Double Fourier Series Solution
of Poisson's Equation on A Sphere

SAMUEL Y. K. YEE

29 October 1980

Approved for public release; distribution unlimited.
This report has been reviewed by the ESD Information Office (OI) and is releasable to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

[Signature]
Chief Scientist

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.
DOUBLE FOURIER SERIES SOLUTION OF POISSON'S EQUATION ON A SPHERE

Samuel Y. K. Ye".

Air Force Geophysics Laboratory (Lyd)
Hanscom AFB
Massachusetts 01731

ABSTRACT

Advances in numerical simulation and prediction in disciplines as diverse as geophysical fluid dynamics, heat transfer, and nuclear and plasma physics have generated, in recent years, considerable interest in the method of solution for Poisson-type equations. A method for the solution of Poisson's equation on the surface of a sphere is given. The method makes use of truncated double Fourier series expansions on the sphere and involves the Galerkin approximation. It has an operation count of approximately $12J^2(1 + \log_2 J)$ for a...
20. Abstract (Continued)

Latitude-longitude grid containing $2J \times (J - 1) + 2$ data points. Numerical results are presented to demonstrate the method's accuracy and efficiency.
Preface

The author would like to thank Donald Aiken for assistance in this work, Jack Mettauer for discussions on the properties of Eqs. (17a) and (17b), and Helen Connell for accurate and efficient typing.
Contents

1. INTRODUCTION ........................................... 7
2. ORDINARY DIFFERENTIAL EQUATIONS WITH VARIABLE COEFFICIENTS ... 9
3. SPECTRAL ALGEBRAIC EQUATIONS ......................... 10
4. TRIDIAGONAL SYSTEMS .................................... 13
5. TEST COMPUTATIONS ..................................... 15
6. REFERENCES ............................................... 19

Tables

1. Accuracy as a Function of N for \(m_1 = 1, m_2 = 2\) .......... 18
2. Accuracy as a Function of \(m_2\) for a \((63 \times 128)\)-grid .... 18
Double Fourier Series Solution of Poisson's Equation on a Sphere

1. INTRODUCTION

Advances in numerical simulation and prediction in disciplines as diverse as geophysical fluid dynamics, heat transfer, and nuclear and plasma physics have generated, over the years, considerable interest in the method of solution for Poisson-type equations. Most research in this area, however, does not deal with spherical geometries. In fact, of a list of 150 articles on "Fast Elliptic Solvers" compiled by Schumann, 1 only two 2,3 deal with the spherical geometry. We present in this paper a new numerical method for the solution of Poisson's equation on a sphere. In contrast to previous methods of Swarztrauber 2 and Yee 3 which are based on finite-difference, this method is based on truncated double Fourier series on spheres (for example, Orszag 4; Boer and Steinberg 5). The use of finite double

(Received for publication 28 October 1980)


Fourier series on a rectangle is not new, but the formulation on a sphere reported here is. In particular, Yee's interpretation of Fourier series on spheres has been implemented to narrow what Swarztrauber referred to as "the theoretical gap which exists between the states of the art for discrete spectral approximations on a sphere and on a rectangle," and to obtain a Galerkin approximation to the solution of Poisson's equation on a sphere.

Let \( u(\theta, \lambda) \) be a scalar function on a sphere where the location of a point is specified by co-latitude \( 0 \leq \theta \leq \pi \) and longitude \( 0 \leq \lambda < 2\pi \). It may then be represented by double Fourier series of the form

\[
  u(\theta, \lambda) = \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} u_{l, m} [(1 - s) \cos \lambda + s \sin \lambda] e^{im\lambda},
\]

(1a)

where

\[
  u_{l, m} = \frac{c}{\pi} \int_{0}^{2\pi} \int_{0}^{\pi} u(\theta, \lambda) e^{-im\lambda} d\lambda \left[ (1 - s) \cos \lambda + s \sin \lambda \right] d\theta,
\]

(1b)

and \( s = 0 \) if \( m \) is even, \( s = 1 \) if \( m \) is odd. Since the transform pair Eqs. (1a) and (1b) is amenable to Fast Fourier Transform (FFT), we have at our disposal an extremely efficient means of obtaining a numerical solution for Poisson's equation on the surface of a sphere.

The procedure involves the expansion of the dependent variables in truncated double Fourier series, the substitution of the truncated series for the dependent variables in the Poisson equation, the application of the Galerkin approximation to obtain in Fourier space a number of linear algebraic systems, the solution of these systems, and the inverse transform of the solution in Fourier space back to physical space.


2. ORDINARY DIFFERENTIAL EQUATIONS WITH VARIABLE COEFFICIENTS

Our goal is to seek, using the FFT technique, an accurate numerical solution to the Poisson equation on the surface of a unit sphere \(0 \leq \theta \leq \pi, 0 \leq \lambda < 2\pi\),

\[
\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial u}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \lambda^2} = f(\theta, \lambda) .
\] (2a)

Here the forcing function \(f(\theta, \lambda)\) must satisfy the compatibility condition (for example, Berg and McGregor\(^9\))

\[
\int_S f(\theta, \lambda) \sin \theta \, ds = 0
\] (2b)

over the surface of the sphere \(S\).

The first step in our solution method is the transformation of this partial differential equation (PDE) to a number of ordinary differential equations (ODE), a standard approach in the method of solution for PDE. A brief outline of this procedure is included here for completeness. We decompose \(u(\theta, \lambda)\) along a given latitude by a truncated Fourier series of the form

\[
u(\theta, \lambda) = \sum_{m=-M}^{M} u_m(\theta) e^{im\lambda},
\] (3a)

where

\[
u_m(\theta) = \frac{cK}{K} \sum_{k=1}^{K} u(\theta, \lambda_k) e^{-im\lambda_k} 
\]  

\[c = \begin{cases} 
1 & \text{for } m = 0 \text{ or } \pm M \\
2 & \text{otherwise}
\end{cases}
\] (3b)

are complex Fourier coefficients, \(M = K/2\) is the cutoff wavenumber, \(\lambda_k = 2\pi k/K\), and \(K\) is the number of data points along a latitude circle. Since a latitude circle degenerates into a single point at the poles, for \(u(\theta, \lambda)\) to have a single value \(u(P)\) there, we must have the pole conditions

where $P = 0$ or $\pi$, and $|m| \leq M$. A Galerkin approximation to Eq. (2a), represented by a set of ODE with variable coefficients, is obtained if we substitute for $u$ and $f$ in Eq. (2a) by truncated series of the form of Eq. (3a), multiply the resulting equations by $\exp(-im\lambda)$, and then integrate over $\lambda$ for the interval $0 \leq \lambda < 2\pi$:

$$
\frac{1}{\sin \theta} \frac{d}{d\theta} \sin \theta \frac{d}{d\theta} u_m(\theta) - \frac{m^2}{\sin^2 \theta} u_m(\theta) = f_m(\theta),
$$

where $0 \leq \theta \leq \pi$, $|m| \leq M$. For convenience of discussion, we rewrite Eq. (5) in the form

$$
u_m''(\theta) + \frac{\cos \theta}{\sin \theta} \nu_m'(\theta) - \frac{m^2}{\sin^2 \theta} u_m(\theta) = f_m(\theta).$$

Equation (6) together with Eq. (4) constitutes a set of one-dimensional Dirichlet boundary-value problems (BVP). Of this set, $2M$ equations have homogeneous boundary conditions and one equation has inhomogeneous boundary conditions given by Eq. (4a). Thus the problem of solving a two-dimensional Poisson equation becomes that of solving a set of one-dimensional variable coefficient BVP of the Helmholtz-type. The pole conditions may now be considered as boundary conditions. For $m = 0$, however, Eq. (6) has no unique solution because the pole values $u(P)$ in Eq. (4a) are among the unknowns to be sought in Eqs. (2a) and (2b). This, of course, indicates nothing more than that solutions to Poisson's equation on a sphere can only be determined to within an additive constant. We shall return to this point later.

3. SPECTRAL ALGEBRAIC EQUATIONS

The next key step in our solution method is based on the fact that $u_m(\theta)$ is a $2\pi$-periodic even (odd) function for even (odd) $m$, and we may approximate $u_m(\theta)$ in truncated half-range expansions of cosine (sine):
\[
\begin{align*}
\mathbf{u}_m(\theta) &= \left\{ \begin{array}{ll}
\sum_{l=0}^{L} u_{l, m} \cos l \theta & \text{for even } m \\
\sum_{l=0}^{L} u_{l, m} \sin l \theta & \text{for odd } m \\
\end{array} \right. \\
\end{align*}
\] (7a)

where

\[
\begin{align*}
u_{l, 0} &= \frac{c}{J} \sum_{j=1}^{J-1} u_{0, j}^{(\theta_j)} \cos \theta_j + \frac{c}{2J} [u(\theta = 0) + (-1)^{l-1} u(\theta = \pi)] \\
u_{l, m} &= \left\{ \begin{array}{ll}
\frac{c}{J} \sum_{j=1}^{J-1} u_{m, j}^{(\theta_j)} \cos \theta_j & \text{for even } m \neq 0 \\
\frac{c}{J} \sum_{j=1}^{J-1} u_{m, j}^{(\theta_j)} \sin \theta_j & \text{for odd } m \\
\end{array} \right. \\
\end{align*}
\] (7b)

\(c = 1\) if \(l = 0\) or \(L\), \(c = 2\) otherwise, \(L = J\) is a cutoff, \(\theta_j = j\pi/J\), and \((J-1)\) is the number of data points between poles.

We consider first the constraints on the use of these expansions to solve Eq. (6). For \(m = 0\), Eq. (6) requires that \(u_0'(0)/\sin \theta\) be finite for all \(\theta\), including \(\theta = 0\) and \(\pi\). The cosine expansions in Eq. (7a) obviously satisfy this requirement. For \(m \neq 0\), Eq. (6) requires at least that \(u_m(0)/\sin \theta\) be finite for all \(\theta\) (Orszag). In the case of odd \(m\), this condition is automatically satisfied by the sine expansions in Eq. (7a). In the case of even \(m \neq 0\), however, we must impose on the cosine expansions the constraints

\[
\begin{align*}
u_m(0) &= \nu_m + w_m = 0 \\
\end{align*}
\] (8a)

and

\[
\begin{align*}
u_m(\pi) &= \nu_m - w_m = 0 \\
\end{align*}
\] (8b)

where

\[
\begin{align*}
\nu_m &= \sum_{l \text{ even}} u_{l, m} \\
\end{align*}
\] (9a)
Equations (8a) and (8b) will hold only if we make both

\[ v_m = 0 \quad (10a) \]

and

\[ w_m = 0 \quad (10b) \]

These are then the "pole conditions" which the spectral equations must satisfy for the absence of singularities at the poles in Eq. (6). Note that these conditions are consistent with Eqs. (4a) and (4b), the pole conditions for the east-west Fourier expansions of \( u(0, \lambda) \).

Just as Eqs. (3a) and (3b) provide the linkage between a two-dimensional PDE and a set of ODE, Eqs. (7a) and (7b) enable us to reduce Eq. (6) to a set of algebraic systems in the Fourier space \( m \leq M, 0 \leq l \leq 4 \):

\[ (l - 2)(l - 1)u_{l - 2, m} - (2l^2 + 4m^2)u_{l, m} + (l + 1)(l + 2)u_{l + 2, m} \]

\[ -f_{l - 2, m} + 2f_{l, m} = f_{l + 2, m} \quad (11) \]

Here \( f_{l, m} \) are complex Fourier coefficients of \( f_m(\theta) \). In the derivation of Eq. (11), we have

1. Made use of the identity \( 2 \sin^2 \theta - 1 = \cos 2\theta \) to rewrite Eq. (4) in the form

\[ u_m''(\theta) = \cos 2\theta u_m''(\theta) + \sin 2\theta u_m'(\theta) - 2\theta u_m'(\theta) - 2\theta^2 u_m(\theta) + (1 - \cos 2\theta)f_m(\theta) \quad (12) \]

2. Expanded \( u_m''(\theta) \), \( u_m'(\theta) \), \( u_m(\theta) \) and \( f_m(\theta) \) in Eq. (12) as truncated series in the form of Eqs. (7a) and (7b).

3. Made use of the trigonometric identities

\[ 2 \cos 2\theta \cos l\theta = \cos (l - 2)\theta + \cos (l + 2)\theta \]

\[ 2 \sin 2\theta \sin l\theta = \cos (l - 2)\theta - \cos (l + 2)\theta \]

\[ 2 \cos 2\theta \sin l\theta = \sin (l - 2)\theta + \sin (l + 2)\theta \]

\[ 2 \sin 2\theta \cos l\theta = -\sin (l - 2)\theta + \sin (l + 2)\theta \]
4. Invoked the Galerkin approximation in the manner that Eq. (3) was derived from Eqs. (2a) and (2b) and dropped the \((L+1)\)th and \((L+2)\)th components which are outside the range of our series expansions defined by Eqs. (7a) and (7b).

We now discuss the pole conditions for Eq. (6) in the context of spectral algebraic systems Eq. (11). As mentioned previously, Eqs. (2a) and (2b) prescribe \(u\) on a sphere only to within an additive constant. To determine \(u\) uniquely, we need to have one additional piece of information on \(u\). Without loss of generality, we shall impose the condition

\[ u(\theta = 0) + u(\theta = \pi) = \text{constant} = 0. \]  

(13)

In Fourier space this condition becomes, through Eqs. (4a) and (7a),

\[ v_0 = \sum_{t \text{ even}} u_{t,0} = 0. \]  

(14)

This is consistent with the pole conditions \(v_m\) for the cases of even \(m \neq 0\) given in Eqs. (10a) and (10b). In solving Eq. (11), we shall use Eq. (14) to compute \(u_{0,0}\), the constant term in our double Fourier series expansions. For even \(m \neq 0\), \(u_{0,m}\) are computed via the pole conditions given in Eq. (10):

\[ u_{0,m} = -\sum_{t \text{ even}} u_{t,m}. \]  

(15)

For odd \(m\), \(u_{0,m}\) are easily found from Eq. (7b) to be

\[ u_{0,m} = u_{L,m} = 0. \]  

(16)

4. TRIDIAGONAL SYSTEMS

With the problem of pole conditions settled, we are now ready to solve linear algebraic systems given in Eq. (11). We note first that for a given \(m\), the even \(t\) and odd \(t\) components are uncoupled. We may therefore break the algebraic system for each \(m\) into two independent subsystems. Furthermore, written in matrix form, these subsystems have a very desirable feature in that their coefficient
matrices are tridiagonal. Very efficient algorithms are available for the solution of such tridiagonal systems (for example, Varga\textsuperscript{10}). We write Eq. (11) in matrix notations:

\begin{align*}
\begin{pmatrix}
b_{2, m} & c_2 \\
a_4 & c_{1,-2} & \ddots & \ddots \\
 & a_1 & b_{1, m} \\
\end{pmatrix}
\begin{pmatrix}
u_2 \\
u_4 \\
\vdots \\
u_{l-2} \\
u_1 \\
\end{pmatrix}
= 
\begin{pmatrix}
h_2 \\
h_4 \\
\vdots \\
h_{l-2} \\
h_1 \\
\end{pmatrix}
\end{align*}

(17a)

\begin{align*}
\begin{pmatrix}
b_{1, m} & c_1 \\
a_3 & c_{1,-1} & \ddots & \ddots \\
 & a_{1,-1} & b_{1,-1, m} \\
\end{pmatrix}
\begin{pmatrix}
u_1 \\
u_3 \\
\vdots \\
u_{l-3} \\
u_{l-1} \\
\end{pmatrix}
= 
\begin{pmatrix}
h_1 \\
h_3 \\
\vdots \\
h_{l-3} \\
h_{l-1} \\
\end{pmatrix}
\end{align*}

(17b)

where

- \(|m| < M\),
- \(a_f = (t - 2)(t - 1)\),
- \(b_{f, m} = 2f^2 - 4m^2\),
- \(c_f = (t + 1)(t + 2)\),
- \(h_{f, m} = f_{f-2, m} + 2f_{f, m} - f_{f+1, m}\),
- \(e_{2, m} = h_2, m - (-1)^f f_0, m\).

In Eqs. (17a) and (17b), we have dropped the \((1, +1, m)\) and \((1, +2, m)\) components because they are beyond the range of our series expansions. On the other hand, the components \((-2, m)\) and \((-1, m)\) have been retained because \(u_m(\theta), f_0, m(\theta)\) are \(2\pi\)-periodic in \(\theta\). Had we used expansions over \(0 \leq \theta < 2\pi\) instead of using half-range expansions as in Eqs. (7a) and (7b), then any component \(|f| \geq 1\) would have

been within the spectral range. We retained the \((-2, m), (-1, m)\) components by the use of the identity
\[
(-1)^m f_{-1, m} = f_{1, m}.
\]
For even \(m\), we have \(L\) unknowns for each \(m\). For odd \(m\), \(u_{1, m}\) are known beforehand to be zero. In this case, we simply set \(u_{1, m} = 0\) and discard the last member in Eq. (17a). Note that the components \(u_{0, m}, |m| \leq M\), are not included in Eqs. (17a) and (17b). These are to be obtained through the pole conditions discussed in Section 3: For \(m = 0\), \(u_{0, 0}\) is given by Eq. (14),
\[
u_{0, 0} = \sum_{\ell \text{ even, } \ell \neq 0} u_{\ell, 0};
\]
for even \(m \neq 0\), \(u_{0, m}\) are given by Eq. (15). For odd \(m\), \(u_{0, m}\) are of course known from Eq. (16) to be zero.

So far we have demonstrated that the Poisson equation on a sphere can be approximated by a number of tridiagonal systems in the Fourier space. At this point, it is appropriate to examine the properties of these systems. A little simple arithmetic with \(a_{\ell}, b_{\ell, m}\) and \(c_{\ell}\) reveals that the coefficient matrices of these systems are strictly diagonally dominant for the cases \(|m| \geq 2\). Unique solutions therefore exist for these cases. For \(m = 0, \pm 1\), the situation is less clearcut. But it can be shown that the determinants of the coefficient matrices for these cases are all nonzero, these algebraic systems therefore also have unique solutions.

5. TEST COMPUTATIONS

As mentioned earlier, very efficient algorithms are available for solution of tridiagonal systems such as Eqs. (17a) and (17b). We shall report in this section some results of our numerical test computations. Before doing so, let us pause for a moment to recapitulate our numerical procedure. We shall also indicate in brackets the approximate minimum number of arithmetic operations required for each step:

1. For each \(j = 1 \text{ through } J - 1\), transform \(f_{j, k}\) to obtain complex Fourier coefficients \(\hat{f}_{m}(0)\): set \(\hat{f}_{0}(0) = f(0), \hat{f}(0) = 0\) or \(\pi, [K(J - 1) \log_2 K]\)

2. For a given even (odd) \(m\), perform cosine (sine) transform on \(\hat{f}_{m}(0)\) to obtain \(f_{j, m}\): \([J \log_2 J]\)
3. Compute $h_{t,m}$ in Eqs. (17a) and (17b). [2J]

4. Solve Eqs. (17a) and (17b) for $u_{t,m'}$ using an algorithm for tridiagonal systems. [10] [4J]

5. For a given even (odd) $m$, perform inverse cosine (sine) transform on $u_{t,m}$ to obtain $u_m(\theta_j)$. [J log$_2$ J]

6. Repeat steps (2) through (5) for all $|m| \leq M$. [Multiply each count in steps (2) through (5) by $K$]

7. Inverse transform $u_m(\theta_j)$, $j = 1, J - 1$, to obtain $u_{j,k}$; set $u(P) = u_0(P)$. [K(J - 1) log$_2$ K]

Thus, for a latitude-longitude grid with $2J \times (J - 1) + 2$ data points ($K = 2J$), the approximate number of arithmetic operations is $12J^2 (1 + \log_2 J)$, or about $6 (1 + \log_2 J)$ operations per data point. Here we assume that $a_{t,m}$, $b_{t,m}$, $c_t$ have been precomputed and $J = 2^p$, where $p$ is a positive integer. An arithmetic operation is defined as a multiplication followed by an addition in the real domain. These counts are comparable to operation counts for the Fourier solution of the Poisson equation on a rectangle (for example, Swarztrauber$^{11}$).

The test function used for program checkout is

$$f(\theta, \lambda) = \sum_{m=1}^{m_2} (m + 1)(m + 2) \cos \theta \sin^m \theta \cos m(\lambda - d_m) - 2 \cos \theta,$$

where $d_m$ are random numbers within the range $(0, 2\pi)$. For this forcing function, Eqs. (2a) and (2b) has the exact solution

$$u_0(\theta, \lambda) = \sum_{m=1}^{m_2} \cos \theta \sin^m \theta \cos m(\lambda - d_m) + \cos \theta.$$

With the exact solutions known, a normalized $L_2$ error norm defined by

$$||E|| = ||u - u_A|| / ||u_A||$$

may then be considered as a measure of the accuracy of our numerical solution $u$. The number of digits of accuracy in $u$ is then given by

Several sets of computations have been made using a CDC 6600 computer for various grid resolutions and various values of \( m_1 \) and \( m_2 \). Some sample results are tabulated in Table 1 for cases in which the data contain only long waves \((m_1 = 1, m_2 = 2)\). We see that 11-, 12-digit accuracies are retained in our numerical solutions for the \( 10^9 \times 10^0 \), \( 5^0 \times 5^0 \), \( 2.8^3 \times 2.8^0 \) resolutions tested. Note that excessive resolutions have been used in these computations in the sense that the tabulated resolutions are capable of resolving waves having wave numbers up to 18, 36, and 64, respectively, while the data contain only wave components up to wave number 2. For the smooth function tested, the slow degradation of accuracy with increasing resolution is due to the increase in the value of the condition number as the order of the coefficient matrices in Eqs. (17a) and (17b) increases. The central processor time required for each solution is given in seconds in the last column in Table 1. It should be noted here that the FFT routine used is one that can accommodate data sets not necessarily having \( 2^p \) pieces of data \((p \) is a positive integer). This special feature of course requires an operation count higher than \( O(\log_2 J) \) in the transforms for cases where \( J \neq 2^p \).

Table 2 shows the accuracy of our numerical results for a \( 63 \times 128 \) grid as a function of \( m_2 \), which may be taken here as the number of components contained in our test function. It is apparent that there is also a gradual degradation in accuracy as the function becomes less smooth. However, even for a function containing components up to the Nyquist frequency, the results still possess more than 9-digit accuracy. Also given in Table 2 is the normalized error for an experiment with \( m_2 = 65 \). In this case, one of the components has a wavelength of less than 2-grid length. The less than 3-digit accuracy in the results is actually not as bad as it may seem. What is demonstrated here is simply that for any forcing function having components with wavelengths less than 2-grid interval, these components are simply truncated by the transform process. For components with wavelength longer than 2-grid interval, the deterioration of accuracy of the solution with decreasing wavelength is much more gradual.
Table 1. Accuracy as a Function of $N$ for $(m_1 = 1, m_2 \leq 2)$

<table>
<thead>
<tr>
<th>$J$</th>
<th>$N$</th>
<th>$Z$ (digits)</th>
<th>$T$ (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>9</td>
<td>11.07</td>
<td>0.59</td>
</tr>
<tr>
<td>36</td>
<td>18</td>
<td>10.94</td>
<td>2.35</td>
</tr>
<tr>
<td>64</td>
<td>32</td>
<td>10.86</td>
<td>3.98</td>
</tr>
</tbody>
</table>

$N$: the size of the tridiagonal system

Table 2. Accuracy as a Function of $m_2$ for a $63 \times 128$-grid

<table>
<thead>
<tr>
<th>$m_2$</th>
<th>$Z$ (digits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>10.81</td>
</tr>
<tr>
<td>32</td>
<td>10.52</td>
</tr>
<tr>
<td>64</td>
<td>9.80</td>
</tr>
<tr>
<td>65</td>
<td>2.92</td>
</tr>
</tbody>
</table>

$m_2$: the highest wavenumber component contained in the forcing function
References


