Spherical harmonic analysis, aliasing, and filtering

Christopher Jekeli

Department of Geodetic Science and Surveying, The Ohio State University, 1958 Neil Ave., Columbus, OH 43210, USA

Received 5 October 1994; Accepted 14 June 1995

Abstract. The currently practiced methods of harmonic analysis on the sphere are studied with respect to aliasing and filtering. It is assumed that a function is sampled on a regular grid of latitudes and longitudes. Then, transformations to and from the Cartesian plane yield formulations of the aliasing error in terms of spherical harmonic coefficients. The following results are obtained: 1) The simple quadratures method and related methods are biased even with band-limited functions. 2) A new method that eliminates this bias is superior to Colombo's method of least squares in terms of reducing aliasing. 3) But, a simple modification of the least-squares model makes it identical to the new method as one is the dual of the other. 4) The essential elimination of aliasing can only be effected with spherical cap averages, not with the often used constant angular block averages.

Introduction

The harmonic (or spectral) representation of functions on the sphere has proved useful in modeling the gravity field, the magnetic field, the topography, the ocean tides, etc, not just for the Earth, but for any roughly spherical planet (Rapp 1977; Schmitz and Cain 1983; Kaula 1993). Such harmonic representations enable a more precise interpretation by wavelength of their magnitudes, their coherence, and their measurability. Usually, these functions are richly endowed with harmonics to a high degree, decaying relatively slowly with increasing detail. For example, the Earth's gravity field, after the first few harmonics (which certainly dominate), decays approximately linearly on the logarithmic scale, and today's representations include over 130 thousand terms (Rapp and Pavlis 1990).

It is a significant consequence of this fine structure in the functions that the harmonic components determined from a finite set of function values are biased estimates of the true components. This is a well known phenomenon in signal analysis, where even though one applies orthogonal operators to the data, they cannot filter out harmonics finer in detail than that dictated by the sampling interval. This is known as "aliasing." Another interpretation, though less precise, is that of attempting to determine all (possibly an infinite number of) harmonics from a given number of function values. In the general case, this is an underdetermined problem and the solution invariably lumps together harmonics in a prescribed way. Spectral aliasing in functions defined on the line or on the Cartesian plane is well understood; it is less obvious on a non-Euclidean surface such as the sphere.

The spherical harmonic analysis is examined here with the object of clearly understanding the effect of aliasing in conventional techniques, as well as in the modern techniques developed by Colombo (1981) that are used to compute today's models. In so doing, a new formula for the analysis is rendered that reduces the aliasing effect better than the least-squares approach. It is, however, also shown that the new technique is equivalent to a modification of the least-squares model, requiring some extra computations. The results also indicate clearly which types of averaging (filtering) do and do not eliminate aliasing. It is assumed throughout that the data are provided discretely at every node of a uniform grid. (Aliasing in the case of a nonuniform data distribution was studied by Sansó 1990.) Though the errors in measuring the function on the sphere are not considered initially, their effect on the derived procedures is discussed in conclusion.
Preparatory Concepts

A periodic function integrable over its period may be represented either in terms of its independent (space or time) variable or, where it is continuous, in terms of its Fourier spectrum, being the set of coefficients in its corresponding Fourier series (Priestley 1981):

\[ g(x) = \sum_{k=-\infty}^{\infty} G_k e^{i2\pi kx/T}, \quad x \in \mathbb{R} \]  
(1a)

\[ G_k = \frac{1}{T} \int_{0}^{T} g(x) e^{-i2\pi kx/T} dx, \quad k \in \mathbb{Z} \]  
(1b)

where \( i^2 = -1 \), \( T \) is the period \((g(x+T) = g(x))\), \( \mathbb{R} \) is the set of real numbers, \( \mathbb{Z} \) is the set of integers, and the integer \( k \) may be termed "frequency," or "wavenumber." In practice, for example from measurements, one knows only discrete values of \( g \) and the corresponding spectrum is the Discrete Fourier Transform (DFT):

\[ g_s = \sum_{s=0}^{K-1} G_k e^{i2\pi ks/K}, \quad s = 0, ..., K-1 \]  
(2a)

\[ G_k = \frac{\Delta x}{T} \sum_{s=0}^{K-1} g_s e^{-i2\pi ks/K}, \quad k = 0, ..., K-1 \]  
(2b)

where it is assumed that the function values are known on a regular grid, i.e., one with constant spacing, \( \Delta x = T/K \). The unsymmetric transform (2a,b), where \( k \) is nonnegative, is preferred here as being more conventional and avoiding the separate treatment of even and odd \( K \). Since the DFT is periodic with frequency period \( K \):

\[ G_{k+K} = G_k, \quad \text{for any } k \]  
(2c)

it is noted that equations (2a,b) are equivalent to the symmetric forms, where \( k \) (and \( s \)) ranges from -(K-1)/2 to (K-1)/2 (if \( K \) is odd). In the notation of (2a,b), the frequencies from (K+1)/2 to K-1 are used to represent the negative frequencies from -(K-1)/2 to -1.

The relationship between the spectra (1b) and (2b) can be found by substituting \( x = s\Delta x \) into (1a):

\[ g_s = g(s\Delta x) = \sum_{j=-\infty}^{\infty} G_j e^{i2\pi js\Delta x/K} = \sum_{k=0}^{K-1} \sum_{j=-\infty}^{\infty} G_{jk+k} e^{i2\pi jk/K} \]  
(3)

from which, by comparing with (2a), one has for any \( k \),

\[ \hat{G}_k = \sum_{j=-\infty}^{\infty} G_{jk+k} = G_k + \sum_{j=0}^{\infty} G_{jk+k} \]  
(4)

The infinite sum on the far right in (4) is an alias of \( G_k \) for any frequency, \( k \), such that \( |k| < k_N = K/2 \), where \( k_N \) is the so-called Nyquist frequency. The determination of the spectrum \( \{G_k\} \) of the function \( g \) from the discrete sequence \( \{g_s\} \), using (2b), is subject to an aliasing error as formulated in (4). Note that if the function \( g \) does not contain harmonics with frequency above the Nyquist frequency, then there is no aliasing error. Such functions will be called band-limited. It is further noted that a function with infinite bandwidth can be filtered to make it band-limited. This procedure is discussed later with respect to the particular application of spherical harmonic analysis.

Obviously, similar formulas hold for periodic functions defined on the Cartesian plane. They are specialized here to the case where \( g \) is a real function of two variables, \( \theta \) and \( \lambda \), and is periodic in \( \theta \) with period \( \pi \) and in \( \lambda \) with period \( 2\pi \). Then

\[ g(\theta, \lambda) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} G_{km} e^{i(2\pi k\theta + m\lambda)}, \quad \theta, \lambda \in \mathbb{R} \]  
(5a)

\[ G_{km} = \frac{1}{2\pi^2} \int_{0}^{\pi} \int_{0}^{2\pi} g(\theta, \lambda) e^{-i(2\pi k\theta + m\lambda)} d\theta d\lambda, \quad k, m \in \mathbb{Z} \]  
(5b)

where, with * denoting the complex conjugate,

\[ G_{km} = G_{km}^* \]  
(5c)

For a regular grid with constant spacings given by \( \Delta \theta = \pi/K \) and \( \Delta \lambda = 2\pi/M \), where \( s = s\Delta \theta \), \( \lambda = t\Delta \lambda \), the Discrete Fourier Transform pair is

\[ g_{s,t} = g(s\Delta \theta, t\Delta \lambda) = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} G_{km} e^{i2\pi(k\theta + m\lambda)/K}, \quad s = 0, ..., K-1, \quad t = 0, ..., M-1 \]  
(6a)

\[ G_{km} = \frac{1}{KM} \sum_{s=0}^{K-1} \sum_{t=0}^{M-1} g_{s,t} e^{-i2\pi(k\theta + m\lambda)/K}, \quad k = 0, ..., K-1, \quad m = 0, ..., M/2 \]  
(6b)

Because the values \( g_{sk} \), are real,

\[ \hat{G}_{km} = G_{km}^* \]  
(6c)

which explains the limit of interest on \( m \). \( M \) is assumed to be even, without loss in generality. By the periodicity in \( k \) and \( m \), \( G_{0,m} = G_{0,M-m} \) and \( G_{k,0} = G_{K-k,0} \), while \( G_{0,0} \) is real. Hence, the complex, discrete spectrum, \( \{G_{km}\} \), contains only \( KM \) independent coefficients. Analogous to (4), the relationship between the total spectrum (5b) and...