Electronic Transactions on Numerical Analysis.
Volume 14, pp. 40-62, 2002.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

ON MULTISCALE DENOISING OF SPHERICAL FUNCTIONS: BASIC THEORY AND NUMERICAL ASPECTS*

W. FREEDEN AND T. MAIER†

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 proof, including non-bandlimited kernel functions, of the pyramid scheme is presented to efficiently remove the noise at different scales using a priori statistical information, i.e. knowledge of the covariance function.

Key words. spherical wavelet theory, scalar multiscale approximation, pyramid scheme, spectral and multiscale variance-covariance model, hard and soft thresholding.

AMS subject classifications. 33C55, 42C40, 62-07, 65T60, 86A25.

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. [6, 8]). However, the amount of localization in space and in frequency is not completely satisfactory in the 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 extent – the data are contaminated with noise. That is, in order to successfully 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. The Euclidean theory of wavelets provides signal processors with new, fast tools that are well–suited for denoising signals (for a survey the reader is e.g. referred to [13] and the references therein).

*Received August 8, 2001. Accepted for publication May 20, 2002. Communicated by Sven Ehrich.
†University of Kaiserslautern, Geomathematics Group, 67653 Kaiserslautern, P.O. Box 3049, Germany, email freeden@mathematik.uni-kl.de, tmaier@mathematik.uni-kl.de, www http://www.mathematik.uni-kl.de/∼wwwgeo; Correspondence to W. Freeden

ETNA
Kent State University
etna@mcs.kent.edu

The objective of this article is to introduce multiscale signal-to-noise thresholding and to provide the wavelet oriented basis of denoising spherical data sets. First, we develop the corresponding theory of denoising spherical functions by non-rotation invariant wavelets (cf. [9] for the rotation-invariant case). 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. [15, 2, 3]) and the stochastic model used in satellite geodesy (see e.g. [14]). Using a multiscale approach we are thus able to include detail information of small spatial extent while suppressing the noise in the approximation appropriately. A simple example of denoising geomagnetic field data will be given as an illustration.

2. Preliminaries. Let $\mathbb{R}^3$ denote 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 this paper we assume this system to be orthonormal in the sense of the $\mathcal{L}^2(\Omega)$–inner product

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

ETNA
Kent State University
etna@mcs.kent.edu

($d\omega$ 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

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

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 [12, 6].

We let

$$Harm_{0,\ldots,m} = \underset{\substack{n=0,\ldots,m\\ k=1,\ldots,2n+1}}{span} (Y_{n,k})\,.$$

Of course,

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

so that

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

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

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

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

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

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

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 consider 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

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

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

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

The series

$$\sum_{(n,k)\in\mathcal{N}} F^\wedge(n,k) Y_{n,k}$$

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

$$H(\xi,\eta) = \sum_{(n,k)\in\mathcal{N}} H^\wedge(n,k) Y_{n,k}(\xi) Y_{n,k}(\eta)$$

such that

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

In the case of rotational invariance of the kernel H, i.e. $H^\wedge(n,k) = H^\wedge(n)$ for $n = 0,1,\ldots, k = 1,\ldots,2n+1$, the last condition is equivalent to $l^2(\mathcal{N})$-summability (cf. [6]), 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 mentioned: (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 strictly monotonically decreasing sequence $\{\rho_j\}_{j\in\mathbb{Z}}$ of real numbers that satisfies

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

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, strictly monotonically decreasing sequence. Let $\{\Phi_{\rho_j}\}_{j\in\mathbb{Z}}$ be a family of square-summable product kernels satisfying the condition $\Phi_{\rho_j}^\wedge(0,1) = 1$ 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) $\lim_{j\to\infty} \left(\Phi_{\rho_j}^\wedge(n,k)\right)^2 = 1$ for all $(n,k) \in \mathcal{N}$,

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

(iii) $\lim_{j\to-\infty} \left(\Phi_{\rho_j}^\wedge(n,k)\right)^2 = 0$ for all $(n,k) \in \mathcal{N}$,

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

$$I_{\rho_j}^{(2)} = \Phi_{\rho_j}^{(2)} * F, \quad j \in \mathbb{Z}$$

is called an *approximate identity in* $\mathcal{L}^2(\Omega)$. It is known (see e.g. [6]) that

$$\lim_{j\to\infty} \|I_{\rho_j}^{(2)}(F) - F\|_{\mathcal{L}^2(\Omega)}$$

$$= \lim_{j\to\infty} \left(\sum_{(n,k)\in\mathcal{N}} (F^\wedge(n,k))^2 \left(1-\Phi^\wedge_{\rho_j}(n,k)\right)^4 \right)^{1/2}$$
$$= 0 \ ,$$

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*

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

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 $J\in\mathbb{Z}$,

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

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 [6] for more details of spherical theory, [5] for the application harmonic theory, and [11] 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 $\mathcal{L}^2(\Omega)$.

ETNA
Kent State University
etna@mcs.kent.edu

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.

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

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

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

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

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

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

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

where $\{\rho_j\}_{j\in\mathbb{Z}}$ is defined as in the Shannon case and $\{\sigma_j\}_{j\in\mathbb{Z}}$ is a strictly 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 strictly 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.

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

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

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

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

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.

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

It should be remarked that, even if we only consider bandlimited functions as data, one might want to use non-bandlimited kernel functions for the multiscale analysis. This is because of the characteristics of the kernel functions in the space domain. Non-bandlimited kernel functions show less oscillations in the space domain than bandlimited kernels, which is surely a desirable feature if data of high degree and order is to be analyzed.

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. [13]).

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

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

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

(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 rotational-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.

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

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

$$\begin{aligned}\tilde{G} = \widetilde{\Lambda F} &= \sum_{(n,k)\in\mathcal{N}} \Lambda^\wedge(n,k) \tilde{F}^\wedge(n,k) Y_{n,k} \\ &= \sum_{(n,k)\in\mathcal{N}} \tilde{G}^\wedge(n,k) Y_{n,k}\end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

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

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

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

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

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*

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

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 $n_K \in \mathbb{N}_0$

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

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

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

Note that this sum, apart from a multiplicative constant, may be understood as a truncated Dirac δ-functional. It is known (see e.g. [10]) 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) $K(\xi \cdot \eta) = \frac{\sigma^2}{(L_{\rho_{J^*}}^{(s)})^{(2)}(1)} (L_{\rho_{J^*}}^{(s)})^{(2)}(\xi \cdot \eta), \qquad (\xi,\eta) \in \Omega \times \Omega,$

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

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

where

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

For the case $k = 0$ this example is known as the *Haar singular integral* (more details about Haar wavelets can be found in [7]).

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* $\mathcal{N}_{\text{res}}$ (with $\mathcal{N}_{\text{res}} \subset \mathcal{N}$). We distinguish the following cases:

(i) *signal dominates noise*

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

(ii) *noise dominates signal*

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

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*

$$\widehat{\Lambda F} = \sum_{(n,k)\in\mathcal{N}} I_{\mathcal{N}_{\rm res}}(n,k) H^\wedge(n,k) \left(\widetilde{\Lambda F}\right)^\wedge (n,k) Y_{n,k}, \tag{7.3}$$

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

$$H^\wedge(n,k) = H^\wedge(n) = \begin{cases} 1 & , \quad (n,k) \in \mathcal{N}_{res} \\ 0 & , \quad (n,k) \notin \mathcal{N}_{res}, \end{cases} \tag{7.4}$$

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

$$\widehat{\Lambda F} = \sum_{(n,k)\in\mathcal{N}} H^\wedge(n) \left(\widetilde{\Lambda F}\right)^\wedge (n,k) Y_{n,k} \tag{7.5}$$

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 rotational-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 (cf. [9])

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

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

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

Expressed in the spectral language of spherical harmonics we get

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

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

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

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 $\eta\in\Omega$

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

The *scale error covariance* is defined by

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

We obviously have in spectral language

$$Cov_{j;\eta}(K) = \sum_{(n,k)\in\mathcal{N}} K^\wedge(n,k)\left(\Psi^\wedge_{\rho_j}(n,k)\right)^2 \quad .$$

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

$$Cov_{j;\eta}(K) = \frac{1}{4\pi}\sum_{(n,k)\in\mathcal{N}} Cov_n(K)\left(\Psi^\wedge_{\rho_j}(n,k)\right)^2 \quad .$$

Our error model is particularly useful for the proper handling of the satellite data in Earth's gravitational or magnetic potential determination (see [5] and the references therein).

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

(i) *signal dominates noise*

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

(ii) *noise dominates signal*

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

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

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

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. $(\widetilde{\Lambda F})_J \simeq \widetilde{\Lambda F}$.

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

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

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

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

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

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

and

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

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, [6] 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 $\eta_i^{N_j} \in \Omega$

$$\tilde{V}_{J_0;\eta} \backsimeq \sum_{i=1}^{N_{J_0}} w_i^{N_{J_0}} \Phi^{(2)}_{\rho_{J_0}}(\eta, \eta_i^{N_{J_0}})(\widetilde{\Lambda F})(\eta_i^{N_{J_0}}),$$

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

('$\backsimeq$' always means that the error is assumed to be negligible). An example (cf. [9]) 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

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

we want to show that the coefficient vectors $\tilde{a}^{N_j} = \left(\tilde{a}_1^{N_j}, \ldots, \tilde{a}_{N_j}^{N_j}\right)^T \in \mathbb{R}^{N_j}$ $j = J_0, \ldots, J-1$, (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 $\tilde{a}^{N_j}$, $j = J_0, \ldots, J-1$, are obtainable by recursion from the values $\tilde{a}_i^{N_J}$.

(ii) For $j = J_0, \ldots, J$

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

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

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

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, the kernel $\Phi^{(2)}_{\rho_J}$ replaces the Dirac–functional δ as follows:

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

where

$$\delta(\xi, \eta) = \delta_\xi(\eta) = \sum_{(n,k)\in\mathcal{N}} Y_{n,k}(\xi) Y_{n,k}(\eta ta)$$

ETNA
Kent State University
etna@mcs.kent.edu

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

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

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 $j = J_0,\dots,J$.
Note that for bandlimited scaling functions the kernels Ξ_j, $j = J_0,\dots,J$, 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, $j = J_0,\dots,J$, 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

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

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 $n = 0,1,\dots,\ \ k = 1,\dots,2n+1$). This finally leads us to the formulae

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

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. [5]).

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:

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

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

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

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

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

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

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 [2, 3])

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

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.

ETNA
Kent State University
etna@mcs.kent.edu

11. Example. In order to illustrate the effectiveness of our multiscale denoising technique we present a simple example using synthetic geomagnetic data. It is clear that the method presented here can also be applied to non-synthetic data, but then it is disproportionately more difficult to compare the results with expected outcomes. For the purpose of our example 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:

$$X \leftrightarrow \varepsilon^t,$$
$$Y \leftrightarrow \varepsilon^\varphi,$$
$$Z \leftrightarrow -\varepsilon^r,$$

where $\varepsilon^t, \varepsilon^\phi$ and ε^r are the usual unit vectors in spherical polar coordinates (for explicit representations see e.g. [6]). 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, which is realistic for the geomagnetic main field) geomagnetic potential due to [1] we calculated the corresponding gradient field in geomagnetic coordinates, i.e. north (ε^t), east (ε^φ) and downward ($-\varepsilon^r$) components, which gave as noise-free data (note that we have used the low degree model just for illustrational purposes; it is clear that multiscale techniques are also valuable tools for handling phenomena of especially high degree). We then added some bandlimited white noise with variance σ and bandwidth n_K of approximately 0.9 and 60, respectively (see section 7.4). This resulted in noise of the order of magnitude 10^0 $[\mathrm{nT}]$ in a field of the order of magnitude 10^4 $[\mathrm{nT}]$. It should be noted that, when looking at the pictures, the noise is not constant at the poles as one should expect it to be. This is due to our routine of adding the noise to the synthetic data. However, our results are not influenced by this, since during the process of decomposition and reconstruction each data point of the rectangular domain is weighted by integration weights due to [4]. These weights are constructed such that the poles do not contribute to the whole integration.

The noise signal then was decomposed and reconstructed using Shannon wavelets up to scale 4 (see construction principles in Section 6.1, where we have chosen $\rho_j = 2^{-j}$ and $j = 1, 2, 3, 4$). During the reconstruction process only those wavelet coefficients containing a predominant amount of the clear signal were 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), $(\Delta\varepsilon^r_{\mathrm{noised}})_{\mathrm{rms}} = 1.13$ $[\mathrm{nT}]$, has been reduced to $(\Delta\varepsilon^r_{\mathrm{denoised}})_{\mathrm{rms}} = 0.35$ $[\mathrm{nT}]$, which is an improvement of about 60 per cent.

Comparing Figs. 11.2 and 11.4 it can be seen 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 shows the functionality of our approach.

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

REFERENCES

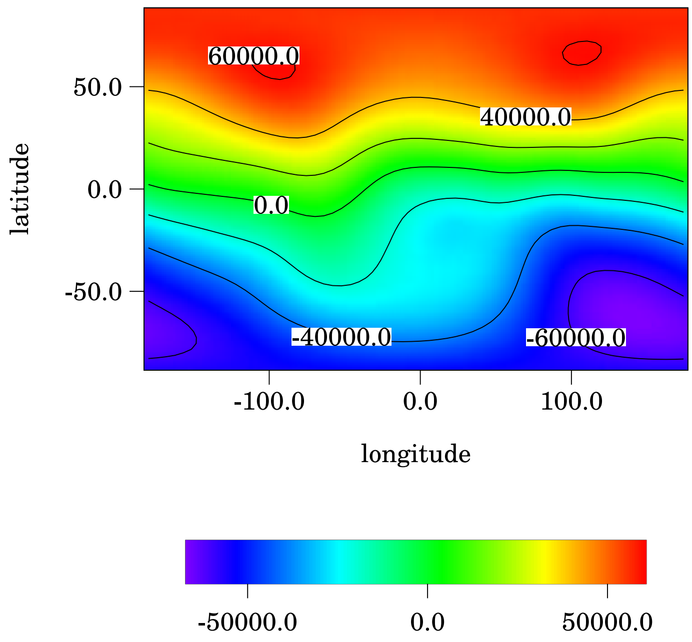

FIG. 11.1. $-\varepsilon^r$ *component of unnoised data [nT]*

[1] C.J. CAIN, D.R. SCHMITZ, AND L. MUTH, *Small-scale features in the earth's magnetic field observed by MAGSAT*. Journal of Geophysical Research, Vol. 89, NO. B2 (1984), pp. 1070-1076.

[2] D.L. DONOHO AND I.M. JOHNSTONE, *Ideal spatial adaptation by wavelet shrinkage*. Biometrika, 81 (1994), pp. 425-455.

[3] ———, *Adapting to unknown smoothness via wavelet shrinkage*, J. Amer. Statistical Association, 90 (1995), pp. 1200-1224.

[4] J.R. DRISCOLL AND D.M. HEALY, *Computing Fourier transforms and convolutions on the 2-sphere*. Adv. Appl. Math., 15 (1994), pp. 202-250.

[5] W. FREEDEN, *Multiscale Modelling of Spaceborne Geodata*, B.G. Teubner, Stuttgart, Leipzig, 1999.

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

[7] W. FREEDEN AND K. HESSE, *On the multiscale solution of satellite problems by use of locally supported kernel functions corresponding to equidistributed data on spherical orbits*. AGTM-Report 233, Laboratory of Technomathematics, Geomathematics Group, University of Kaiserslautern, 2000.

[8] W. FREEDEN, V. MICHEL, *Constructive Approximation and Numerical Methods in Geodetic Research Today - an Attempt at a Categorization Based on an Uncertainty Principle*, J. of Geodesy, 73 (1999), pp.452-465.

[9] W. FREEDEN, V. MICHEL, AND M. STENGER, *Multiscale signal-to-noise thresholding*, Acta Geod. Geoph. Hung., Vol 36(1) (2001), pp. 55-86.

[10] N.N. LEBEDEW, *Spezielle Funktionen und ihre Anwendung*, BI-Wissenschaftsverlag, Mannheim, Wien, Zürich, 1973.

[11] V. MICHEL, *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, 1999.

[12] C. MÜLLER, *Spherical Harmonics*. Lecture Notes in Mathematics 17, Springer, Berlin, Heidelberg, 1966.

[13] R.T. OGDEN, *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhäuser, Boston, Basel, Berlin, 1997.

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

[15] J.B. WEAVER, X. YANSUN, D.M. HEALY JR., AND L.D. CROMWELL, *Filtering noise from images with wavelet transforms*, Magnetic Resonance in Medicine, 24 (1991), pp. 288-295.

ETNA
Kent State University
etna@mcs.kent.edu

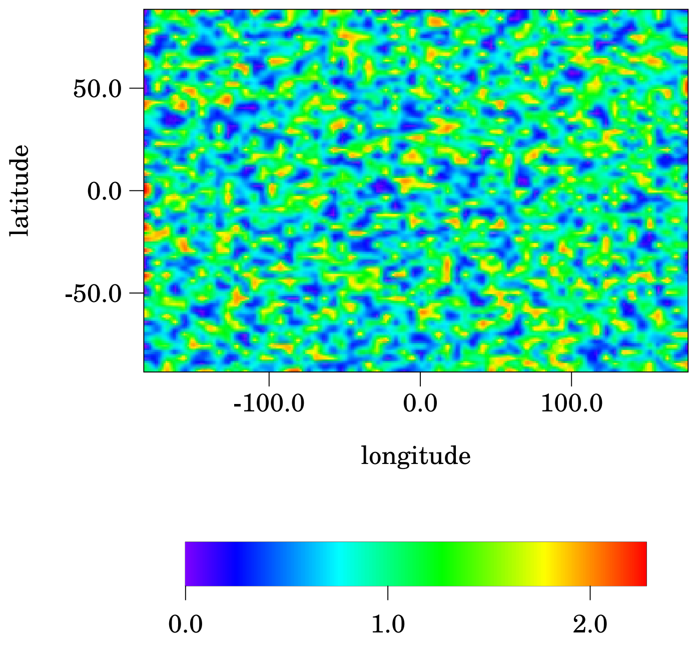

FIG. 11.2. *Absolute value of added noise [nT]*

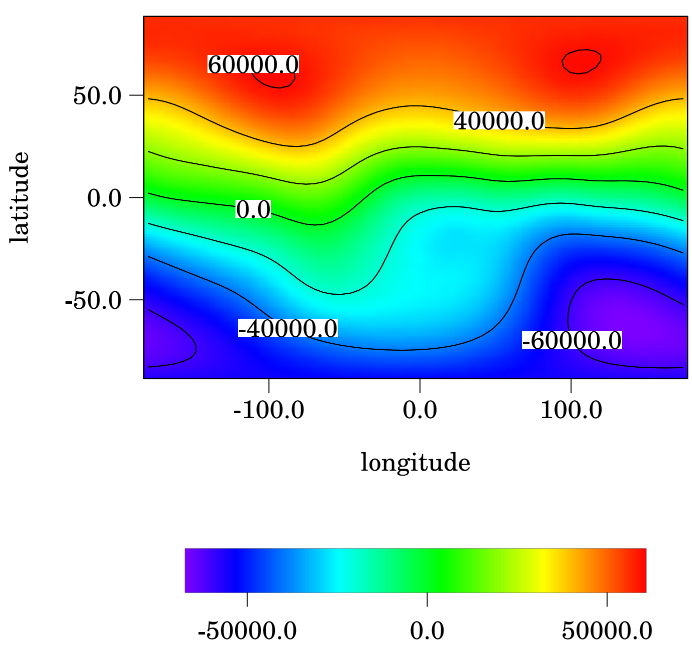

FIG. 11.3. $-\varepsilon^r$ *component of denoised data [nT]*

ETNA
Kent State University
etna@mcs.kent.edu

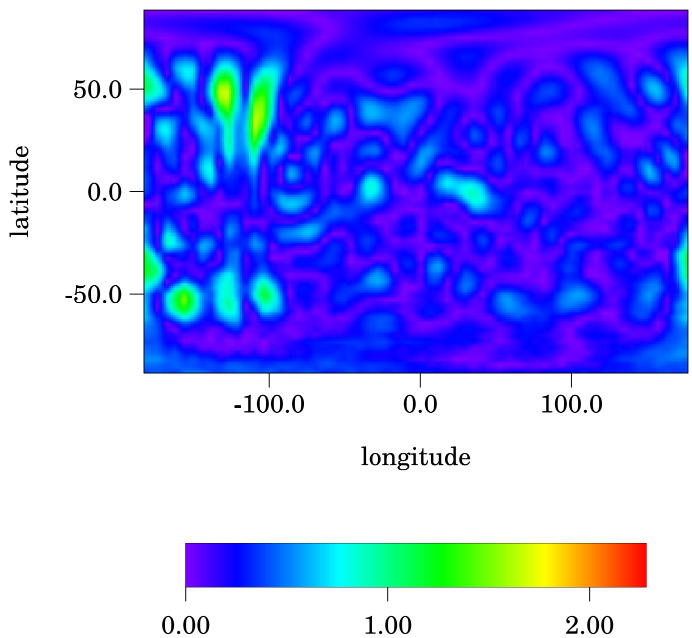

FIG. 11.4. *Absolute error of denoised* $-\varepsilon^r$ *component [nT]*