Journal of Computational Physics **159,** 213–230 (2000)
doi:10.1006/jcph.2000.6431, available online at http://www.idealibrary.com on IDEAL®

Generalized Discrete Spherical Harmonic Transforms

Paul N. Swarztrauber[1] and William F. Spotz[2]

National Center for Atmospheric Research (sponsored by the National Science Foundation), Boulder, Colorado 80307

Received April 15, 1999; revised November 18, 1999

Two generalizations of the spherical harmonic transforms are provided. First, they are generalized to an arbitrary distribution of latitudinal points θ_i. This unifies transforms for Gaussian and equally spaced distributions and provides transforms for other distributions commonly used to model geophysical phenomena. The discrete associated Legendre functions $\bar{P}_n^m(\theta_i)$ are shown to be orthogonal, to within roundoff error, with respect to a weighted inner product, thus providing the forward transform to spectral space. Second, the representation of the transforms is also generalized to rotations of the discrete basis set $\bar{P}_n^m(\theta_i)$. A discrete function basis is defined that provides an alternative to $\bar{P}_n^m(\theta_i)$. On a grid with N latitudes, the new basis requires $O(N^2)$ memory compared to the usual $O(N^3)$. The resulting transforms differ in spectral space but provide identical results for certain applications. For example, a forward transform followed immediately by a backward transform projects the original discrete function in a manner identical to the existing transforms. Namely, they both project the original function onto the same smooth least squares approximation without the high frequencies induced by the closeness of the points in the neighborhood of the poles. Finally, a faster projection is developed based on the new transforms.

1. INTRODUCTION

Harmonic transforms are used on the sphere in the same way that Fourier transforms are used on a rectangle in Cartesian coordinates. On the rectangle, Fourier transforms are used in both the x and y directions. On the sphere, however, Fourier transforms are used in the longitudinal direction, while Legendre transforms are used in the latitudinal

[1] Supported in part by the DOE and UCAR Cooperative Agreement for the Climate Change Prediction Program under Cooperative Agreement DE-FC03-97ER62402.

[2] Supported by the Division of Mathematical Sciences at the National Science Foundation.

direction. On a grid with N latitudinal points and at least $2N - 2$ longitudinal points, the complete set of spectral coefficients are computed by $2N - 2$ discrete forward Legendre transforms, each corresponding to a rectangular matrix with $N - m$ rows and N columns, where $m = 0, \ldots, N - 1$ are the longitudinal wave numbers. With the exception of $m = 0$ and $N - 1$, two transforms are required for each m.

Except where noted, it is assumed that a complete set of spectral coefficients is computed, which is often called the "linear grid." Computational economies are evident for truncations that require less than a complete set. The backward transform, from spectral to Fourier space, also consists of $2N - 2$ matrices but with N rows and $N - m$ columns. The forward and backward transforms between spectral and Fourier space are called the Legendre analysis and synthesis, respectively. The combined transforms, namely, a forward plus a backward transform, define a projection operator more commonly referred to as the "filter" [5, 11].

The Legendre transforms are not one-to-one because the number of spectral coefficients is about half the number of points on the sphere and consequently, unlike the Fourier transform on a rectangle, a forward followed by a backward transform (projection) does not necessarily reconstruct the original data. Although initially somewhat disconcerting, this property of the combined transforms has been found to be quite useful for time-dependent models of geophysical processes. In particular, a weighted least-squares approximation to the original data is obtained [9] that removes the high frequencies and the resulting time step restriction that are induced by the closeness of the points in the neighborhood of the poles. Projecting the dependent variables in this manner permits the use of a larger time step based on the spacing of the equatorial grid points [5, 7].

This approach also requires fewer harmonic transforms since additional harmonic transforms are not needed to evaluate spatial derivatives. Rather, the derivatives can be evaluated using fast methods based on double Fourier series with model results that are identical to the traditional spherical harmonic spectral method [7]. Or, if spectral accuracy can be relaxed somewhat, the derivatives can be computed by yet faster methods based on high-order finite differences. Although the projection approach provides increased speed, the time required by the projection itself remains $O(N^3)$ and has therefore become the focus of efforts to further speed the computations. A guide to the pseudospectral method itself is given in [3], which includes an application to numerical weather prediction.

The literature contains fast projection methods based on the multipole method [2, 5, 11]. Although these methods have the potential to be $O(N^2 \log N)$, in practice they perform like efficient $O(N^3)$ methods, at least for current and expected values of N [8]. Here we proceed in the latter direction with the development of a faster $O(N^3)$ projection that can halve the number of computations for the Legendre transforms. This variant is also memory-efficient with an $O(N^2)$ memory requirement, compared to the existing $O(N^3)$ requirement. The goal of an $O(N^2 \log N)$ harmonic spectral method remains elusive; however, the "projection" method provides a new avenue of research. Perhaps the development of a fast projection will prove to be easier than the development of a fast harmonic transform.

The Legendre transforms consist of N matrices, which, as currently posed, require memory proportional to N^3. This can be viewed as excessive when compared to the $O(N^2)$ memory requirement of a discrete function on an equiangular grid. Therefore, often the choice is to compute the elements of these matrices at run time rather than precompute and store for repeated use later. This seems reasonable since the induced computational overhead is about 25%, which may be tolerated in exchange for the sizeable memory reduction. Interestingly, accuracy is not necessarily reduced by this choice. From an accuracy point of view it would seem preferable to precompute the transform matrices in 64-bit precision and store them in full 32-bit precision for subsequent use. However, current methods for computing these matrices at run time are extremely accurate [10]. Nevertheless, if precomputation is preferred, the reduced $O(N^2)$ memory requirement further motivates this approach.

The harmonic transforms are generalized to arbitrary θ_i in Section 2. This unifies the harmonic transforms for Gaussian and equally spaced grids. It also provides new transforms for other common distributions such as equally spaced without pole points, which is close to a Guassian distribution, or a shifted equally spaced latitudinal grid that also excludes the pole points. Indeed, the work formally extends to an arbitrary latitudinal distribution but with grid-specific considerations that are discussed in Appendix A.2. The resulting discrete associated Legendre functions are shown to be orthogonal with respect to a weighted inner product, which thereby provides the forward transform into spectral space.

A variant of the Legendre transforms based on functions $Q_n^0(\theta)$ and $Q_n^1(\theta)$ that are linear combinations (rotations) of the associated Legendre functions $\overline{P}_n^m(\theta)$ is developed in Section 3. The resulting Legendre-type transforms differ in spectral space but are identical in physical space, where they give the same results as the usual Legendre transforms. For example, a projection based on the Q functions gives the same results (to roundoff) as one based on the $\overline{P}_n^m(\theta)$. The advantage of the former is a significant reduction in memory from $O(N^3)$ to $O(N^2)$. If used to approximate the derivatives of a discrete function on the sphere, this approach will give the same results as the traditional approach and may be preferable for certain applications where the spectral coefficients can be made invisible to the user.

A faster projection that can reduce the computation required by the Legendre transforms by as much as 50% is developed in Section 4. This would at first seem obvious by simply combining the forward and backward matrices. However, because these matrices are rectangular, such a combination for $m > N/2$ would actually increase the amount of computation. We take an alternate approach in which the projection is represented in terms of the orthogonal complement of the discrete Legendre functions, which becomes the preferred approach for $m < N/2$.

The accuracy of projections based on the traditional and generalized harmonic transforms is compared in Section 5. Both traditional and generalized harmonic transforms are developed for five latitudinal point distributions. Traditional projections are compared to projections based on $O(N^2)$ representations of the generalized discrete Legendre transforms and their orthogonal complements. A summary of results is given in Section 6. Computational methods, theorem proofs, and relevant formulae are provided in the Appendix.

2. GENERALIZED DISCRETE HARMONIC TRANSFORMS

Given the discrete function $f_{i,j}$, defined at latitudes θ_i, $i = 1, \ldots, N$, and longitudes θ_j, $j = 1, \ldots, 2N-2$, the forward harmonic transform or harmonic analysis consists of determining coefficients $a_{m,n}$ and $b_{m,n}$ such that $f_{i,j}$ can be synthesized by

$$f_{i,j} = \sum_{n=0}^{N-1} \sum_{m=0}^{n} \overline{P}_n^m(\theta_i)(a_{m,n} \cos m\phi_j + b_{m,n} \sin m\phi_j). \tag{2.1}$$

The analysis consists of two phases. First we use the fast Fourier transform (FFT) to compute

$$a_m(\theta_i) = \frac{1}{2N-2}\sum_{j=1}^{2N-2} f_{i,j}\cos m\phi_j \quad \text{and} \quad b_m(\theta_i) = \frac{1}{2N-2}\sum_{j=1}^{2N-2} f_{i,j}\sin m\phi_j. \quad (2.2)$$

If the θ_i are Gaussian distributed then the desired coefficients are given in the second phase by

$$a_{m,n} = \sum_{i=1}^{N} w_i a_m(\theta_i)\bar{P}_n^m(\theta_i) \quad \text{and} \quad b_{m,n} = \sum_{i=1}^{N} w_i b_m(\theta_i)\bar{P}_n^m(\theta_i), \quad (2.3)$$

where the w_i are the Gaussian weights. Once the $a_{m,n}$ and $b_{m,n}$ are determined, the harmonic synthesis or backward transform is given by (2.1). Here we focus on the computationally intensive transforms (2.3) between Fourier and harmonic spaces because the Fourier transforms (2.2) are fast and not relevant to the work presented here.

As mentioned in the Introduction, the forward harmonic transform followed immediately by a backward transform will not, in general, reconstruct the discrete function but rather will provide a least-squares approximation to $f_{i,j}$ that is quite useful, for the reasons stated earlier. We begin with a study of this combination, called the Legendre projection, that will later yield the generalized harmonic transforms, which include both Gauss and equally spaced grids as well as other common grids used in computational geophysics. Indeed, the work generalizes to any latitudinal distribution with considerations that are grid-specific and discussed later.

The projection occurs in the Legendre transforms between Fourier and harmonic space. The computation of both $a_{m,n}$ and $b_{m,n}$ in (2.3) is by application of the matrix operator $\mathbf{P}_m^T\mathbf{W}$, where $\mathbf{W}$ is an $N \times N$ diagonal matrix of Gaussian weights w_i and $\mathbf{P}_m$ is the $N \times (N-m)$ matrix

$$\mathbf{P}_m = \begin{bmatrix} \bar{P}_m^m(\theta_1) & \cdots & \bar{P}_{N-1}^m(\theta_1) \\ \vdots & & \vdots \\ \bar{P}_m^m(\theta_N) & \cdots & \bar{P}_{N-1}^m(\theta_N) \end{bmatrix}, \quad (2.4)$$

whose entries $\bar{P}_n^m(\theta_i)$ are tabulations of the normalized associated Legendre functions

$$\bar{P}_n^m(\theta) = \frac{1}{2^n n!}\left[\frac{2n+1}{2}\frac{(n-m)!}{(n+m)!}\right]^{1/2}\cos^m\theta\,\frac{d^{n+m}}{dx^{n+m}}(x^2-1)^n; \quad x = \sin\theta. \quad (2.5)$$

The Legendre projection combines the analysis and synthesis,

$$\mathbf{F}_m = \mathbf{P}_m\mathbf{P}_m^T\mathbf{W}, \quad (2.6)$$

into a single $N \times N$ matrix for each longitudinal wave number m. The discrete Legendre functions are orthogonal with respect to Gaussian quadrature. That is, $\mathbf{P}_m^T\mathbf{W}\,\mathbf{P}_m = \mathbf{I}_{N-m\times N-m}$. Then, because a matrix commutes with its inverse, $\mathbf{P}_0\mathbf{P}_0^T\mathbf{W} = \mathbf{I}_{N\times N}$ and therefore

$$\mathbf{W}_{N\times N} = \left(\mathbf{P}_0\mathbf{P}_0^T\right)^{-1}, \quad (2.7)$$

and, from (2.6), $\mathbf{F}_0 = \mathbf{I}_{N\times N}$. For exposition the dimensions of a matrix may be included with its definition.

Unless otherwise noted, weighted orthogonality is assumed throughout. $\mathbf{W}$ defines a weighted inner product $(\mathbf{u},\mathbf{v})_{\mathbf{W}} = (\mathbf{u}^T\mathbf{W}\mathbf{v})$ that is fundamental to the generalized harmonic transforms presented here. Of course it is possible to select a distribution of points that yields a singular or near-singular matrix on the right side of (2.7). The management of this potential problem is discussed in Appendix A.2.

Consider now the elements of $\mathbf{P}_0\mathbf{P}_0^T$ but for an arbitrary distribution of latitudes θ_i,

$$\left(\mathbf{P}_0\mathbf{P}_0^T\right)_{i,j} = \sum_{k=0}^{N-1} \bar{P}_k^0(\theta_i)\bar{P}_k^0(\theta_j). \tag{2.8}$$

A simple proof of the Christoffel–Darboux formula is given in Appendix A.3, which provides the following alternate representation of the elements:

$$\left(\mathbf{P}_0\mathbf{P}_0^T\right)_{i,j} = \frac{N}{\sqrt{4N^2-1}}\frac{\bar{P}_N^0(\theta_i)\bar{P}_{N-1}^0(\theta_j) - \bar{P}_{N-1}^0(\theta_i)\bar{P}_N^0(\theta_j)}{\sin\theta_i - \sin\theta_j}. \tag{2.9}$$

The diagonal elements can be computed either from (2.8) or by applying l'Hôpital's rule to (2.9). Using the latter approach,

$$\left(\mathbf{P}_0\mathbf{P}_0^T\right)_{i,i} = \frac{N}{\sqrt{4N^2-1}\cos\theta_i}\left[\bar{P}_{N-1}^0(\theta_i)\frac{d}{d\theta}\bar{P}_N^0(\theta_i) - \bar{P}_N^0(\theta_i)\frac{d}{d\theta}\bar{P}_{N-1}^0(\theta_i)\right]. \tag{2.10}$$

From (2.9) we are motivated to select θ_i as the zeros of $\bar{P}_N^0(\theta_i)$ because the resulting matrix is then diagonal. The resulting θ_i are known as the Gaussian distribution and $(\mathbf{W})_{i,i}$ are the Gaussian weights obtained from (2.10) as

$$(\mathbf{W})_{i,i} = \sqrt{4N^2-1}\cos\theta_i \Big/ N\bar{P}_{N-1}^0(\theta_i)\frac{d}{d\theta}\bar{P}_N^0(\theta_i). \tag{2.11}$$

The Gaussian points θ_i can be computed accurately and efficiently as the eigenvalues of a symmetric tridiagonal matrix as described in [10]. The weights can be computed from (2.8) or (2.9) or by a third method also described in [10] and implemented in subroutine `gaqd` in `SPHEREPACK` [1], at (A.5.1). `SPHEREPACK` contains programs for the harmonic transforms on Gaussian and equally spaced grids as well as a number of other harmonic transforms and related computations that can assist model development.

Although the spherical harmonics are not polynomials for odd m, the product of two such harmonics is a polynomial for which Gauss quadrature is exact. Therefore, for all $m \le N-1$ the spherical harmonics $\mathbf{P}_m$ are weighted orthogonal with respect to the Gauss weights. However, unlike the Gaussian distribution, other distributions require two weight matrices, namely, $\mathbf{W}_0 = (\mathbf{P}_0\mathbf{P}_0^T)^{-1}$ and $\mathbf{W}_1 = (\tilde{\mathbf{P}}_1\tilde{\mathbf{P}}_1^T)^{-1}$, for m even and odd, respectively. The tilde notation is used because $\mathbf{P}_1$ must be augmented to an independent set of N vectors before inversion is possible. Indeed, for an arbitrary distribution, $\mathbf{P}_0$ may also be singular as discussed in Appendix A.2.

THEOREM 1. *Let* $\mathbf{P}_m$ *be the* $N\times(N-m)$ *matrix defined by* (2.4) *and let* $\mathbf{W}_l$ $(l = 0, 1)$ *be the* $N\times N$ *matrices defined in the paragraph preceeding Theorem* 1; *then for all* $0 < m <$

$N-1$ and any distribution θ_i such that the corresponding weight matrices $\mathbf{W}_0$ and $\mathbf{W}_1$ exist, the associated Legendre functions $\mathbf{P}_m$ are weighted orthogonal. That is,

$$\mathbf{P}_m^T \mathbf{W}_l \mathbf{P}_m = \mathbf{I}_{(N-m)\times(N-m)} \quad l = \begin{cases} 0, & m \text{ even} \\ 1, & m \text{ odd}. \end{cases} \tag{2.12}$$

Most of the proof is in Appendix A.4, where it is shown that the four-point recurrence for the associated Legendre functions corresponds to an orthogonal transform. That is, $\mathbf{P}_{m+1} = \mathbf{P}_{m-1}\mathbf{B}_m$, where $\mathbf{B}_m = (\mathbf{X}_m^{-1}\mathbf{Y}_m)^T$ is $(N-m+1)\times(N-m-1)$ and $\mathbf{B}_m^T\mathbf{B}_m = \mathbf{I}_{(N-m-1)\times(N-m-1)}$. $\mathbf{X}_m$ and $\mathbf{Y}_m$ are defined in Appendix A.4 at (A.4.3). Let $\mathbf{M}_m$ be the $N\times(N-m)$ matrix $\mathbf{B}_{l+1}\mathbf{B}_{l+3}\cdots\mathbf{B}_{m-1}$. Then $\mathbf{M}_m^T\mathbf{M}_m = \mathbf{I}_{(N-m)\times(N-m)}$ and $\mathbf{P}_m = \mathbf{P}_l\mathbf{M}_m$. Theorem 1 follows because, by construction, $\mathbf{P}_l$ is weighted orthogonal with respect to $\mathbf{W}_l$.

The forward transform to spectral space, or the Legendre analysis, is then immediately evident as

$$\mathbf{Z}_m^T = \mathbf{P}_m^T \mathbf{W}_l, \tag{2.13}$$

where l is 0 (1) if m is even (odd). The backward transform or Legendre synthesis is simply $\mathbf{P}_m$. Like (2.6), the projection combines the two and from (2.12)

$$\mathbf{F}_m^2 = \mathbf{F}_m, \tag{2.14}$$

which demonstrates that $\mathbf{F}_m$ is a projection operator onto the discrete associated Legendre functions $\mathbf{P}_m$. This result is the key to the stability of the spectral transform method and remains an attribute of the generalized projections that are developed here.

For arbitrary θ_i, $\mathbf{W}_l$ is not diagonal; however, when it is combined with $\mathbf{P}_m^T$ as in (2.13), the compute time for the forward transform is the same as that required for a Gaussian distribution of points. For equally spaced θ_i, the resulting Legendre analysis is identical to that given in [6, 9].

If the matrices are stored, the harmonic transforms require double the memory required by transforms on a Gaussian grid, because both $\mathbf{P}_m$ and $\mathbf{Z}_m$ must be stored. However, if they are computed at run time, the memory requirements for both the Gaussian and equally spaced grids are $O(N^2)$ rather than $O(N^3)$. This is often the preferred approach since the elements in both matrices can be computed efficiently with three multiplications and two additions, as described in Appendix A.4 and implemented in `SPHEREPACK` (A.5.1). However, this computation can be eliminated, while at the same time retaining $O(N^2)$ storage, by using a variant of the Legendre transforms that is developed in the next section.

3. A VARIANT OF THE LEGENDRE TRANSFORMS WITH $O(N^2)$ REPRESENTATION

We begin with the construction of the orthogonal complement of $\mathbf{P}_m^T\mathbf{W}$. That is, for each m we will determine m orthogonal vectors $\mathbf{q}$ such that $\mathbf{P}_m^T\mathbf{W}\mathbf{q} = 0$. The memory requirement for the complete set is $O(N^2)$ compared with $O(N^3)$ for $\mathbf{P}_m$. In what follows, a discrete function will be called an even (odd) vector if it is even (odd) about the equator ($\theta = 0$). For a distribution of latitudinal points that is symmetric about the equator, the even (odd) classification can be used to halve the computational time as demonstrated below. The development will be presented by example for the case $N = 8$ with multiple references to

TABLE I
Bases for the Orthogonal Complement of $\mathbf{P}_m^T\mathbf{W}$ on an Eight-Point Latitudinal Distribution

$\mathbf{p}_0^0$	$\mathbf{p}_1^0$	$\mathbf{p}_2^0$	$\mathbf{p}_3^0$	$\mathbf{p}_4^0$	$\mathbf{p}_5^0$	$\mathbf{p}_6^0$	$\mathbf{p}_7^0$
$\mathbf{q}_0^1$	$\mathbf{p}_1^1$	$\mathbf{p}_2^1$	$\mathbf{p}_3^1$	$\mathbf{p}_4^1$	$\mathbf{p}_5^1$	$\mathbf{p}_6^1$	$\mathbf{p}_7^1$
$\mathbf{q}_0^0$	$\mathbf{q}_1^0$	$\mathbf{p}_2^2$	$\mathbf{p}_3^2$	$\mathbf{p}_4^2$	$\mathbf{p}_5^2$	$\mathbf{p}_6^2$	$\mathbf{p}_7^2$
$\mathbf{q}_0^1$	$\mathbf{q}_1^1$	$\mathbf{q}_2^1$	$\mathbf{p}_3^3$	$\mathbf{p}_4^3$	$\mathbf{p}_5^3$	$\mathbf{p}_6^3$	$\mathbf{p}_7^3$
$\mathbf{q}_0^0$	$\mathbf{q}_1^0$	$\mathbf{q}_2^0$	$\mathbf{q}_3^0$	$\mathbf{p}_4^4$	$\mathbf{p}_5^4$	$\mathbf{p}_6^4$	$\mathbf{p}_7^4$
$\mathbf{q}_0^1$	$\mathbf{q}_1^1$	$\mathbf{q}_2^1$	$\mathbf{q}_3^1$	$\mathbf{q}_4^1$	$\mathbf{p}_5^5$	$\mathbf{p}_6^5$	$\mathbf{p}_7^5$
$\mathbf{q}_0^0$	$\mathbf{q}_1^0$	$\mathbf{q}_2^0$	$\mathbf{q}_3^0$	$\mathbf{q}_4^0$	$\mathbf{q}_5^0$	$\mathbf{p}_6^6$	$\mathbf{p}_7^6$
$\mathbf{q}_0^1$	$\mathbf{q}_1^1$	$\mathbf{q}_2^1$	$\mathbf{q}_3^1$	$\mathbf{q}_4^1$	$\mathbf{q}_5^1$	$\mathbf{q}_6^1$	$\mathbf{p}_7^7$

Table I. The orthogonal complement is spanned by the discrete functions $\mathbf{q}_n^0$ and $\mathbf{q}_n^1$ in the last two rows of Table I. They are computed in the following sequence using the Gram–Schmidt orthogonalization with respect to the weighted inner product $(\mathbf{u}, \mathbf{v})_{\mathbf{W}}$. The vectors $\mathbf{p}_n^m$ are the discrete Legendre functions or the columns of $\mathbf{P}_m$ defined in (2.4).

1. In the second row of Table I, $\mathbf{q}_0^1$ is uniquely determined as an odd vector that is orthogonal to $\mathbf{p}_2^1$, $\mathbf{p}_4^1$, and $\mathbf{p}_6^1$. It is implicitly orthogonal to the remaining $\mathbf{p}_n^1$ for $n = 1, 3, 5$, and 7 because they are even vectors. It is also orthogonal to the remaining $\mathbf{p}_n^m$ in Table I with odd m because they are linear combinations of $\mathbf{p}_n^1$ as discussed in Appendix A.4. Therefore $\mathbf{q}_0^1$ can also be selected as a member of the orthogonal complement to $\mathbf{P}_m^T\mathbf{W}$ for $m = 3, 5$, and 7, where it appears in the corresponding rows of the first column in Table I.
2. In the third row, $\mathbf{q}_0^0$ is computed as orthogonal to $\mathbf{p}_n^2$ for $n = 2, 4$, and 6. Also $\mathbf{q}_1^0$ is computed as orthogonal to $\mathbf{p}_n^2$ for $n = 3, 5$, and 7. For the reason stated in 1 above, both are also orthogonal to the remaining $\mathbf{p}_n^m$ for $m = 4$ and 6 where they appear in the first two columns of Table I.
3. This process continues down Table I with two new vectors being added to each row. For example, in the fifth row the vectors $\mathbf{q}_2^0$ and $\mathbf{q}_3^0$ are added but with the requirement that they are orthogonal to all other vectors in the row with the same parity. Therefore, $\mathbf{q}_2^0$ must be computed not only as orthogonal to $\mathbf{p}_4^4$ and $\mathbf{p}_6^4$ but also as orthogonal to $\mathbf{q}_0^0$.

This completes the construction of the orthogonal complement, from which a variant to the Legendre functions can be developed that requires only $O(N^2)$ memory. The variant is determined from the last two rows in Table I, which provide both the analysis and synthesis as presented in the following two theorems.

First define $\mathbf{q}_6^0 = \mathbf{p}_6^6$, $\mathbf{q}_7^0 = \mathbf{p}_7^6$, and $\mathbf{q}_7^1 = \mathbf{p}_7^1$. Next assume general N and define the following matrices that are formed from vectors in the last two rows of Table I,

$$\mathbf{Q}_m^l = \begin{bmatrix}\mathbf{q}_m^l & \cdots & \mathbf{q}_{N-1}^l\end{bmatrix}_{N\times(N-m)}, \tag{3.1}$$

where $l = 0$ (1) if m is even (odd).

THEOREM 2. *There exists an $(N-m)\times(N-m)$ l_2 orthonormal matrix $\mathbf{H}_m^l$ such that the Legendre synthesis $\mathbf{P}_m$ in (2.4) is given by $\mathbf{P}_m = \mathbf{Q}_m^l\mathbf{H}_m^l$.*

The proof, together with the definition of $\mathbf{H}_m^l$, is given in Appendix A.1.

Because $(\mathbf{H}_m^l)^T\mathbf{H}_m^l = \mathbf{I}_{(N-m)\times(N-m)}$ we have the corollary that $\mathbf{Q}_m^l = \mathbf{P}_m(\mathbf{H}_m^l)^T$ and therefore we can now interpret the discrete functions $\mathbf{q}_n^l$ as continuously differentiable functions

$Q_n^l(\theta)$. Like the associated Legendre functions, the $Q_n^l(\theta)$ can be differentiated, integrated, or interpolated in any manner that may be required by an application.

Next let $\mathbf{r}_n^l$ be the rows of $(\mathbf{Q}_0^l)^{-1}$ and define the last $N-m$ rows by

$$\mathbf{R}_m^l = \begin{bmatrix} \mathbf{r}_m^l & \cdots & \mathbf{r}_{N-1}^l \end{bmatrix}_{N\times(N-m)}. \tag{3.2}$$

THEOREM 3. *The Legendre analysis* (2.13) *is given by* $(\mathbf{Z}_m^l)^T = (\mathbf{H}_m^l)^T(\mathbf{R}_m^l)^T$.

The proof, together with the definition of $\mathbf{H}_m^l$, is given in Appendix A.1. For all m, $\mathbf{H}_m^l$ requires $O(N^3)$ memory; however, it may not be relevant for applications that do not require explicit calculation of the spectral coefficients. For example, the projection $\mathbf{F}_m$ in (2.6), which consists of an analysis followed immediately by a synthesis, has a representation as the $O(N^2)$ projection

$$\mathbf{F}_m = \mathbf{Q}_m^l\left(\mathbf{R}_m^l\right)^T, \tag{3.3}$$

where $\mathbf{Q}_m^l$ and $\mathbf{R}_m^l$ are both $N\times(N-m)$ matrices given by (3.1) and (3.2) respectively. Although m ranges from 0 to $N-1$, they require only $2N^2$ locations because they are all generated from $\mathbf{Q}_l^l$ and $\mathbf{R}_l^l$, $l=0,1$, by deleting m columns. Both $\mathbf{Q}_m^l$ and $\mathbf{R}_m^l$ require $2N^2$ locations and hence the projection as well as the Legendre type transforms require $4N^2$ locations. This can be halved taking advantage of symmetries.

Next suppose we wish to approximate the θ derivative of a discrete function defined on the surface of the sphere. The traditional approach is to first analyze and then synthesize but with $\mathbf{P}_m$ replaced by its derivative $\dot{\mathbf{P}}_m$. By combining Theorems 2 and 3 and replacing Q_m^l with $\dot{\mathbf{Q}}_m^l$ we determine that an approximate latitudinal derivative can be obtained by the application of $\dot{\mathbf{Q}}_m^l(\mathbf{R}_m^l)^T$. Like the projection, this computation is independent of $\mathbf{H}_m^l$. $\dot{\mathbf{Q}}_m^l$ can be precomputed and stored in $2N^2$ locations for subsequent use.

4. A FASTER PROJECTION

In Section 2 we observed that $\mathbf{F}_0=\mathbf{I}_{N\times N}$, which considerably facilitates its application for $m=0$. However, for $m>0$, $\mathbf{F}_m$ is a full $N\times N$ matrix that does not admit an obvious computational saving. However, the rank of $\mathbf{F}_m$ is $N-m$, and hence the rank of its orthogonal complement is m, which can be used to speed its computation for $m<N/2$. We proceed now to develop an alternate form of $\mathbf{F}_m$.

THEOREM 4. *The projection* $\mathbf{F}_m$ *has the alternate form* $\mathbf{F}_m=\mathbf{I}_{N\times N}-\mathbf{G}_m$, *where* $\mathbf{G}_m$ *is equal to the product of the first m columns of* $\mathbf{Q}_0^l$ *times the first m rows of* $\mathbf{R}_0^l$.

First define the orthogonal complement of $\mathbf{Q}_m$ and $\mathbf{R}_m$ as

$$\bar{\mathbf{Q}}_m^l = \begin{bmatrix} \mathbf{q}_0^l & \cdots & \mathbf{q}_{m-1}^l \end{bmatrix}_{N\times m} \quad \text{and} \quad \bar{\mathbf{R}}_m^l = \begin{bmatrix} \mathbf{r}_0^l & \cdots & \mathbf{r}_{m-1}^l \end{bmatrix}_{N\times m}. \tag{4.1}$$

Then, from (3.2), (3.3), and by inspection

$$\mathbf{I}_{N\times N} = \mathbf{F}_0 = \mathbf{Q}_0^l\left(\mathbf{R}_0^l\right)^T = \mathbf{Q}_m^l\left(\mathbf{R}_m^l\right)^T + \bar{\mathbf{Q}}_m^l\left(\bar{\mathbf{R}}_m^l\right)^T, \tag{4.2}$$

or

$$\mathbf{F}_m = \mathbf{I} - \mathbf{G}_m, \tag{4.3}$$

where $\mathbf{G}_m = \bar{\mathbf{Q}}_m^l(\bar{\mathbf{R}}_m^l)^T$, which completes the proof.

This provides an alternate way of applying the projection $\mathbf{F}_m$; namely, in terms of its orthogonal complement. For example, consider the case $m = 1$, for which $\mathbf{F}_1$ is the product of $\mathbf{Q}_1$ and $\mathbf{R}_1^T$, which are almost full matrices. However, $\mathbf{G}_1$ is the outer product of the vectors $\mathbf{q}_0^1$ and $\mathbf{r}_0^1$. Assume the latter is precomputed, then for arbitrary $\mathbf{x}$, $(\mathbf{I} - \mathbf{G}_1)\mathbf{x}$ requires $2N$ flops compared to $2N^2$ flops for computing $\mathbf{F}_1\,\mathbf{x}$. The alternate $(\mathbf{I} - \mathbf{G}_m)\mathbf{x}$ is more efficient for $m < N/2$ and $\mathbf{F}_m\mathbf{x}$ is preferred for $m \geq N/2$. This alternate approach is also slightly more accurate, which can be determined by comparing Tables II and III in the next section.

This alternative provides a computational savings of 50% if all harmonics are included in the projection. The savings is less when fewer harmonics are included. For example, if two-thirds of the harmonics are included, then the dimension of the orthogonal complement increases with crossover at at $m = N/6$. The overall savings when using the two-thirds rule is about 12.5%. The two-thirds rule can be implemented either by truncating the discrete functions in $\mathbf{P}_m$ or by enlarging its orthogonal complement—as an example of the latter, by including the vectors in the last two columns of Table I.

5. COMPUTATIONAL EXPERIMENTS

The developments of the previous sections have been implemented in a single program, which accepts any latitudinal distribution θ_i and computes (a) the weight matrices $\mathbf{W}_l$ as described in Appendix A.3, (b) the orthogonal complement $\mathbf{Q}_m^l$, (c) the weight matrix (A.1.9) in terms of the memory efficient alternative Legendre functions $\mathbf{q}_n^l$, and (d) projections based on both the traditional Legendre functions $\mathbf{p}_n^m$ and alternatives $\mathbf{q}_n^l$. The accuracy of these projections is listed in Tables II and III for five different latitudinal distributions.

Table II contains the maximum error of the projection based on the application of $\mathbf{P}_m\mathbf{P}_m^T\mathbf{W}$ to 20 random vectors. The implementation is sequential with the vector being first multiplied by the $m \times N$ analysis matrix $\mathbf{P}_m^T\mathbf{W}$ followed by the $N \times m$ synthesis matrix $\mathbf{P}_m$. Accuracy is computed as the maximum difference between the calculations in 32- and 64-bit precision.

Table III contains the accuracy of the same projection based on the memory-efficient alternatives to the discrete Legendre functions and their orthogonal complements. That is, for $m \leq N/2$, the projection is computed from (4.3), and for $m > N/2$, the projection is computed from (3.3). In all cases the elements of the matrices are computed in 64-bit precision but stored in 32-bit precision before application to the same random vectors used in Table II. The accuracy of the alternate approach, as given in Table III is computed as its maximum difference from the traditional approach in 64-bit precision. The accuracy of the alternate approach appears slightly better.

TABLE II

Error in Computing the Legendre Projection Based on Traditional Legendre Functions $\mathbf{p}_n^m$ for Five Different N-Point Latitudinal Distributions

Distribution	$N = 32$	$N = 64$	$N = 128$
Equally spaced with poles	6.56×10^{-7}	8.34×10^{-7}	2.00×10^{-6}
Equally spaced without poles	6.56×10^{-7}	8.34×10^{-7}	1.31×10^{-6}
Shifted equally spaced	5.66×10^{-7}	1.07×10^{-6}	1.31×10^{-6}
Gauss distributed	4.77×10^{-7}	5.96×10^{-7}	1.54×10^{-6}
Random distribution	8.34×10^{-7}	1.25×10^{-6}	1.88×10^{-6}

TABLE III
Error in Computing the Legendre Projection Based on the Alternative Legendre Functions $\mathbf{q}_n^0$ and $\mathbf{q}_n^1$ for Five Different *N*-Point Latitudinal Distributions

Distribution	$N = 32$	$N = 64$	$N = 128$
Equally spaced with poles	3.57×10^{-7}	4.77×10^{-7}	6.56×10^{-7}
Equally spaced without poles	3.58×10^{-7}	6.56×10^{-7}	1.01×10^{-6}
Shifted equally spaced	3.87×10^{-7}	5.36×10^{-7}	7.15×10^{-7}
Gauss distributed	2.98×10^{-7}	5.07×10^{-7}	7.15×10^{-7}
Random distribution	3.57×10^{-7}	5.36×10^{-7}	7.74×10^{-7}

For each point distribution, a discrete harmonic analysis and synthesis were determined with characteristics identical to the traditional Gaussian or equally spaced transforms. That is, in each case the discrete functions $\mathbf{p}_n^m$ or $\mathbf{q}_n^l$ were orthonormal under the computed weight function and $\mathbf{F}_m$ defined a projection operator to machine roundoff error. The five different latitudinal distributions are listed below.

1. An equally spaced distribution of points that includes the pole points. That is, $\theta_i = \pi/2 - (i-1)\pi/(N-1)$, for $i = 1, \ldots, N$. The resulting transforms are identical to the existing transforms in [6, 9].
2. An equally spaced distribution of points, like that given above but excluding the pole points. That is, $\theta_i = \pi/2 - i\pi/(N+1)$, for $i = 1, \ldots, N$. This distribution is close to the Gaussian distribution in 4 below.
3. A shifted equally spaced distribution like 1 above but with the first and last points at a distance $\pi/(2N)$ from the poles. That is, $\theta_i = \pi/2 - (i - .5)\pi/N$, for $i = 1, \ldots, N$. The pole points are not included.
4. A Gauss distribution of N points. The resulting transforms are identical to the traditional transforms based on Gaussian quadrature.
5. Like 1 above, but with a 10% random perturbation of the nonpole points. Here we are not able to take advantage of the parity of the vector functions, which doubles the time required to compute the orthogonal complement $\mathbf{q}_n^l$.

6. SUMMARY

Here we first summarize the results of the preceding sections

1. The Legendre transforms are generalized to an arbitrary latitudinal distribution of points, thereby unifying the transforms based on Gauss and equally spaced distribution as well as providing new transforms for other grid distributions used to model geophysical processes. The resulting discrete Legendre functions are orthogonal with respect to a weighted inner product and define projection operators.
2. Memory efficient alternative Legendre transforms are developed whose coefficients in spectral space are rotations of the traditional spectral coefficients. These transforms require $O(N^2)$ memory compared to the traditional $O(N^3)$ requirement. They provide identical results for computations that do not require the explicit computation of the traditional spectral coefficients. Examples include the computation of the Legendre projection or latitudinal derivatives.

3. The Legendre projection consists of a forward transform followed immediately by a backward transform and therefore the spectral coefficients are not explicitly required. A faster projection is developed based on the alternative Legendre transforms and their orthogonal complement. A computational savings of up to 50% can be realized.

4. The accuracy of projections based on the traditional and alternative discrete Legendre functions are compared in Tables II and III. The projections based on the memory-efficient alternative to the Legendre transforms are slightly more accurate than the traditional projection. Therefore, like the traditional projection, the alternate projection projects any discrete function onto a smooth least-squares approximation in a manner identical to the existing harmonic projection.

In [7] it was shown that the stability and accuracy of the spectral transform method was determined by the implicit application of the harmonic projection. That is, it was determined that spectral transform accuracy and stability can be obtained by projecting the dependent variables only and using (say) double Fourier series to compute spatial derivatives. This resulted in a computational savings, since fewer Legendre transforms are required. It also focused attention on the projection as the most computation-intensive part of the model dynamics and therefore as having a significant potential to provide additional savings. Here that potential has been realized by doubling its speed, which makes the cost of the projection comparable to that of a single Legendre transform. In turn, this further increases the speed of the projection method to between double and triple that of the traditional spectral transform method, while at the same time reducing the memory requirement from $O(N^3)$ to $O(N^2)$.

APPENDIX

Proofs, Formulae, and Computational Methods

A.1. *Proofs of Theorems 2 and 3 in Section 3*

In this section we prove Theorems 2 and 3 in Section 3 that provide a memory efficient alternative to the traditional Legendre transforms. In particular we derive an orthonormal matrix $\mathbf{H}_m$ that rotates the coefficients in spectral space in such a way that both the resulting analysis and synthesis have an $O(N^2)$ representation. Alternatively, we define a variant of the Legendre analysis, with $O(N^2)$ representation, that computes spectral coefficients that are a rotation of the coefficients determined by the usual Legendre analysis. The resulting transforms differ in spectral space but are identical in physical space and may therefore be attractive for many applications, including weather and climate modeling, where spectral space can be made invisible to the user. The presentation is by example; for $N = 8$, it will be shown that all of the Legendre transforms and projections can be expressed in terms of the last two rows in Table I.

We begin by demonstrating that $\mathbf{P}_4$, as defined in (2.4) and (3.1), is a rotation of

$$\mathbf{Q}_4^0 = \begin{bmatrix} \mathbf{q}_4^0 & \mathbf{q}_5^0 & \mathbf{p}_6^6 & \mathbf{p}_7^6 \end{bmatrix}_{8\times 4}, \tag{A.1.1}$$

which is located in the next to last row of Table I. To this end we define the matrix

$$\mathbf{D} = \begin{bmatrix} \bar{\mathbf{Q}}_4^0 & \mathbf{Q}_4^0 \end{bmatrix}_{8\times 8}, \tag{A.1.2}$$

where $\bar{\mathbf{Q}}_4^0$ is defined in (4.1). Therefore $\mathbf{D}$ corresponds to the matrix with columns defined by the next to last row of Table I. Because $\mathbf{D}$ is weighted orthogonal,

$$\mathbf{I}_{8\times 8} = \mathbf{D}^T\mathbf{W}_0\mathbf{D} = \mathbf{W}_0\mathbf{D}\mathbf{D}^T = \mathbf{D}\mathbf{D}^T\mathbf{W}_0. \tag{A.1.3}$$

Then, by inspection,

$$\mathbf{I}_{8\times 8} = \bar{\mathbf{Q}}_4^0\left(\bar{\mathbf{Q}}_4^0\right)^T\mathbf{W}_0 + \mathbf{Q}_4^0\left(\mathbf{Q}_4^0\right)^T\mathbf{W}_0. \tag{A.1.4}$$

A similar development for the fifth row of Table 1 yields

$$\mathbf{I}_{8\times 8} = \bar{\mathbf{Q}}_4^0\left(\bar{\mathbf{Q}}_4^0\right)^T\mathbf{W}_0 + \mathbf{P}_4^0\left(\mathbf{P}_4^0\right)^T\mathbf{W}_0. \tag{A.1.5}$$

Subtracting (A.1.5) from (A.1.4), we obtain

$$\mathbf{Q}_4^0\left(\mathbf{Q}_4^0\right)^T\mathbf{W}_0 = \mathbf{P}_4\mathbf{P}_4^T\mathbf{W}_0. \tag{A.1.6}$$

Recalling that the columns of both $\mathbf{Q}_4^0$ and $\mathbf{P}_4$ are orthogonal with respect to $\mathbf{W}_0$, we obtain

$$\mathbf{P}_4 = \mathbf{Q}_4^0\mathbf{H}_4^0 \quad \text{and} \quad \mathbf{Q}_4^0 = \mathbf{P}_4\left(\mathbf{H}_4^0\right)^T, \tag{A.1.7}$$

where $\mathbf{H}_4^0 = (\mathbf{Q}_4^0)^T\mathbf{W}_0\mathbf{P}_4$ is l_2 orthogonal, i.e., $(\mathbf{H}_4^0)^T\mathbf{H}_4^0 = \mathbf{I}_{4\times 4}$. Therefore $\mathbf{P}_4$ can be represented as a rotation of the last four discrete functions in the next to last row of Table I.

The generalization is now clear; for even m, $\mathbf{P}_m$ can be represented as a rotation of the $N-m$ discrete functions $\mathbf{q}_n^0$, $n = m, \dots, N$. For odd m, $\mathbf{P}_m$ can be represented as a rotation of the $N-m$ discrete functions $\mathbf{q}_n^1$, $n = m, \dots, N$ in the last row of Table I. If N is odd the last two rows must be switched. In general,

$$\mathbf{P}_m = \mathbf{Q}_m^l\mathbf{H}_m^l \quad \text{where} \quad \mathbf{H}_m^l = \left(\mathbf{Q}_m^l\right)^T\mathbf{W}_l\mathbf{P}_m, \tag{A.1.8}$$

where l is 0 or 1 if m is even or odd, respectively. $\mathbf{W}_l$ can also be expressed in terms of the discrete functions $\mathbf{q}_n^l$. From (2.7) and (A.1.8),

$$\mathbf{W}_l = \left[\mathbf{Q}_0^l\left(\mathbf{Q}_0^l\right)^T\right]_{N\times N}^{-1}, \tag{A.1.9}$$

where for even N

$$\mathbf{Q}_0^0 = \left[\mathbf{q}_0^0 \quad \mathbf{q}_1^0 \cdots \mathbf{q}_{N-3}^0 \quad \mathbf{p}_{N-2}^{N-2} \quad \mathbf{p}_{N-1}^{N-2}\right] \tag{A.1.10}$$

$$\mathbf{Q}_0^1 = \left[\mathbf{q}_0^1 \quad \mathbf{q}_1^1 \cdots \mathbf{q}_{N-3}^1 \quad \mathbf{q}_{N-2}^1 \quad \mathbf{p}_{N-1}^{N-1}\right] \tag{A.1.11}$$

and for odd N

$$\mathbf{Q}_0^0 = \left[\mathbf{q}_0^0 \quad \mathbf{q}_1^0 \cdots \mathbf{q}_{N-3}^0 \quad \mathbf{q}_{N-2}^0 \quad \mathbf{p}_{N-1}^{N-1}\right] \tag{A.1.12}$$

$$\mathbf{Q}_0^1 = \left[\mathbf{q}_0^1 \quad \mathbf{q}_1^1 \cdots \mathbf{q}_{N-3}^1 \quad \mathbf{p}_{N-2}^{N-2} \quad \mathbf{p}_{N-1}^{N-2}\right]. \tag{A.1.13}$$

This completes the proof of Theorem 2. Consider now the following proof of Theorem 3 beginning with the $O(N^2)$ representations of the Legendre transforms. From (2.13), (A.1.8), and (A.1.9) the traditional forward transform or Legendre analysis is given by

$$\left(\mathbf{Z}_m^l\right)^T = \mathbf{P}_m^T \mathbf{W}_l = \left(\mathbf{H}_m^l\right)^T \left(\mathbf{Q}_m^l\right)^T \left[\left(\mathbf{Q}_0^l\right)^T\right]^{-1} \left(\mathbf{Q}_0^l\right)^{-1} \tag{A.1.14}$$

or

$$\left(\mathbf{Z}_m^l\right)^T = \left(\mathbf{H}_m^l\right)^T [\mathbf{O} \quad \mathbf{I}]_{(N-m)\times N} \left(\mathbf{Q}_0^l\right)^{-1}, \tag{A.1.15}$$

where $\mathbf{O}$ is an $(N-m)\times(N-m)$ matrix with zero entries. Then, recalling Definition (3.2), we obtain the desired result,

$$\mathbf{Z}_m^T = \left(\mathbf{H}_m^l\right)^T \left(\mathbf{R}_m^l\right)^T, \tag{A.1.16}$$

which completes the proof of Theorem 3.

A.2. Computing the Weight Matrix W

The weight matrix $\mathbf{W}$ in (2.7) is diagonal for a Gaussian distribution defined as the zeros of $\overline{P}_N^m(\theta_i) = 0$. From (A.3.1) this implies that $x_i = \sin\theta_i$ must be an eigenvalue of the symmetric tridiagonal matrix $[\alpha_n^m, 0, \alpha_{n+1}^m]$, $n = m, \ldots, N-1$. Then, given the Gaussian points θ_i, the weights can be computed directly from (2.7) or (2.11) or by a third method described in [10] and implemented in subroutine `gaqd` in `SPHEREPACK`, which is available directly on the World Wide Web. Programs are also included for transforms on an equally spaced grid.

For equally spaced θ_i (including the poles) two weight matrices are required; namely, $\mathbf{W}_0$ for even m and $\mathbf{W}_1$ for odd m. $\mathbf{W}_0^{-1} = \mathbf{P}_0\mathbf{P}_0^T$ has full rank. First compute the $\mathbf{P}_0^T = \mathbf{QR}$ decomposition where $\mathbf{Q}$ is orthogonal and $\mathbf{R}$ is upper triangular. Then $\mathbf{W}_0 = (\mathbf{R}^T)^{-1}\mathbf{R}^{-1}$. The $\mathbf{QR}$ decomposition can be computed using `EISPACK`, available from Netlib. However, $\mathbf{W}_1^{-1} = \mathbf{P}_1\mathbf{P}_1^T$ has rank $N-2$ because $\mathbf{p}_{N-1}^1$ is linearly dependent on the remaining $\mathbf{p}_n^m$. First define $\tilde{\mathbf{p}}_n^1$ with length $N-2$ as $\mathbf{p}_n^1$ but without its first and last zero components. Then define $\tilde{\mathbf{P}}_1 = [\tilde{\mathbf{p}}_1^1 \cdots \tilde{\mathbf{p}}_{N-2}^1]$. Then $\mathbf{W}_1 = (\tilde{\mathbf{P}}_1\tilde{\mathbf{P}}_1^T)^{-1}$. In this manner $\mathbf{W}_1$ is defined on the latitudinal grid minus the points at the pole. It can be extended to the poles by giving each the weight of 1.

For the general distribution of latitudinal points, the "extent" of the linear independence of the $\mathbf{p}_n^l$ ($l = 0, 1$) is unknown. Consequently $\mathbf{W}_l = (\mathbf{P}_l\mathbf{P}_l^T)^{-1}$ may not exist, as was the case for $\mathbf{W}_1$ above. The goal here is to determine the number of independent $\mathbf{p}_n^l$ and the extent to which they are independent so that an informed decision can be made about the number of $\mathbf{p}_n^l$ to retain or, when possible, redistributing θ_i.

To deal effectively with these considerations we choose to use the singular value decomposition (SVD). The subroutine `dsvdc` in `LINPACK` is available from Netlib and provides matrices $\mathbf{U}$, $\mathbf{S}$, and $\mathbf{V}$ such that

$$\mathbf{P}_l = \mathbf{U}_{N\times(N-l)}\mathbf{S}_{(N-l)\times(N-l)}\mathbf{V}_{(N-l)\times(N-l)}^T, \tag{A.2.1}$$

where $\mathbf{U}^T\mathbf{U} = \mathbf{I}_{(N-l)\times(N-l)}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}_{(N-l)\times(N-l)}$, and $\mathbf{S}$ is a diagonal matrix with positive or zero singular values $s_{i,i}$.

At this point, the number of strongly independent $\mathbf{p}_n^l$ (say N') can be determined by the number of largest singular values. This is somewhat subjective but usually easy to determine because the singular values will generally attenuate smoothly until a sudden drop-off. N' can then be chosen to include only the larger singular values. Where optional, N' may be increased by moving or adding latitudinal points.

Next define $\mathbf{P}_l' = [\mathbf{P}_0^l \ \cdots \ \mathbf{P}_{N'-1}^l]$ and compute its singular value decomposition,

$$\mathbf{P}_l' = \mathbf{U}'_{N\times N'}\mathbf{S}'_{N'\times N'}\mathbf{V}'^T_{N'\times N'}. \tag{A.2.2}$$

Now expand $\mathbf{U}'$ to a square l_2 orthonormal matrix $\tilde{\mathbf{U}}_l$ by adding vectors using Gram–Schmidt orthogonalization. Also change zero singular values to $s_i = 1$. Finally, the $N \times N$ weight matrices are given by

$$\mathbf{W}_l = \tilde{\mathbf{U}}_l\mathbf{S}^{-2}(\tilde{\mathbf{U}}_l)^T. \tag{A.2.3}$$

The resulting weight matrices are nonsingular and therefore define a weighted norm $\|\mathbf{u}\|^2_{\mathbf{W}_l} = \mathbf{u}^T\mathbf{W}_l\mathbf{u}$. Also, from (A.2.2) and (A.2.3), it can be verified that $\mathbf{P}_l'^T\mathbf{W}_l\mathbf{P}_l' = \mathbf{I}_{(N'-l)\times(N'-l)}$.

We close this appendix with the following observation. If $\mathbf{P} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ then the weight matrix is $\mathbf{W} = \mathbf{U}\mathbf{S}^{-2}\mathbf{U}^T$, which can be quite ill conditioned. However, the analysis matrix $\mathbf{Z} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T$ is better conditioned and the condition of the projection matrix $\mathbf{F} = \mathbf{U}\mathbf{U}^T$ is best possible and independent of the condition of either $\mathbf{P}$ or $\mathbf{W}$. Therefore, in theory, the projection may be well conditioned even if the weight and analysis matrices are ill-conditioned. Although it may be possible to formulate the harmonic projection in these terms, it is not clear at the time of this writing whether the $O(N^2)$ memory requirement can be retained. Furthermore, a reformulation may not be necessary because the numerical experiments in Section 5 demonstrate that the $O(N^2)$ formulation is already somewhat more accurate than the traditional formulation. This topic is the focus of current research.

A.3. *The Christoffel–Darboux Formula*

Here we follow the proof in Hildebrand [4] with application to the associated Legendre functions. We begin with the well-known three-term recursion formula,

$$x\bar{P}_n^m = \alpha_n^m\bar{P}_{n-1}^m + \alpha_{n+1}^m\bar{P}_{n+1}^m, \quad \text{where} \quad \alpha_n^m = \left[\frac{(n-m)(n+m)}{(2n-1)(2n+1)}\right]^{1/2}. \tag{A.3.1}$$

We wish to determine a closed form of

$$S(x, y) = \sum_{k=m}^{N-1}\bar{P}_k^m(x)\bar{P}_k^m(y). \tag{A.3.2}$$

From (A.3.1),

$$xS(x, y) = \sum_{k=m}^{N-1}\left[\alpha_k^m\bar{P}_{k-1}^m(x) + \alpha_{k+1}^m\bar{P}_{k+1}^m(x)\right]\bar{P}_k^m(y) \tag{A.3.3}$$

and

$$yS(x, y) = \sum_{k=m}^{N-1}\bar{P}_k^m(x)\left[\alpha_k^m\bar{P}_{k-1}^m(y) + \alpha_{k+1}^m\bar{P}_{k+1}^m(y)\right]. \tag{A.3.4}$$

Most of the terms cancel in the difference of (A.3.3) and (A.3.4) and we obtain the desired result,

$$(x-y)\sum_{k=m}^{N-1} \bar{P}_k^m(x)\bar{P}_k^m(y) = \alpha_N^m\left[\bar{P}_N^m(x)\bar{P}_{N-1}^m(y) - \bar{P}_{N-1}^m(x)\bar{P}_N^m(y)\right]. \tag{A.3.5}$$

A.4. *Orthogonal Four-Term Recurrence Transform*

The discrete functions $\mathbf{q}_n^0$ and $\mathbf{q}_n^1$ that provide the memory-efficient alternative to $\mathbf{p}_n^m$ are developed in three steps in Section 3, prior to Table I. A key factor in facilitating their development is that the set $\mathbf{p}_n^{m+1}$ is a linear combination of the $\mathbf{p}_n^{m-1}$. This is provided by the following four-term recurrence. If we define $a_n = [n(n+1)]^{1/2}$ and $b_n = [(2n+3)/(2n-1)]^{1/2}$, then the normalized version of (5.14) in [10] is

$$\mathbf{p}_{n+1}^{m+1} = a_{n+m+1}^{-1}\left(b_n a_{n+m-1}\mathbf{p}_{n-1}^{m-1} - a_{n-m+1}\mathbf{p}_{n+1}^{m-1} + b_n a_{n-m-1}\mathbf{p}_{n-1}^{m+1}\right). \tag{A.4.1}$$

This recurrence relation is also quite stable and provides an accurate method for computing the $\mathbf{p}_n^m$ starting with $\mathbf{p}_n^0$ and $\mathbf{p}_n^1$. An accurate method for computing the latter is also given in [10]. The rows of $\mathbf{Z}_m$ also satisfy (A.4.1), which is initialized by the rows of $\mathbf{Z}_0$ and $\mathbf{Z}_1$, which can be computed by (2.13) or the formulas in [10]. The recurrence (A.4.1) requires eight flops for each component; however, if the coefficients are precomputed and stored in two-dimensional arrays, only five flops are required.

The stability of (A.4.1) results from the orthogonality of the recurrence relation. That is, the transform from $\mathbf{p}_n^{m-1}$ to $\mathbf{p}_n^{m+1}$ corresponds to an orthogonal transform. In addition, the proof of (2.12) in Theorem 1 also depends heavily on its orthogonality, which we proceed now to prove. For $m = 1, \ldots, N-1$, Eq. (A.4.1) can be written in matrix form,

$$\mathbf{X}_m\mathbf{p}^{m+1} = \mathbf{Y}_m\mathbf{p}^{m-1} \quad \text{or} \quad \mathbf{p}^{m+1} = \mathbf{X}_m^{-1}\mathbf{Y}_m\mathbf{p}^{m-1}, \tag{A.4.2}$$

where for arbitrary θ, $\mathbf{p}^m = [P_m^m(\theta)\cdots P_{N-1}^m(\theta)]^T$. $\mathbf{X}_m$ is an $(N-m-1)\times(N-m-1)$ matrix with elements $x_{i,j}$ and $\mathbf{Y}_m$ is an $(N-m-1)\times(N-m+1)$ matrix with elements $y_{i,j}$, where

$$x_{i,j} = \begin{cases} a_{2m+i}, & j=i \\ -b_{m+i-1}a_{i-2}, & j=i-2 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad y_{i,j} = \begin{cases} b_{m+i-1}a_{2m+i-2}, & j=i \\ -a_i, & j=i+2 \\ 0, & \text{otherwise.} \end{cases} \tag{A.4.3}$$

The goal is now to show that $\mathbf{X}_m^{-1}\mathbf{Y}_m$ is orthogonal or that $\mathbf{I}_{(N-m-1)\times(N-m-1)} = \mathbf{X}_m^{-1}\mathbf{Y}_m(\mathbf{X}_m^{-1}\mathbf{Y}_m)^T = \mathbf{X}_m^{-1}\mathbf{Y}_m\mathbf{Y}_m^T(\mathbf{X}_m^{-1})^T$. Equivalently we will show that

$$\mathbf{Y}_m\mathbf{Y}_m^T\left(\mathbf{X}_m^{-1}\right)^T = \mathbf{X}_m. \tag{A.4.4}$$

Let $\mathbf{X}_m^{-1}$ have elements $\xi_{i,j}$. Because of its simple two-diagonal form (diagonal and subdiagonal at a distance of 2 from the diagonal) the elements can be given recursively as

$$\xi_{i,j} = \begin{cases} 1/x_{i,i}, & j=i \\ \frac{-x_{i,i-2}\xi_{i-2,j}}{x_{i,i}}, & j<i,\ (i,j) \text{ same parity} \\ 0, & \text{otherwise.} \end{cases} \tag{A.4.5}$$

$\mathbf{X}_m^{-1}$ is a “checkerboard” lower triangular matrix.

Now define $\mathbf{E}_m = \mathbf{Y}_m \mathbf{Y}_m^T$, which is symmetric pentadiagonal with zero sub- and super-diagonal. Let $\mathbf{E}_m$ have elements $e_{i,j}$; then

$$e_{i,j} = \begin{cases} (y_{i,i})^2 + (y_{i,i+2})^2, & j = i \\ y_{i,i+2} y_{i+2,i+2}, & j = i+2 \\ y_{i,i} y_{i-2,i}, & j = i-2 \\ 0, & \text{otherwise.} \end{cases} \tag{A.4.6}$$

Next define $\hat{\mathbf{X}}_m = \mathbf{E}_m (\mathbf{X}_m^{-1})^T$ with elements $\hat{x}_{i,j}$. By inspection, $\hat{\mathbf{X}}_m$ is also a "checkerboard" matrix with zeros below the subdiagonal located at distance 2 from the diagonal.

To prove (A.4.4), $\hat{\mathbf{X}}_m$ must be shown to equal $\mathbf{X}_m$. To this end we devote the rest of this section to three separate cases.

First Case: The Subdiagonal. From the definition of $\hat{\mathbf{X}}_m$, (A.4.5), and (A.4.6),

$$\begin{aligned} \hat{x}_{i+2,i} &= e_{i+2,i}\xi_{i,i} = y_{i+2,i+2} y_{i,i+2}/x_{i,i} \\ &= -b_{m+i+1} a_{2m+i} a_{i+2}/a_{2m+i} = -b_{m+i+1} a_{i+2} = x_{i+2,i}. \end{aligned} \tag{A.4.7}$$

Second Case: The Diagonal.

$$\hat{x}_{i,i} = \begin{cases} e_{i,i}\xi_{i,i}, & i \le 2 \\ e_{i,i-2}\xi_{i,i-2} + e_{i,i}\xi_{i,i} & i > 2. \end{cases} \tag{A.4.8}$$

We prove only the more complicated case $i > 2$. From (A.4.5), (A.4.6), and (A.4.8),

$$\hat{x}_{i,i} = \frac{1}{x_{i,i}} \left[(y_{i,i})^2 + (y_{i,i+2})^2 - y_{i,i} y_{i-2,i} x_{i,i-2} \xi_{i-2,i-2} \right], \tag{A.4.9}$$

$$= \frac{1}{x_{i,i}} \left[(y_{i,i})^2 + (y_{i,i+2})^2 - y_{i,i} y_{i-2,i} \frac{x_{i,i-2}}{x_{i-2,i-2}} \right], \tag{A.4.10}$$

then from (A.4.3)

$$\hat{x}_{i,i} = \frac{1}{a_{2m+i}} \left[b_{m+i-1}^2 \left(a_{2m+i-2}^2 - a_{i-2}^2 \right) + a_i^2 \right]. \tag{A.4.11}$$

Finally, using the definitions of a_n and b_n prior to (A.4.1)

$$\hat{x}_{i,i} = \frac{(2m+i)(2m+i+1)}{a_{2m+i}} = \frac{a_{2m+i}^2}{a_{2m+i}} = a_{2m+i} = x_{i,i}. \tag{A.4.12}$$

The proof for the subcase $i \le 2$ proceeds in a similar manner using $a_{i-2} = [(i-2)(i-1)]^{1/2} = 0$.

Third Case: Above the Diagonal. If $j > i$ and have the same parity, then

$$\hat{x}_{i,j} = \begin{cases} e_{i,i}\xi_{j,i} + e_{i,i+2}\xi_{j,i+2}, & i \le 2 \\ e_{i,i-2}\xi_{j,i-2} + e_{i,i}\xi_{j,i}, + e_{i,i+2}\xi_{j,i+2} & i > 2. \end{cases} \tag{A.4.13}$$

We again consider only the more complicated case. From the recursive definition (A.4.5),

$$\xi_{j,i} = \frac{x_{j,j-2}}{x_{j,j}} \frac{x_{j-2,j-4}}{x_{j-2,j-2}} \cdots \frac{x_{i+2,i}}{x_{i+2,i+2}} \frac{(-1)^{(i-j)/2}}{x_{i,i}}. \tag{A.4.14}$$

Substituting into (A.4.13)

$$\hat{x}_{i,j} = \xi_{j,i+2}\left\{y_{i,i}y_{i-2,i}\frac{x_{i+2,i}}{x_{i,i}}\frac{x_{i,i-2}}{x_{i-2,i-2}} - \left[(y_{i,i})^2 + (y_{i,i+2})^2\right]\frac{x_{i+2,i}}{x_{i,i}} + y_{i,i+2}y_{i+2,i+2}\right\}, \tag{A.4.15}$$

then from (A.4.3)

$$\hat{x}_{i,j} = \xi_{j,i+2}\left\{\left[b^2_{m+i-1}\left(a^2_{2m+i-2} - a^2_{i-2}\right) + a^2_i\right]\frac{b_{m+i+1}a_i}{a_{2m+i}} - a_i b_{m+i+1}a_{2m+i}\right\}. \tag{A.4.16}$$

Using the definitions of a_n and b_n prior to (A.4.1),

$$b^2_{m+i-1}\left(a^2_{2m+i-2} - a^2_{i-2}\right) + a^2_i = (2m+i)(2m+i+1) = a^2_{2m+i}, \tag{A.4.17}$$

and substituting into (A.4.16),

$$\hat{x}_{i,j} = \xi_{j,i+2}\left(a^2_{2m+i}\frac{b_{m+i+1}a_i}{a_{2m+i}} - a_i b_{m+i+1}a_{2m+i}\right) = 0. \tag{A.4.18}$$

The proof for the subcase $i \le 2$ proceeds in a similar manner using $a_{i-2} = [(i-2)(i-1)]^{1/2} = 0$.

This completes the proof of (A.4.4) and the fact that the four-term recurrence relation corresponds to an orthogonal transform.

A.5. *Resource Centers*

SPHEREPACK is available at

$$\text{http://www.scd.ucar.edu/css/software/spherepack} \tag{A.5.1}$$

Netlib routines are available at

$$\text{http://www.netlib.org/index.html} \tag{A.5.2}$$

ACKNOWLEDGMENT

The authors thank John Boyd for his meticulous review of the manuscript and a number of very helpful comments.

REFERENCES

1. J. C. Adams and P. N. Swarztrauber, *SPHEREPACK 2.0: A Model Development Facility* (NCAR Technical Note NCAR/TN-436-STR, September 1997).
2. J. P. Boyd, Multipole expansions and pseudospectral cardinal functions: A new generalization of the fast Fourier transform, *J. Comput. Phys.* **103**, 184 (1992).
3. B. Fornberg, *A Practical Guide to Pseudospectral Methods* (Cambridge Univ. Press, New York, 1996).
4. F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw–Hill, New York, 1956.
5. R. Jakob-Chien and B. K. Alpert, A fast spherical filter with uniform resolution, *J. Comput. Phys.* **145**, 580 (1997).

6. B. Machenhauer and R. Daley, A baroclinic primitive equation model with spectral representation in three dimensions (Institute for Theoretical Meteorology, Copenhagen University, 1972), Report 4.
7. W. F. Spotz, M. A. Taylor, and P. N. Swarztrauber, Fast shallow-water equation solvers in latitude–longitude coordinates, *J. Comput. Phys.* **145**, 432 (1998).
8. W. F. Spotz and P. N. Swarztrauber, A performance comparison of associated Legendre projections, in preparation.
9. P. N. Swarztrauber, On the spectral approximation of discrete scalar and vector functions on the sphere, *SIAM J. Numer. Anal.* **16**, 934 (1979).
10. P. N. Swarztrauber, The vector harmonic transform method for solving partial differential equations in spherical geometry, *Mon. Weather Rev.* **121**, 3415 (1993).
11. N. Yarvin and V. Rokhlin, A generalized one-dimensional fast multipole method with applications to filtering of spherical harmonics, *J. Comput. Phys.* **147**, 594 (1998).