases}$$

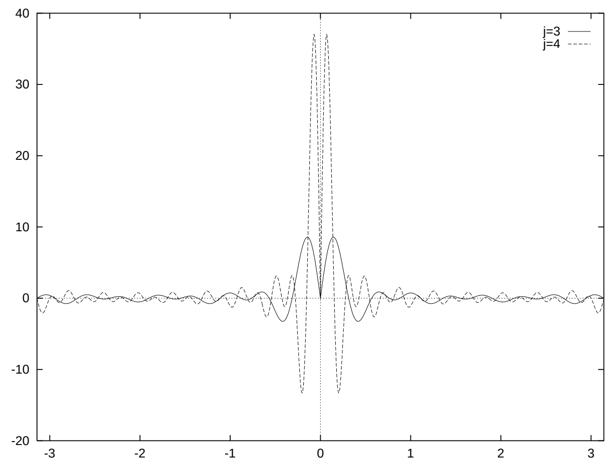

FIG. 6.2. *Shannon wavelet of type i=2 for j=3,4*

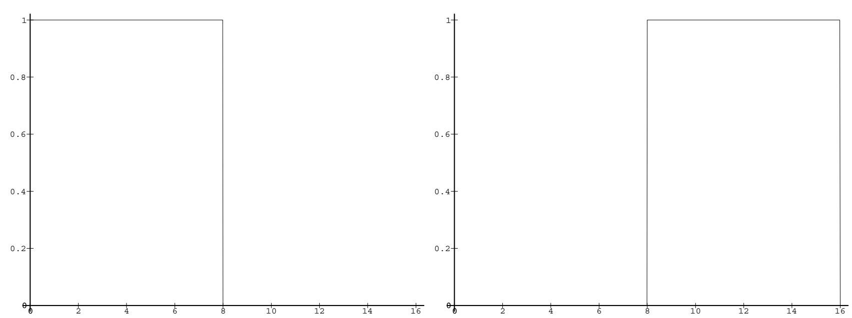

FIG. 6.3. *Generator of Shannon scaling function (left) and wavelet (right) for* $j = 2, 3$

The reconstruction via the corresponding wavelets $\mathbf{\Psi}_j$ yields

$$\mathbf{\Psi}_j * \mathbf{\Psi}_j * f = \sum_{i=1}^{3} \sum_{n=2^j}^{2^{j+1}-1} \sum_{l=1}^{2n+1} (f^{(i)})^\wedge(n, l) y_{n,l}^{(i)},$$

i.e. we get a "piece" out of the Fourier expansion of f in terms of vector spherical harmonics. Within this "piece", all Fourier coefficients are fully reconstructed and are not affected by weights. Note that the vectorial kernels can be expressed as a vectorial part times a scalar sum. What we have plotted in Figure 6.2 is the absolute value of the vectorial part times the scalar sum, resulting in the 'dip' that can be observed. The dip is due to the fact, that for type 2 and 3 vectorial wavelets the vectorial part results in the zero vector when the argument ξ tends towards η.

6.2.2. CP-Wavelet. The Shannon wavelet shows serious oscillations (cf. Figure 6.2). This phenomenon is usually not very helpful for computational purposes. If one is willing to give up the orthogonality of the detail spaces d_j, the number of oscillations can be reduced drastically by a modification of the generator $(\Phi_0)^\wedge(n)$. A good choice (as we know from numerical experiences) is the c(ubic) p(olynomial)-wavelet, briefly called CP-wavelet.

$$(\Phi_0^{(i)})^\wedge(x) = \begin{cases} (1-x)^2(1+2x) & , \quad x \in [0,1) \\ 0 & , \quad x \in [1,\infty), \end{cases}$$

$i = 1,2,3$, which is constructed in such a way that

$$(\Phi_0^{(i)})^\wedge(1) = ((\Phi_0^{(i)})^\wedge(x))'(1) \ = ((\Phi_0^{(i)})^\wedge(x))'(0) \ = 0.$$

For the CP-wavelet and its generator see Figures 6.4, 6.5. Of course, a great number of other choices of $(\Phi_0^{(i)})^\wedge(n)$, $i = 1,2,3$, can be used (cf. W. Freeden, M. Schreiner (1997), S. Beth (2000)). We omit these considerations.

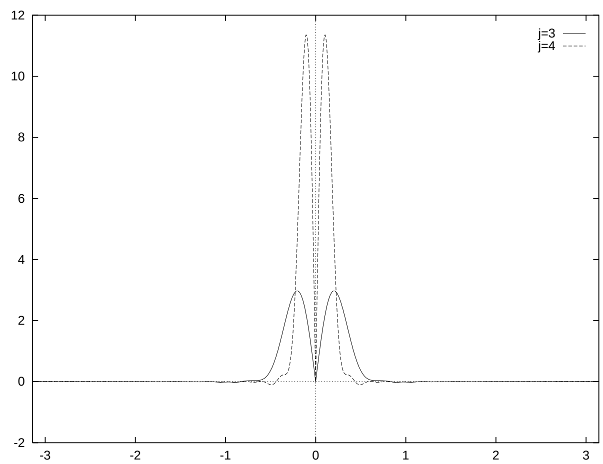

Fig. 6.4. *CP wavelet of type i=3 for j=3,4*

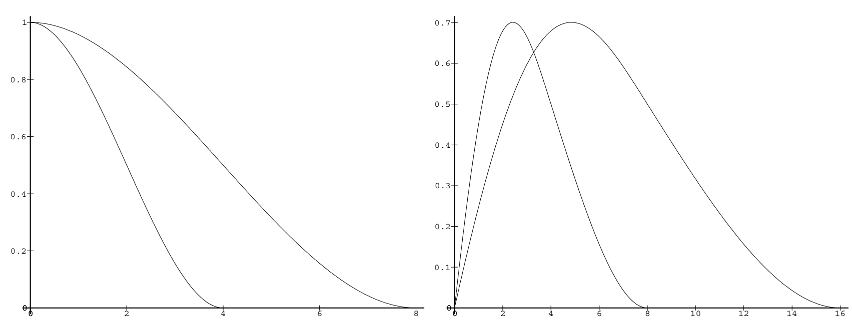

Fig. 6.5. *Generator of CP scaling function (left) and wavelet (right) for $j = 2, 3$*

7. Spectral Approximation. Let us think of an 'output signal' g as produced by a linear operator $\mathbf{\Lambda}$ applied to an 'input signal' f

$$\mathbf{\Lambda} g = f,$$

where $\mathbf{\Lambda}$ is an operator mapping $l^2(\Omega)$ into itself such that

$$\mathbf{\Lambda} y_{n,k}^{(i)} = \left(\mathbf{\Lambda}^{(i)}\right)^{\wedge}(n,k)\; y_{n,k}^{(i)}$$

for $i = 1,2,3$; $n = 0_i, 0_i+1, \ldots$; $k = 1, \ldots, 2n+1$ (with the so-called symbol $\left\{\left(\mathbf{\Lambda}^{(i)}\right)^{\wedge}(n,k)\right\}_{n\in\mathbb{N}_0}$, $i = 1,2,3$, being sequences of real numbers).

7.1. Signal–to–Noise Response. In practise, an error-affected 'output signal' is observed

$$\tilde{g} = g + \tilde{\varepsilon},$$

where $\tilde{\varepsilon}$ is the *observation noise.* In analogy to the scalar case (cf. W. Freeden et al. (2000)) and in accordance with the approach used in physical geodesy (see e.g. R. Rummel (1997)) we assume that

$$Cov[\tilde{g}(\xi), \tilde{g}(\eta)] = E[\tilde{\varepsilon}(\xi), \tilde{\varepsilon}(\eta)] = \mathbf{k}(\xi,\eta), \quad (\xi,\eta) \in \Omega \times \Omega,$$

is known, where $\mathbf{k}(\cdot,\cdot) : \Omega \times \Omega \to \mathbb{R}^{3\times 3}$ is explicitly given by the product kernel

$$\begin{aligned}
\mathbf{k}(\xi,\eta) &= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^{\wedge}(n,k)(\mu_n^{(i)})^{-1} o_\xi^{(i)} o_\eta^{(i)} Y_{n,k}(\xi) Y_{n,k}(\eta) \\
&= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^{\wedge}(n,k) y_{n,k}^{(i)}(\xi) \otimes y_{n,k}^{(i)}(\eta), \quad (\xi,\eta) \in \Omega^2,
\end{aligned}$$

and the symbol $\left\{(\mathbf{k}^{(i)})^{\wedge}(n,k)\right\}_{n\in\mathbb{N}}$, $i = 1,2,3$, satisfying the conditions:

$$\text{(C1)} \quad (\mathbf{k}^{(i)})^{\wedge}(n,k) \geq 0 \;\text{ for } n = 0_i, 0_i+1, \ldots,\; i \in \{1,2,3\},$$

$$\text{(C2)} \quad \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\mathbf{k}^{(i)}\right)^{\wedge}(n,k) \sup_{\eta\in\Omega} \left(y_{n,k}^{(i)}(\eta)\right)^2 < \infty \;.$$

7.1.1. Degree Variances. Any 'output function' (output signal) can be expanded into an orthogonal series in terms of vector spherical harmonics:

$$\begin{aligned}
\tilde{g} \;=\; \widetilde{\mathbf{\Lambda} f} &= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\mathbf{\Lambda}^{(i)}\right)^{\wedge}(n,k)(\tilde{f}^{(i)})^{\wedge}(n,k) y_{n,k}^{(i)} \\
&= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\tilde{g}^{(i)})^{\wedge}(n,k) y_{n,k}^{(i)},
\end{aligned}$$

where the equality has to be understood in the sense of $\|\cdot\|_{l^2(\Omega)}$.

The *signal degree and order variances of type i* are defined by

$$
\begin{aligned}
Var_{n,k}^{(i)}(\widetilde{\mathbf{\Lambda} f}) &= \left(\left((\widetilde{\mathbf{\Lambda} f})^{(i)}\right)^\wedge (n,k)\right)^2 \\
&= \int_\Omega \int_\Omega (\widetilde{\mathbf{\Lambda} f})(\xi) \cdot \left(y_{n,k}^{(i)}(\xi) \otimes y_{n,k}^{(i)}(\eta)\right) (\widetilde{\mathbf{\Lambda} f})(\eta) \; d\omega(\xi) \; d\omega(\eta) \\
&= \int_\Omega \int_\Omega \left(y_{n,k}^{(i)}(\xi) \otimes y_{n,k}^{(i)}(\eta)\right) \cdot \left((\widetilde{\mathbf{\Lambda} f})(\xi) \otimes (\widetilde{\mathbf{\Lambda} f})(\eta)\right) \; d\omega(\xi) \; d\omega(\eta).
\end{aligned}
$$

Accordingly the *signal degree variances of type i* are given by

$$
\begin{aligned}
Var_n^{(i)}(\widetilde{\mathbf{\Lambda} f}) &= \sum_{k=1}^{2n+1} Var_{n,k}^{(i)}(\widetilde{\mathbf{\Lambda} f}) \\
&= \sum_{k=1}^{2n+1} \left(\left((\widetilde{\mathbf{\Lambda} f})^{(i)}\right)^\wedge (n,k)\right)^2 \\
&= \frac{2n+1}{4\pi} \int_\Omega \int_\Omega (\widetilde{\mathbf{\Lambda} f})(\xi) \cdot \mathbf{p}_n^{(i,i)}(\xi,\eta) (\widetilde{\mathbf{\Lambda} f})(\eta) \; d\omega(\eta) \; d\omega(\xi),
\end{aligned}
$$

while the *signal degree variances* read as follows:

$$
Var_n(\widetilde{\mathbf{\Lambda} f}) = \sum_{i=1}^{3} Var_n^{(i)}(\widetilde{\mathbf{\Lambda} f}).
$$

Obviously, by virtue of Parseval's identity, we obtain

$$
\left\| \widetilde{\mathbf{\Lambda} f} \right\|_{l^2(\Omega)} = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} Var_{n,k}^{(i)}(\widetilde{\mathbf{\Lambda} f}).
$$

7.1.2. Degree Error Covariances. In analogy to the scalar case, the spectral error theory is based on the *degree and order error covariance of type i* given by

$$
\begin{aligned}
& Cov_{n,k}^{(i)}(\mathbf{k}) \\
&= \int_\Omega \int_\Omega \left(y_{n,k}^{(i)}(\xi) \otimes y_{n,k}^{(i)}(\eta)\right) \cdot \mathbf{k}(\xi,\eta) \; d\omega(\xi) \; d\omega(\eta) \\
&= \sum_{l=1}^{3} \sum_{p=0_i}^{\infty} \sum_{q=1}^{2p+1} (\mathbf{k}^{(l)})^\wedge(p,q) \int_\Omega \int_\Omega \left(y_{n,k}^{(i)}(\eta) \cdot y_{p,q}^{(l)}(\eta)\right) \left(y_{n,k}^{(i)}(\xi) \cdot y_{p,q}^{(l)}(\xi)\right) \; d\omega(\xi) \; d\omega(\eta) \\
&= (\mathbf{k}^{(i)})^\wedge(n,k).
\end{aligned}
$$

Moreover, we have

$$
Cov_n^{(i)}(\mathbf{k}) = \sum_{k=1}^{2n+1} Cov_{n,k}^{(i)}(\mathbf{k}) = \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^\wedge(n,k) \tag{7.1}
$$

and

$$
Cov_n(\mathbf{k}) = \sum_{i=1}^{3} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^\wedge(n,k)
$$

for $n = 0_i, 0_i + 1, \ldots$.

7.1.3. Spectral Estimation. The signal-to-noise relation is determined by the *degree and order resolution set* $\mathcal{N}_{res}^{(i)}$ of type i:

1. *signal dominates noise*

$$Var_{n,k}^{(i)}(\widetilde{\mathbf{\Lambda} f}) \geq Cov_{n,k}^{(i)}(\mathbf{k}), \quad (n,k) \in \mathcal{N}_{res}^{(i)} \ .$$

2. *noise dominates signal*

$$Var_{n,k}^{(i)}(\widetilde{\mathbf{\Lambda} f}) < Cov_{n,k}^{(i)}(\mathbf{k}), \quad (n,k) \notin \mathcal{N}_{res}^{(i)} \ ,$$

for $n = 0_i, 1, 2 \ldots, \ \ k = 1 \ldots 2n+1$.

7.2. Tensor (Radial Basis) Multiscale Approximation. Let $\mathbf{\Phi}_0$ be a scale discrete scaling functions, and let $\mathbf{\Psi}_0$ be the corresponding mother wavelet.

Under this assumption it is easily seen that $\tilde{g}$ can be represented in vectorial multiscale approximation as follows

$$\tilde{g} = \widetilde{\mathbf{\Lambda} f} = \sum_{j=-\infty}^{\infty} \mathbf{\Psi}_j * \mathbf{\Psi}_j * (\widetilde{\mathbf{\Lambda} f}) \ , \tag{7.2}$$

where this equality is understood in the $\|\cdot\|_{l^2(\Omega)}$-sense. This result is equivalent to the fact that

$$\lim_{N\to\infty} \left\| \widetilde{\mathbf{\Lambda} f} - \left((\widetilde{\mathbf{\Lambda} f})_{J_0} + \sum_{j=J_0}^{N} \mathbf{\Psi}_j * \mathbf{\Psi}_j * (\widetilde{\mathbf{\Lambda} f}) \right) \right\|_{l^2(\Omega)} = 0 \tag{7.3}$$

for every $J_0 \in \mathbb{Z}$.

7.2.1. Tensor Based Scale and Position Variances. Denote by $l^{(2)}(\mathbb{Z} \times \Omega)$ the space of fields $h : \mathbb{Z} \times \Omega \to \mathbb{R}^3$ satisfying the inequality

$$\sum_{j=-\infty}^{\infty} \int_{\Omega} \big(h(j;\eta) \cdot h(j;\eta) \big) \ d\omega(\eta) < \infty. \tag{7.4}$$

The space $l^{(2)}(\mathbb{Z} \times \Omega)$ is a Hilbert space equipped with the inner product

$$(h_1, h_2)_{l^{(2)}(\mathbb{Z}\times\Omega)} = \sum_{j=-\infty}^{\infty} \int_{\Omega} \big(h_1(j;\eta) \cdot h_2(j;\eta) \big) \ d\omega(\eta) \tag{7.5}$$

corresponding to the norm

$$\|h\|_{l^{(2)}(\mathbb{Z}\times\Omega)} = \left(\sum_{j=-\infty}^{\infty} \int_{\Omega} \big| h(j;\eta) \big|^2 d\omega(\eta) \right)^{1/2} . \tag{7.6}$$

The multiscale approximation of an 'output field' $\tilde{g} = \widetilde{\mathbf{\Lambda} f} \in l^2(\Omega)$ immediately leads us to

$$\left(\widetilde{\mathbf{\Lambda} f}, \widetilde{\mathbf{\Lambda} f} \right)_{l^2(\Omega)}$$

$$= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\left((\widetilde{\mathbf{\Lambda} f})^{(i)} \right)^{\wedge} (n,k) \right)^2 \sum_{j=-\infty}^{\infty} \left(\left(\Psi_j^{(i)} \right)^{\wedge} (n) \right)^2$$

$$= \sum_{i=1}^{3} \sum_{j=-\infty}^{\infty} \int_{\Omega} \left| \int_{\Omega} \mathbf{\Psi}_j^{(i)}(\eta,\xi) \, (\widetilde{\mathbf{\Lambda} f})(\xi) d\omega(\xi) \right|^2 \, d\omega(\eta)$$

$$= \sum_{i=1}^{3} \sum_{j=-\infty}^{\infty} \int_{\Omega} \left(\int_{\Omega} \int_{\Omega} \left(\mathbf{\Psi}_j^{(i)}(\xi,\eta) \mathbf{\Psi}_j^{(i)}(\eta,\zeta) \right) \cdot \right.$$

$$\left. \cdot \left((\widetilde{\mathbf{\Lambda} f})(\xi) \otimes (\widetilde{\mathbf{\Lambda} f})(\zeta) \right) \, d\omega(\xi) \, d\omega(\zeta) \right) d\omega(\eta)$$

$$= \sum_{i=1}^{3} \sum_{j=-\infty}^{\infty} \int_{\Omega} (\widetilde{\mathbf{\Lambda} f})(\xi) \cdot \left(\mathbf{\Psi}_j^{(i)} * \mathbf{\Psi}_j^{(i)} * \widetilde{\mathbf{\Lambda} f} \right)(\xi) \, d\omega(\xi)$$

$$= \sum_{i=1}^{3} \sum_{j=-\infty}^{+\infty} \int_{\Omega} \widetilde{\mathbf{\Lambda} f}(\xi) \cdot r_j^{(i)}(f)(\xi) \, d\omega(\xi) \ .$$

Consequently, the *tensor based signal scale and space variance at position* $\eta \in \Omega$, *scale* $j \in \mathbb{Z}$ *and type* $i \in \{1,2,3\}$ is defined by

$$TVar_{j;\eta}^{(i)} \left(\widetilde{\mathbf{\Lambda} f} \right) = \left| \int_{\Omega} \mathbf{\Psi}_j^{(i)}(\eta,\xi) \left(\widetilde{\mathbf{\Lambda} f} \right)(\xi) \, d\omega(\xi) \right|^2$$

$$= \int_{\Omega} \mathbf{\Psi}_j^{(i)}(\eta,\xi) \left(\widetilde{\mathbf{\Lambda} f} \right)(\xi) \, d\omega(\xi) \cdot \int_{\Omega} \mathbf{\Psi}_j^{(i)}(\eta,\zeta) \left(\widetilde{\mathbf{\Lambda} f} \right)(\zeta) \, d\omega(\zeta)$$

$$= \int_{\Omega} \int_{\Omega} \mathbf{\Psi}_j^{(i)}(\eta,\xi) \left(\widetilde{\mathbf{\Lambda} f} \right)(\xi) \cdot \mathbf{\Psi}_j^{(i)}(\eta,\zeta) \left(\widetilde{\mathbf{\Lambda} f} \right)(\zeta) \, d\omega(\xi) \, d\omega(\zeta)$$

$$= \int_{\Omega} \int_{\Omega} \left(\mathbf{\Psi}_j^{(i)}(\xi,\eta) \mathbf{\Psi}_j^{(i)}(\eta,\zeta) \right) \cdot \left(\left(\widetilde{\mathbf{\Lambda} f} \right)(\xi) \otimes \left(\widetilde{\mathbf{\Lambda} f} \right)(\zeta) \right) \, d\omega(\xi) \, d\omega(\zeta),$$

while

$$TVar_{j;\eta}(\widetilde{\mathbf{\Lambda} f}) = \sum_{i=1}^{3} TVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda} f}).$$

The *tensor based signal scale variance of type* i is defined by

$$TVar_j^{(i)}(\widetilde{\mathbf{\Lambda} f}) = \int_{\Omega} TVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda} f}) \, d\omega(\eta),$$

and the *tensor based signal scale variance* is given by

$$TVar_j(\widetilde{\mathbf{\Lambda} f}) = \sum_{i=1}^{3} TVar_j^{(i)}(\widetilde{\mathbf{\Lambda} f}).$$

In conclusion, the identity

$$\left\| \widetilde{\mathbf{\Lambda} f} \right\|_{l^2(\Omega)}^2 = \sum_{j=-\infty}^{\infty} TVar_j(\widetilde{\mathbf{\Lambda} f})$$

$$= \sum_{j=-\infty}^{\infty} \int_{\Omega} TVar_{j;\eta}(\widetilde{\mathbf{\Lambda} f})\ d\omega(\eta)$$
$$= \left\| TVar_{\cdot;\cdot}(\widetilde{\mathbf{\Lambda} f}) \right\|^2_{l^2(\mathbb{Z}\times\Omega)}$$

holds true. In spectral language we have

$$TVar^{(i)}_{j,\eta}(\widetilde{\mathbf{\Lambda} f}) = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\left(\Psi^{(i)}_j \right)^{\wedge} (n) \right)^2 \left(\left((\widetilde{\mathbf{\Lambda} f})^{(i)} \right)^{\wedge} (n,k) \right)^2 \left(y^{(i)}_{n,k}(\eta) \right)^2 \tag{7.7}$$

and

$$TVar^{(i)}_{j}(\widetilde{\mathbf{\Lambda} f}) = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\left(\Psi^{(i)}_j \right)^{\wedge} (n,k) \right)^2 \left(\left(\left(\widetilde{\mathbf{\Lambda} f} \right)^{(i)} \right)^{\wedge} (n,k) \right)^2 . \tag{7.8}$$

7.2.2. Noise Model. Let $\mathbf{k} : (\xi,\eta) \mapsto \mathbf{k}(\xi,\eta), \quad (\xi,\eta) \in \Omega \times \Omega$, satisfy the conditions (C1) and (C2). The (tensor) multiscale error theory is based on the *tensor based scale and space error covariance at position $\eta \in \Omega$, scale j, and type i :*

$$TCov^{(i)}_{j;\eta}(\mathbf{k}) = \int_{\Omega} \int_{\Omega} \left(\mathbf{\Psi}^{(i)}_j(\xi,\eta) \mathbf{\Psi}^{(i)}_j(\eta,\zeta) \right) \cdot \mathbf{k}(\xi,\zeta)\ d\omega(\xi)\ d\omega(\zeta). \tag{7.9}$$

Furthermore,

$$TCov_{j;\eta}(\mathbf{k}) = \sum_{i=1}^{3} TCov^{(i)}_{j;\eta}(\mathbf{k}). \tag{7.10}$$

The *tensor based scale error covariance of type i* is defined by

$$TCov^{(i)}_{j}(\mathbf{k}) = \int_{\Omega} TCov^{(i)}_{j;\eta}(\mathbf{k})\ d\omega(\eta), \tag{7.11}$$

whereas

$$TCov_{j}(\mathbf{k}) = \sum_{i=1}^{3} TCov^{(i)}_{j}(\mathbf{k}). \tag{7.12}$$

Obviously, we have

$$TCov^{(i)}_{j;\eta}(\mathbf{k}) = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^{\wedge}(n,k) \left((\Psi^{(i)}_j)^{\wedge}(n) \right)^2 \left(y^{(i)}_{n,k}(\eta) \right)^2 , \tag{7.13}$$

$$TCov^{(i)}_{j}(\mathbf{k}) = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^{\wedge}(n,k) \left((\Psi^{(i)}_j)^{\wedge}(n) \right)^2 \tag{7.14}$$

and

$$TCov_{j}(\mathbf{k}) = \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^{\wedge}(n,k) \left(\left(\Psi^{(i)}_j \right)^{\wedge} (n) \right)^2 . \tag{7.15}$$

Note that the multiscale noise model is able to specify pointwise dependent error covariances (which is not possible in spectral theory by means of vector spherical harmonics).

7.2.3. Tensor Based Scale and Space Estimation. Signal and noise scale 'intersect' at the so-called *tensor based scale and space resolution set* $\mathcal{TZ}_{res}^{(i)}(\eta)$ *of type* i *at position* η. We distinguish the following cases:

1. *signal dominates noise*

$$TVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda}f}) \geq TCov_{j;\eta}^{(i)}(\mathbf{k}), \quad (j;\eta) \in \mathcal{TZ}_{res}^{(i)}(\eta).$$

2. *noise dominates signal*

$$TVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda}f}) < TCov_{j;\eta}^{(i)}(\mathbf{k}), \quad (j;\eta) \notin \mathcal{TZ}_{res}^{(i)}(\eta).$$

7.3. Vector (Radial Basis) Multiscale Approximation. Let Φ_0 again be a vectorial scale discrete scaling function, and let Ψ_0 be the corresponding vectorial mother wavelet. Then we know from our multiscale approach that

$$\tilde{g} = \widetilde{\mathbf{\Lambda}f} = \sum_{i=1}^{3} \sum_{j=-\infty}^{+\infty} \Psi_j^{(i)} \star \Psi_j^{(i)} * (\widetilde{\mathbf{\Lambda}f}) \tag{7.16}$$

(in the $\|\cdot\|_{l^2(\Omega)}$–sense).

7.3.1. Vector Based Signal Scale and Space Variances. The multiscale approximation of an 'output field' $\tilde{g} = \widetilde{\mathbf{\Lambda}f} \in l^2(\Omega)$ now leads us to

$$\begin{aligned}
&\left(\widetilde{\mathbf{\Lambda}f}, \widetilde{\mathbf{\Lambda}f}\right)_{l^2(\Omega)} \\
&= \sum_{i=1}^{3} \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\left(\left(\widetilde{\mathbf{\Lambda}f}\right)^{(i)}\right)^{\wedge}(n,k)\right)^2 \sum_{j=-\infty}^{+\infty} \left(\left(\Psi_j^{(i)}\right)^{\wedge}(n)\right)^4 \\
&= \sum_{i=1}^{3} \sum_{j=-\infty}^{+\infty} \int_\Omega \left| \int_\Omega (\Psi_j^{(i)})^{(2)}(\xi,\eta) \cdot (\widetilde{\mathbf{\Lambda}f})(\xi) \; d\omega(\xi) \right|^2 \; d\omega(\eta) \; .
\end{aligned}$$

This gives rise to introduce the *vector based signal scale and space variance at position* $\eta \in \Omega$, *scale* $j \in \mathbb{Z}$ *and type* $i \in \{1,2,3\}$ by

$$\begin{aligned}
VVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda}f}) &= \left| \int_\Omega \Psi_j^{(i)}(\xi,\eta) \cdot (\widetilde{\mathbf{\Lambda}f})(\xi) \; d\omega(\xi) \right|^2 \\
&= \int_\Omega \int_\Omega \left(\Psi_j^{(i)}(\xi,\eta) \cdot (\widetilde{\mathbf{\Lambda}f})(\xi)\right) \left(\Psi_j^{(i)}(\zeta,\eta) \cdot (\widetilde{\mathbf{\Lambda}f})(\zeta)\right) \; d\omega(\xi) \; d\omega(\zeta) \\
&= \int_\Omega \int_\Omega \left(\Psi_j^{(i)}(\xi,\eta) \otimes \Psi_j^{(i)}(\zeta,\eta)\right) \cdot \left((\widetilde{\mathbf{\Lambda}f})(\xi) \otimes (\widetilde{\mathbf{\Lambda}f})(\zeta)\right) \; d\omega(\xi) \; d\omega(\zeta) \; .
\end{aligned}$$

The *vector based signal scale variance of type* i is given by

$$VVar_j^{(i)}(\widetilde{\mathbf{\Lambda}f}) = \int_\Omega VVar_{j,\eta}^{(i)}(\widetilde{\mathbf{\Lambda}f}) d\omega(\eta), \tag{7.17}$$

while the *vector based signal scale variance* reads as follows

$$VVar_j(\widetilde{\mathbf{\Lambda}f}) = \sum_{i=1}^{3} VVar_j^{(i)}(\widetilde{\mathbf{\Lambda}f}) \; . \tag{7.18}$$

Obviously, $VVar^{(i)}_{j;\eta}(\widetilde{\mathbf{\Lambda} f})$ can be represented in terms of scalar spherical harmonics as follows:

$$VVar^{(i)}_{j;\eta}(\widetilde{\mathbf{\Lambda} f}) = \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} \left(\left(\Psi^{(i)}_j \right)^\wedge (n) \right) \left(\left(\widetilde{\mathbf{\Lambda} f} \right)^{(i)} \right)^\wedge (n,k) \big(Y_{n,k}(\eta) \big)^2 \ . \tag{7.19}$$

Hence it follows that

$$TVar^{(i)}_j(\widetilde{\mathbf{\Lambda} f}) = VVar^{(i)}_j(\widetilde{\mathbf{\Lambda} f}), \qquad i = 1,2,3, \tag{7.20}$$

and

$$TVar_j(\widetilde{\mathbf{\Lambda} f}) = VVar_j(\widetilde{\mathbf{\Lambda} f}) \ . \tag{7.21}$$

In conclusion,

$$\|\widetilde{\mathbf{\Lambda} f}^{(i)}\|_{l^2(\Omega)} = \|TVar^{(i)}.;.(\widetilde{\mathbf{\Lambda} f})\|_{l^2(\mathbb{Z}\times\Omega)} = \|VVar^{(i)}.;.(\widetilde{\mathbf{\Lambda} f})\|_{l^2(\mathbb{Z}\times\Omega)} \ . \tag{7.22}$$

7.3.2. Noise Mode. Let $\mathbf{k} : (\xi,\eta) \mapsto \mathbf{k}(\xi,\eta)$, $(\xi,\eta) \in \Omega\times\Omega$, satisfy the condition (C1) and (C2). The (vector) multiscale error theory is based on the *vector based scale and space error covariance at position* $\eta \in \Omega$, *scale* j, *and type* i*:*

$$VCor^{(i)}_{j;\eta}(\mathbf{k}) = \int_\Omega \int_\Omega \left(\Psi^{(i)}_j(\xi,\eta) \otimes \Psi^{(i)}_j(\zeta,\eta) \right) \cdot \mathbf{k}(\xi,\zeta) \ d\omega(\xi) \ d\omega(\zeta) \ . \tag{7.23}$$

Furthermore,

$$VCov_{j;\eta}(\mathbf{k}) = \sum_{i=1}^{3} VCov^{(i)}_{j;\eta}(\mathbf{k}) \ . \tag{7.24}$$

The *vector based scale error covariance of type* i is defined by

$$VCov^{(i)}_j(\mathbf{k}) = \int_\Omega VCov^{(i)}_{j;\eta}(\mathbf{k}) \ d\omega(\eta) \ . \tag{7.25}$$

It is not hard to verify that

$$\begin{aligned}
&VCov^{(i)}_{j;\eta}(\mathbf{k}) \\
&= \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^\wedge(n) \int_\Omega \int_\Omega \left(\Psi^{(i)}_j(\xi,\eta) \otimes \Psi^{(i)}_j(\zeta,\eta) \right) \cdot \left(y^{(i)}_{n,k}(\xi) \otimes y^{(i)}_{n,k}(\zeta) \right) \ d\omega(\xi) \ d\omega(\zeta) \\
&= \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^\wedge(n) \int_\Omega \Psi^{(i)}_j(\xi,\eta) \cdot y^{(i)}_{n,k}(\xi) \ d\omega(\xi) \int_\Omega \Psi^{(i)}_j(\zeta,\eta) \cdot y^{(i)}_{n,k}(\zeta) \ d\omega(\zeta) \\
&= \sum_{n=0_i}^{\infty} \sum_{k=1}^{2n+1} (\mathbf{k}^{(i)})^\wedge(n) \left(\left(\Psi^{(i)}_j \right)^\wedge (n) \right)^2 (Y_{n,k}(\eta))^2 \ .
\end{aligned}$$

This shows us that

$$TCov^{(i)}_j(\mathbf{k}) = VCov^{(i)}_j(\mathbf{k}), \qquad i = 1,2,3, \tag{7.26}$$

and

$$TCov_j(\mathbf{k}) = VCov_j(\mathbf{k}) \ . \tag{7.27}$$

7.3.3. Vector Scale and Space Estimation. Signal and noise scale 'intersect' at the so–called *vector based scale and resolution* set $\mathcal{VZ}_{\mathrm{res}}^{(i)}(\eta)$ *of type* i *at position* η. We distinguish the following case:

1. *signal dominates noise*

$$VVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda} f}) \geq VCov_{j;\eta}^{(i)}(\mathbf{k}), \qquad (j;\eta) \in \mathcal{VZ}_{\mathrm{res}}^{(i)}(\eta) \ . \tag{7.28}$$

2. *noise dominates signal*

$$VVar_{j;\eta}^{(i)}(\widetilde{\mathbf{\Lambda} f}) < VCov_{j;\eta}^{(i)}(\mathbf{k}), \qquad (j;\eta) \in \mathcal{VZ}_{\mathrm{res}}^{(i)}(\eta) \ . \tag{7.29}$$

8. Selective Multiscale Reconstruction. The vector variant of the multiscale approximation can be formulated by replacing the (unknown) error–free field $(\mathbf{\Lambda} f)^{(i)}, i \in \{1,2,3\}$, being approximated by

$$(\mathbf{\Lambda} f)^{(i)}_J = \Phi^{(i)}_{J_0} \star \Phi^{(i)}_{J_0} * (\mathbf{\Lambda} f) + \sum_{j=J_0}^{J-1} \Psi^{(i)}_j \star \Psi^{(i)}_j * (\mathbf{\Lambda} f) \tag{8.1}$$

by the error–affected field $(\widetilde{\mathbf{\Lambda} f})^{(i)}$, $i \in \{1,2,3\}$, such as

$$(\widetilde{\mathbf{\Lambda} f})^{(i)}_J = \Phi^{(i)}_{J_0} \star \Phi^{(i)}_{J_0} * (\widetilde{\mathbf{\Lambda} f}) + \sum_{j=J_0}^{J-1} \Psi^{(i)}_j \star \Psi^{(i)}_j * (\widetilde{\mathbf{\Lambda} f}), \tag{8.2}$$

$J > J_0$. The coefficients at position $\eta \in \Omega$

$$\tilde{p}^{(i)}_{J_0}(\widetilde{\mathbf{\Lambda} f})(\eta) = \int_\Omega \Phi^{(i)}_{J_0}(\eta,\zeta) \int_\Omega \Phi^{(i)}_{J_0}(\xi,\zeta) \cdot (\widetilde{\mathbf{\Lambda} f})(\xi)\ d\omega(\xi)\ d\omega(\zeta),$$
$$\tilde{r}^{(i)}_{j}(\widetilde{\mathbf{\Lambda} f})(\eta) = \int_\Omega \Psi^{(i)}_{j}(\eta,\zeta) \int_\Omega \Psi^{(i)}_{j}(\xi,\zeta) \cdot (\widetilde{\mathbf{\Lambda} f})(\xi)\ d\omega(\xi)\ d\omega(\zeta),$$

$j = J_0, \ldots, J-1$, have to be calculated by approximate integration in combination with the denoising procedure mentioned in Section 7.3.

We base our considerations on the approximate integration formulae with weights $v_s^{N_j}$, $w_l^{L_j} \in \mathbb{R}$ and knots $\zeta_s^{N_j}, \xi_l^{L_j} \in \Omega$, $s = 1, \ldots, N_j$; $l = 1, \ldots, L_j$, of the form:

$$\tilde{p}^{(i)}_{J_0}(f)(\eta) \simeq \sum_{s=1}^{N_{J_0}} (v^{(i)})_s^{N_{J_0}} \Phi^{(i)}_{J_0}(\eta, \zeta_s^{N_{J_0}}) (\tilde{a}^{(i)})_s^{N_{J_0}},$$
$$\tilde{r}^{(i)}_{j}(f)(\eta) \simeq \sum_{s=1}^{N_j} (v^{(i)})_s^{N_j} \Psi^{(i)}_j(\eta, \zeta_s^{N_j}) (\tilde{b}^{(i)})_s^{N_j}, \quad j = J_0, \ldots, J-1,$$

where

$$(\tilde{a}^{(i)})_s^{N_{J_0}} \simeq \sum_{l=1}^{L_{J_0}} (w^{(i)})_l^{L_{J_0}} \Phi^{(i)}_{J_0}(\xi_l^{L_{J_0}}, \zeta_s^{N_{J_0}}) \cdot (\widetilde{\mathbf{\Lambda} f})(\xi_l^{L_{J_0}}),$$
$$(\tilde{b}^{(i)})_s^{N_j} \simeq \sum_{l=1}^{L_j} (w^{(i)})_l^{L_j} \Psi^{(i)}_j(\xi_l^{L_j}, \zeta_s^{N_j}) \cdot (\widetilde{\mathbf{\Lambda} f})(\xi_l^{L_j}) \ .$$

The sign '$\simeq$' always means that the error is assumed to be negligible. A simple example is equidistribution (i.e. $(v^{(i)})_s^{N_j} = \frac{4\pi}{N_j}, s = 1, \ldots, N_j$; $(w^{(i)})_l^{L_j} = \frac{4\pi}{L_j}$, $l = 1, \ldots, L_j$).

Since the large 'true' coefficients are the ones that should be included in a selective reconstruction for estimating an unknown field, it is quite natural to include only coefficients $(\tilde{a}^{(i)})_s^{N_{J_0}}, (\tilde{b}^{(i)})_s^{N_J}$ larger than some specified threshold value.

In our context a 'larger' coefficient is taken to be one that satisfies the estimates

$$\left(\left(\tilde{a}^{(i)} \right)_s^{N_{J_0}} \right)^2 = \left| \int_\Omega \Phi^{(i)}_{J_0}(\alpha, \zeta_s^{N_{J_0}}) \cdot (\widetilde{\mathbf{\Lambda} f})(\alpha)\ d\omega(\alpha) \right|^2$$
$$= \int_\Omega \int_\Omega \left(\Phi^{(i)}_{J_0}(\alpha, \zeta_s^{N_{J_0}}) \otimes \Phi^{(i)}_{J_0}(\beta, \zeta_s^{N_{J_0}}) \right) \cdot \left((\widetilde{\mathbf{\Lambda} f})(\alpha) \otimes (\widetilde{\mathbf{\Lambda} f})(\beta) \right)\ d\omega(\alpha)\ d\omega(\beta)$$

$$
\begin{aligned}
&\simeq \sum_{p=1}^{L_{J_0}} \sum_{q=1}^{L_{J_0}} (w^{(i)})_p^{L_{J_0}} (w^{(i)})_q^{L_{J_0}} \left(\Phi_{J_0}^{(i)} \left(\xi_p^{L_{J_0}}, \zeta_s^{N_{J_0}} \right) \otimes \Phi_{J_0}^{(i)} \left(\xi_q^{L_{J_0}}, \zeta_s^{N_{J_0}} \right) \right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \cdot \left((\widetilde{\mathbf{\Lambda} f})(\xi_p^{L_{J_0}}) \otimes (\widetilde{\mathbf{\Lambda} f})(\xi_q^{L_{J_0}}) \right) \\
&\geq \sum_{p=1}^{L_{J_0}} \sum_{q=1}^{L_{J_0}} (w^{(i)})_p^{L_{J_0}} (w^{(i)})_q^{L_{J_0}} \left(\Phi_{J_0}^{(i)} \left(\xi_p^{L_{J_0}}, \zeta_s^{N_{J_0}} \right) \otimes \Phi_{J_0}^{(i)} \left(\xi_q^{L_{J_0}}, \zeta_s^{N_{J_0}} \right) \right) \cdot \mathbf{k}(\xi_p^{L_{J_0}}, \xi_q^{L_{J_0}}) \\
&\simeq \int_\Omega \int_\Omega \Phi_{J_0}^{(i)}(\alpha, \zeta_s^{N_{J_0}}) \otimes \Phi_{J_0}^{(i)}(\beta, \zeta_s^{N_{J_0}}) \cdot \mathbf{k}(\alpha, \beta) \; d\omega(\alpha) \; d\omega(\beta) \\
&= \kappa_{\Phi_{J_0}^{(i)}}(\zeta_s^{N_{J_0}})
\end{aligned}
$$

and

$$
\begin{aligned}
&\left(\left(\tilde{b}^{(i)} \right)_s^{N_j} \right)^2 = \left| \int_\Omega \Psi_j^{(i)}(\alpha, \zeta_s^{N_j}) \cdot (\widetilde{\mathbf{\Lambda} f})(\alpha) \; d\omega(\alpha) \right|^2 \\
&= \int_\Omega \int_\Omega \left(\Psi_j^{(i)}(\alpha, \zeta_s^{N_j}) \otimes \Psi_j^{(i)}(\beta, \zeta_s^{N_j}) \right) \cdot \left((\widetilde{\mathbf{\Lambda} f})(\alpha) \otimes (\widetilde{\mathbf{\Lambda} f})(\beta) \right) \; d\omega(\alpha) \; d\omega(\beta) \\
&\simeq \sum_{p=1}^{L_j} \sum_{q=1}^{L_j} (w^{(i)})_p^{L_j} (w^{(i)})_q^{L_j} \left(\Psi_j^{(i)} \left(\xi_p^{L_j}, \zeta_s^{N_j} \right) \otimes \Psi_j^{(i)} \left(\xi_q^{L_j}, \zeta_s^{N_j} \right) \right) \cdot \left((\widetilde{\mathbf{\Lambda} f})(\xi_p^{L_j}) \otimes (\widetilde{\mathbf{\Lambda} f})(\xi_q^{L_j}) \right) \\
&\geq \sum_{p=1}^{L_j} \sum_{q=1}^{L_j} (w^{(i)})_p^{L_j} (w^{(i)})_q^{L_j} \left(\Psi_j^{(i)} \left(\xi_p^{L_j}, \zeta_s^{N_j} \right) \otimes \Psi_j^{(i)} \left(\xi_q^{L_j}, \zeta_s^{N_j} \right) \right) \cdot \mathbf{k}(\xi_p^{L_j}, \xi_q^{L_j}) \\
&\simeq \int_\Omega \int_\Omega \Psi_j^{(i)}(\alpha, \zeta_s^{N_j}) \otimes \Psi_j^{(i)}(\beta, \zeta_s^{N_j}) \cdot \mathbf{k}(\alpha, \beta) \; d\omega(\alpha) \; d\omega(\beta) \\
&= \kappa_{\Psi_j^{(i)}}(\zeta_s^{N_j}) \; .
\end{aligned}
$$

For the threshold values $\kappa_{\Phi_{J_0}^{(i)}}(\zeta_s^{N_{J_0}})$, $s = 1, \ldots, N_{J_0}$; $\kappa_{\Psi_j^{(i)}}$, $j = 1, \ldots, N_j$, $i \in \{1, 2, 3\}$, such an estimator can be written in the form

$$
\begin{aligned}
(\widehat{\mathbf{\Lambda} f})_J^{(i)}(\eta) &= \sum_{s=1}^{N_{J_0}} I_{\{((\tilde{a}^{(i)})_s^{N_{J_0}})^2 \geq \kappa_{\Phi_{J_0}^{(i)}}(\zeta_s^{N_{J_0}})\}} (v^{(i)})_s^{N_{J_0}} \Phi_{J_0}^{(i)}(\eta, \zeta_s^{N_{J_0}}) (\tilde{a}^{(i)})_s^{N_{J_0}} \\
&+ \sum_{j=J_0}^{J-1} \sum_{s=1}^{N_j} I_{\{((b^{(i)})_s^{N_j})^2 \geq \kappa_{\Psi_j^{(i)}}(\zeta_s^{N_j})\}} (v^{(i)})_s^{N_j} \Psi_j^{(i)}(\eta, \zeta_s^{N_j}) (\tilde{b}^{(i)})_s^{N_j},
\end{aligned}
$$

where I_A denotes the indicator function of the set A ('$\wedge$' denotes 'estimated'). In other words, the large coefficients relative to the threshold values are kept intact, and the small coefficients are set to zero. The *thresholding estimators* of the true coefficients $p_{J_0}^{(i)}(f)(\eta)$, $r_j^{(i)}(f)(\eta)$ can thus be written in the form

$$
\hat{p}_{J_0}^{(i)}(f)(\eta) = \sum_{s=1}^{N_{J_0}} \delta^{\text{hard}}_{\kappa_{\Phi_{J_0}^{(i)}}(\zeta_s^{N_{J_0}})} \left(\left(\left(\tilde{a}^{(i)} \right)_s^{N_{J_0}} \right)^2 \right) (v^{(i)})_s^{N_{J_0}} \Phi_{J_0}^{(i)}(\eta, \zeta_s^{N_{J_0}}) (\tilde{a}^{(i)})_s^{N_{J_0}},
$$

$$
\hat{r}_j^{(i)}(f)(\eta) = \sum_{s=1}^{N_j} \delta^{\text{hard}}_{\kappa_{\Phi_j^{(i)}}(\zeta_s^{N_j})} \left(\left(\left(\tilde{b}^{(i)} \right)_s^{N_j} \right)^2 \right) (v^{(i)})_s^{N_j} \Psi_j^{(i)}(\eta, \zeta_s^{N_j}) (\tilde{b}^{(i)})_s^{N_j},
$$

where the function $\delta_\lambda^{\mathrm{hard}}$ is the *hard thresholding function*

$$\delta_\lambda^{\mathrm{hard}}(x) = \begin{cases} 1 & , \quad |x| \geq \lambda \\ 0 & , \quad |x| < \lambda \end{cases} . \tag{8.3}$$

The 'keep or kill' hard thresholding operation is not the only reasonable way to estimate the coefficients. Recognizing that the coefficients $\tilde{p}_{J_0}^{(i)}(f)(\eta)$, $\tilde{r}_j^{(i)}(f)(\eta)$ consist of both a signal portion and a noise portion, it might be desirable to attempt to isolate the signal contribution by removing the noisy part. This idea leads to the *soft thresholding function* (cf. D.L. Donoho, I.M. Johnstone (1994, 1995))

$$\delta_\lambda^{\mathrm{soft}}(x) = \begin{cases} \max\{0, 1 - \frac{\lambda}{|x|}\} & , \quad x \neq 0 \\ 0 & , \quad x = 0, \end{cases}$$

which can also be used in the above identities. When soft thresholding is applied to a set of empirical coefficients, only coefficients greater than the threshold (in absolute value) are included, but their values are 'shrunk' toward zero by an amount equal to the threshold λ.

Summarizing our results we finally obtain the following thresholding multiscale estimator

$$\begin{aligned} (\widehat{\mathbf{\Lambda} f})_J^{(i)}(\eta) = & \sum_{s=1}^{N_{J_0}} \delta_{\kappa_{\Phi_{J_0}^{(i)}}(\zeta_s^{N_{J_0}})} \left(\left(\left(\tilde{a}^{(i)} \right)_s^{N_{J_0}} \right)^2 \right) (v^{(i)})_s^{N_{J_0}} \Phi_{J_0}^{(i)}(\eta, \zeta_s^{N_{J_0}}) (\tilde{a}^{(i)})_s^{N_{J_0}} \\ & + \sum_{j=J_0}^{J-1} \sum_{s=1}^{N_j} \delta_{\kappa_{\Psi_j^{(i)}}(\zeta_s^{N_j})} \left(\left(\left(\tilde{b}^{(i)} \right)_s^{N_j} \right)^2 \right) (v^{(i)})_s^{N_j} \Psi_j^{(i)}(\eta, \zeta_s^{N_j}) (\tilde{b}^{(i)})_s^{N_j} . \end{aligned}$$

In doing so $(\widehat{\mathbf{\Lambda} f})_J$ first is approximated by a thresholded $(\widehat{\mathbf{\Lambda} f})_{J_0}$ which represents the denoised smooth components of the data. Then the coefficients at higher resolutions are thresholded so that the noise is suppressed but the fine–scale details are included in the calculation.

9. Numerical Example. In order to illustrate the the effectiveness of our multiscale denoising technique we present an example using synthetic geomagnetic data. For this purpose we introduce the usual moving local triad on the unit sphere, i.e. $\varepsilon^r, \varepsilon^\varphi$ and ε^t which, using spherical polar coordinates (see (2.2)), can be expressed in terms of the cartesian basis as follows:

$$\begin{aligned}
\varepsilon^r(\varphi,t) &= \sqrt{(1-t^2)}\left(cos(\varphi)\varepsilon^1 + sin(\varphi)\varepsilon^2\right) + t\varepsilon^3,\\
\varepsilon^\varphi(\varphi,t) &= -sin(\varphi)\varepsilon^1 + cos(\varphi)\varepsilon^2,\\
\varepsilon^t(\varphi,t) &= -t\left(cos(\varphi)\varepsilon^1 + sin(\varphi)\varepsilon^2\right) + \sqrt{1-t^2}\varepsilon^3.
\end{aligned}$$

Remark 9.1. Note that identifying 0 degree longitude with Greenwich and 0 degree latitude with the equator one can identify this local triad with the so-called geomagnetic coordinates $\{X,Y,Z\}$ (see Backus, G. *et al. (1996)) via $X = -\varepsilon^t, Y = \varepsilon^\varphi$ and $Z = -\varepsilon^r$. This means that X is always directed towards the geographic northpole, Y is always pointing into the geographic east direction while Z is always directed radially downward onto the Earth's body.*

9.1. Input Data. To generate an unnoised vectorial data set, we use a bandlimited geomagnetic potential due to C.J. Cain et al. (1984) and calculate the corresponding gradient field in geomagnetic coordinates $-\varepsilon^t$, ε^φ and $-\varepsilon^r$. This vector field also serves as a reference field for testing the quality of the denoised data. In a second step we add some so-called *bandlimited white noise* of variance σ and bandlimit n_K of approximately 2.9 and 60, respectively, to each of the three field components, i.e. $-\varepsilon^t$, ε^φ and $-\varepsilon^r$. Note that bandlimited white noise is characterized by the following symbol of the covariance kernel function:

$$\mathbf{k}^\wedge(n,k) = \begin{cases} \frac{\sigma^2}{(n_K+1)^2} & , \quad n \le n_K, \quad k = 1,\ldots,2n+1 \\ 0 & , \quad n > n_K, \quad k = 1,\ldots,2n+1. \end{cases}$$

This procedure resulted in noise of the order of magnitude 10^0 [nT] in field components of the order of magnitude 10^4 [nT] (see Figure 9.1 for a plot of the noise).

9.2. Results. In what follows we restrict ourselves to the radial component ($-\varepsilon^r$) and one of the tangential components (ε^φ), the results for the $-\varepsilon^t$ component are similar and will therefore be omitted.

In order to denoise the vector field we have decomposed and reconstructed the noised input data using spherical vectorial Shannon wavelets of type $i = 1, 2$ up to scale 3. Note that, since the input data is a gradient field, we need not use type 3 vector wavelets. During the reconstruction process only those wavelet coefficients containing a predominant amount of the clear signal have been used in accordance to our considerations in Section 7.3.3, i.e. we have used hard thresholding denoising.

In Figure 9.2 we show the unnoised and the noised $-\varepsilon^r$ component, while Figure 9.3 shows the corresponding values of the ε^φ component of the input data.

The reconstructed and denoised $-\varepsilon^r$ and ε^φ components as well as the corresponding errors (w.r.t. the unnoised data) can be seen in Figures 9.4 and 9.5, respectively.

Using our vectorial multiscale denoising technique, the root-mean-square error of the noised $-\varepsilon^r$ component (w.r.t. the unnoised data) $(\Delta\varepsilon^r_{noised})_{rms} = 4.9$ [nT] has been reduced to $(\Delta\varepsilon^r_{denoised})_{rms} = 0.6$ [nT], which is an improvement of about 87 per cent. In the ε^φ component we have $(\Delta\varepsilon^\varphi_{noised})_{rms} = 4.9$ [nT] and $(\Delta\varepsilon^r_{denoised})_{rms} =$

0.5 [nT] which is an improvement of about 89 per cent. As can be expected, comparing Figures 9.1 and 9.5 we see how the comparatively rough structure of the noise has been smoothed out by the denoising procedure. This example obviously demonstrates the functionality and efficiency of our multiscale approach.

For illustrational purposes we furthermore present Figures 9.6 (scale $j = 1$), 9.7 (scale $j = 2$) and 9.8 (scale $j = 3$) which give an impression of what happens on the different scales during the denoising process. The plots on the left hand side show the difference between the undenoised and the denoised partial wavelet reconstruction of the noised data. The right hand sides give an illustration of the difference between the denoised partial reconstruction of the noised data and a partial reconstruction of an unnoised data set. On the left of Figure 9.6 one can hardly see any structure and the order of magnitude of the plotted difference is 10^{-6}. This is understandable if one takes into account that the noise is very small compared to the vector field at scale $j = 1$. Nevertheless we can see an error in the reconstruction (right hand side of Figure 9.6) which, though small in magnitude, has a large spatial structure. This large spatial extend reflects the typical size of spatial features at this scale. The difference between the denoised and the undenoised partial reconstruction at scale $j = 2$ is of the order 10^{-3} and can be seen on the left of Figure 9.7. This increasing difference shows that the noise plays a more important role at scale $j = 2$ than at scale $j = 1$ since the noise at this scale becomes comparable with the vector field. Of course, this results in an increased - but still small - error of the denoised reconstruction with respect to the corresponding unnoised data (Figure 9.7, right). Again the spatial extend of the visible features is correspondent to the typical lenghtscales at scale $j = 2$. Going up to scale $j = 3$ we obtain similar effects, i.e. an increasing difference between the denoised and the undenoised partial reconstruction (order of magnitude 10^0, Figure 9.8, left) and a larger error with respect to the unnoised data (Figre 9.8, right). Again the reason is the noise which, at this scale, is of the same order of magnitude like the vector field.

Last but not least we need to mention that for physical reasons it is clear that in the case of real satellite data we cannont assume the input data to be a pure gradient field. This means that type 3 vector wavelets need to be used in the denoising process, too. However, this will not be of any difficulty since various vector wavelet modelling of all types ($i = 1, 2, 3$) have been already succesfully applied to geomagnetic satellite data e.g. by Maier, T. (1999) and Bayer, M. et al. (2000). As far as the noise is concerned the use of 'bandlimited white noise' is a good and widespread tool for testing the effectiveness and efficiency of the numerical procedures. More realistic noise models need to be tested as soon as they are available for recent satellite missions like Oersted or CHAMP, for example.

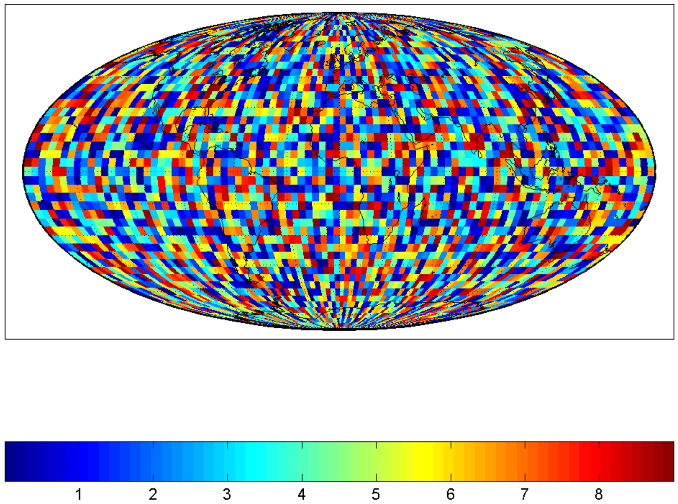

FIG. 9.1. *Absolute value of noise [nT]*

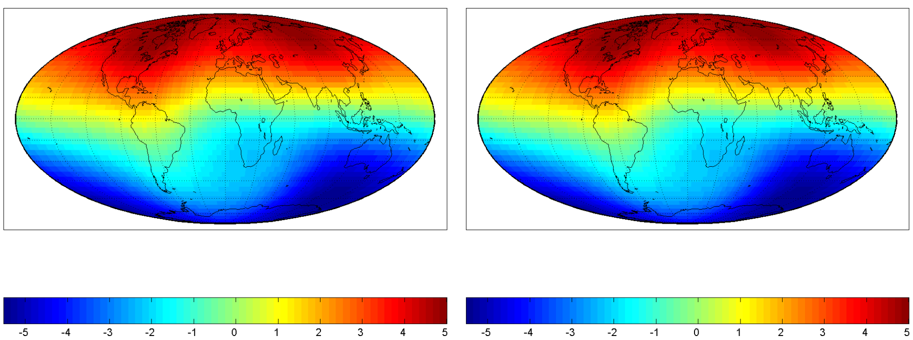

FIG. 9.2. *Unnoised (left) and noised (right)* $-\varepsilon^r$ *component in 10000 [nT]*

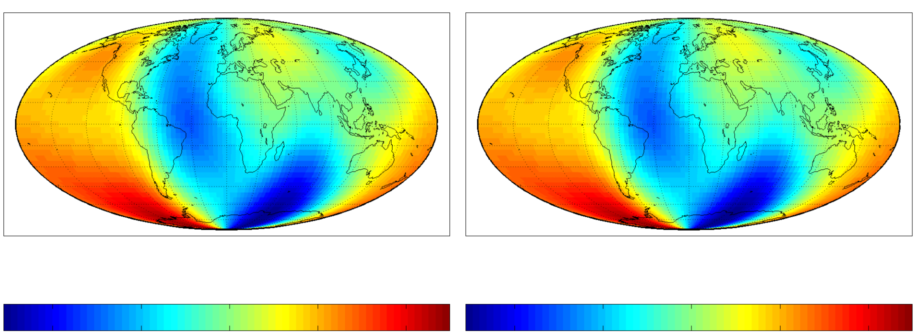

FIG. 9.3. *Unnoised (left) and noised (right)* ε^φ *component in 10000 [nT]*

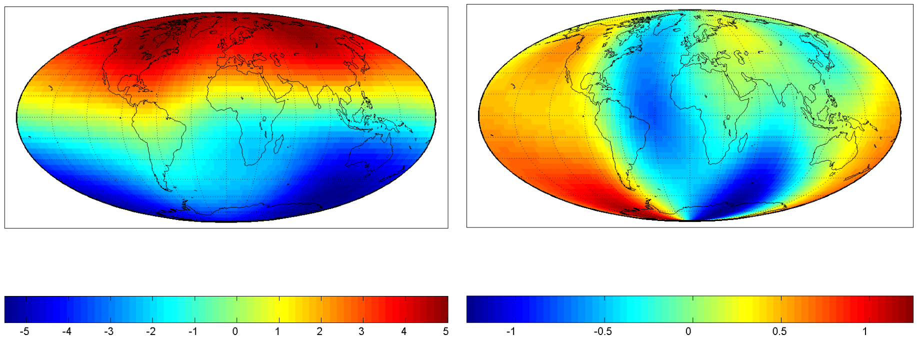

FIG. 9.4. *Denoised* $-\varepsilon^r$ *component (left) and* ε^φ *component (right) in 10000 [nT]*

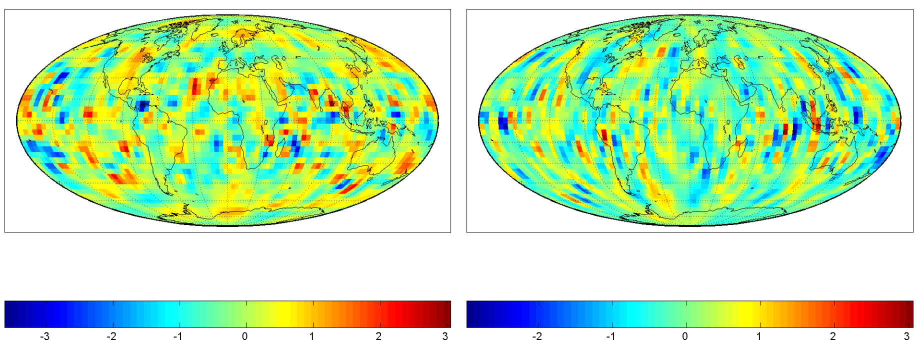

FIG. 9.5. *Error of denoised* $-\varepsilon^r$ *(left) and* ε^φ *(right) component w.r.t. unnoised data [nT]*

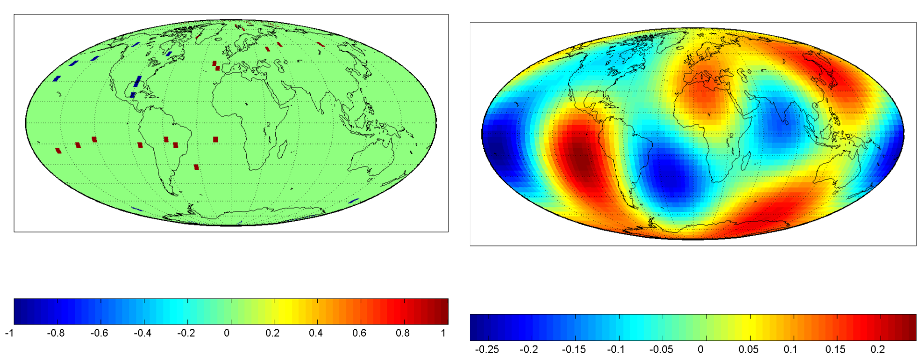

FIG. 9.6. *difference (left,* 10^{-6} *[nT]) and error plot (right, [nT]), scale* $j = 1$

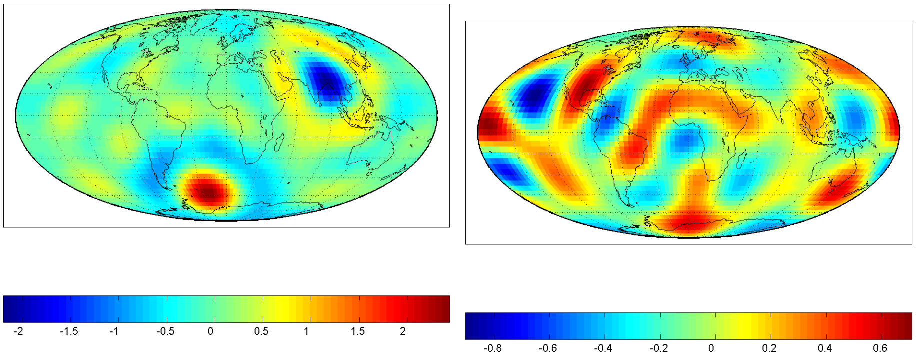

FIG. 9.7. *difference (left, 10^{-3} [nT]) and error plot (right, [nT]), scale $j = 2$*

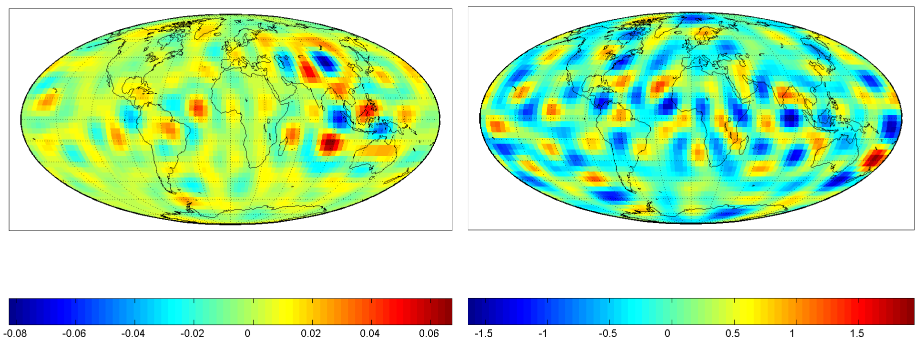

FIG. 9.8. *difference (left, 10^{0} [nT]) and error plot (right, [nT]), scale $j = 3$*

Acknowledgements.

The support by German Research Foundation (DFG contract No. FR 761-10-1) is gratefully acknowledged.

References.

BACKUS, G., PARKER, R., CONSTABLE, C. (1996) *Foundation of Geomagnetism.* Cambridge University Press.

BAYER, M. (2000) Geomagnetic Field Modelling From Satellite Data by First and Second Generation Vector Wavelets. Doctoral Thesis, University of Kaiserslautern, Geomathematics Group, Shaker, Aachen.

BAYER, M., BETH, S., FREEDEN, W. (1998) *Geophysical Field Modelling by Multisresolution Analysis.* Acta Geod. Geoph. Hung., **33**, 289–319.

BAYER, M., FREEDEN, W., MAIER, T. (2000) *A Vector Wavelet Approach to Iono- and Magnetospheric Geomagnetic Satellite Data.* Journal of Atmospheric and Solar-Terrestrial Physics, accepted for publication.

BEN MENAHEM, A., SINGH, S.J. (1981) *Seismic Waves and Sources.* Springer, Berlin, Heidelberg.

BETH, S. (2000) Multiscale Approximation by Vector Radial Basis Functions on the Sphere. Doctoral Thesis, University of Kaiserslautern, Geomathematics Group, Shaker, Aachen.

CAIN, C.J., SCHMITZ, D.R., MUTH, L. (1984) *Small-Scale Features in the Earth's Magnetic Field Observed by MAGSAT.* Journal of Geophysical Research, **89**, 1070-1076.

DONOHO, D.L., JOHNSTONE, I.M. (1994) *Ideal Spatial Adaptation by Wavelet Shrinkage.* Biometrika, **81**, 425-455.

DONOHO, D.L., JOHNSTONE, I.M. (1995) *Adapting to Unknown Smoothness Via Wavelet Shrinkage.* J. Amer. Statistical Association, **90**, 1200-1224.

FREEDEN, W. (1980) *On Integral Formulas on the (Unit) Sphere and Their Application to Numerical Computation of Integrals.* Computing, **25**, 131-146.

FREEDEN, W. (1999): Multiscale Modelling of Spaceborne Geodata. B.G. Teubner, Stuttgart, Leipzig.

FREEDEN, W., GERVENS, T., SCHREINER, M. (1998) Constructive Approximation on the Sphere (With Applications to Geomathematics). Oxford Science Publications, Clarendon.

FREEDEN, W., GLOCKNER, O., THALHAMMER, M. (1999) *Multiscale Gravitational Field Recovery from GPS–Satellite–to–Satellite Tracking.* Studia Geophys. et Geodaet. **43**, 229–264.

FREEDEN, W., HESSE, K., (2000) *Locally Supported Difference Wavelets on the Sphere.* AGTM-Report 225, Laboratory of Technomathematics, Geomathematics Group, University of Kaiserslautern.

FREEDEN, W., MICHEL, V., STENGER, M., (2000) *Multiscale Signal–to–Noise Thresholding.* AGTM-Report 224, Laboratory of Technomathematics, Geomathematics Group, University of Kaiserslautern.

LEBEDEW N.N. (1973) Spezielle Funktionen und ihre Anwendung. BI–Wissenschaftsverlag, Mannheim, Wien, Zürich.

MAIER, T. (1999) *Multiscale Analysis of the Geomagnetic Field.* Diploma Thesis, University of Kaiserslautern, Geomathematics Group.

MAIER, T. AND BAYER, M. (1998) *Multiscale Analysis of the Geomagnetic Field.* W. Freeden (Ed.): Progress in Geodetic Science (at GW98), Shaker Verlag, Aachen, 274–283.

Michel, V. (1999) A Multiscale Method for the Gravimetry Problem — Theoretical and Numerical Aspects of Harmonic and Anharmonic Modelling. Doctoral Thesis, University of Kaiserslautern, Geomathematics Group, Shaker Verlag, Aachen.

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

Narcovich, F.J., Ward, J.D. (1996) *Nonstationary Wavelets on the m-Sphere for Scattered Data*, Appl. Comp. Harm. Anal, 3, 324–336

Ogden, R.T. (1997) Essential Wavelets for Statistical Applications and Data Analysis. Birkhäuser, Boston, Basel, Berlin.

Rummel, R. (1979) *Determination of the Short–Wavelength Components of the Gravity Field from Satellite–to–Satellite Tracking or Satellite Gradiometry.* Manuscr. Geod., **4**, 107–148

Rummel, R. (1997) *Spherical Spectral Properties of the Earth's Gravitational Potential and its First and Second Derivatives.* In: Lecture Notes in Earth Sciences, **65**, 359–404.

Weaver, J.B., Yansun, X., Healy, D.M.Jr., Cromwell L.D. (1991): *Filtering Noise From Images With Wavelet Transforms.* Magnetic Resonance in Medicine, **24**, 288-295.