A Method for Direct Solution of a Steady Linearized Spectral General Circulation Model

EDWIN K. SCHNEIDER

Department of Meteorology, University of Maryland, College Park, Maryland

(Manuscript received 3 January 1989, in final form 14 April 1989)

ABSTRACT

A steady linearized version of a general circulation model (GCM) is a potentially useful tool for diagnosis and understanding of the time-mean solutions of the GCM. A method is developed for direct solution of the linearized equations. The method is efficient in the case that the GCM employs spectral horizontal discretization and finite differences in the vertical sigma coordinate, with seven or more vertical levels.

The method uses a system of equations that is equivalent to the linearized GCM, but with an expanded number of variables. The resulting system of equations can then be solved by sparse matrix techniques. The method is applicable to general basic states.

1. Introduction

A scheme is developed for numerical solution of the steady or constant frequency linearized primitive equations in sigma coordinates. The scheme applies when finite differences are used for vertical discretization. The primary use foreseen for this method is for models in which the horizontal discretization is fully spectral and the number of vertical levels exceeds seven. This is the case, for example, for a diagnostic stationary wave model of a moderate vertical resolution spectral general circulation model (GCM). For linear models in which the horizontal discretization is performed by finite differencing or semispectrally, other numerical schemes will be more efficient.

The first application of a steady linearized model of the primitive equations in sigma coordinates, using vertical finite differencing and a spectral representation in the horizontal, was discussed by Hoskins and Karoly (1981). In that study, the matrix governing the steady linear model was generated in part by computing the time tendencies of a linearized adiabatic inviscid spectral GCM. It was convenient to assume that the solution was given in terms of the GCM prognostic variables relative vorticity, divergence, temperature, and ln(surface pressure). The matrix that is generated using this procedure is not sparse, and has order that is proportional to \( L \), the number of vertical levels. Therefore, the number of operations required to solve the system increases proportional to \( L^3 \).

Here, a method for organizing the variables so that the number of operations required for solution increases proportional to \( L \) rather than \( L^3 \) is developed.

This method requires assuming an expanded set of variables for the linear solution that includes the GCM diagnostic variables geopotential and vertical sigma velocity, as well as the prognostic variables. With this set of variables, the steady linear model matrix becomes block tridiagonal with additional nonzero blocks in the far right column. Each block involves only the variables at a single level. The simplified structure of the matrix is due to the system of equations at level \( k \) involving only the variables at level \( k \), the adjacent levels \( k + 1 \) and \( k - 1 \), and the variable \( \ln(\text{surface pressure}) \). The linear systems solved with the prognostic variables only and the expanded variable set are equivalent. The sparse matrix expanded variable set problem can then be solved by a variant of the Lindzen–Kuo algorithm (Lindzen and Kuo 1969).

The scheme developed here becomes cost-effective when \( L \) exceeds about seven. The method can be applied to solve the linearized primitive equations with either zonally symmetric or zonally asymmetric basic states.

2. Derivation

We begin with the primitive equations on a sphere with vertical coordinate \( \sigma = p/p_s \), where \( p \) is pressure and \( p_s \) is surface pressure. The vertical discretization used by the National Meteorological Center (NMC) medium-range weather prediction model (Sela 1980) is introduced for concreteness, and the steady, linear vertically discretized primitive equation system in which the equations applied to level \( k \) depend only on variables defined at levels \( k - 1, k, k + 1 \), and \( q = \ln(p) \) is developed. Special treatment is required for the hydrostatic equation because of the use of energy conserving vertical differencing (Arakawa and Lamb 1977) in the NMC model.
a. The prognostic primitive equations

The primitive equations on a sphere can be written

\[
\frac{\partial \mathbf{T}}{\partial t} + \alpha(A, B) = F_t
\]

(1)

\[
\frac{\partial D}{\partial t} + \alpha(B, -A) + \nabla^2 (E + \phi) = F_D
\]

(2)

\[
\frac{dT}{dt} + \frac{\partial \phi}{\partial \sigma} \frac{dT}{\partial \sigma} (T \sigma^{-s}) - \kappa T \frac{dq}{dt} = \frac{Q}{c_p}
\]

(3)

\[
\frac{dq}{dt} + D + \frac{\partial \phi}{\partial \sigma} = 0
\]

(4)

\[
\frac{\partial \phi}{\partial \sigma} = -\frac{RT}{\sigma}.
\]

(5)

The dependent variables are \( \mathbf{T} \), the vertical component of the relative vorticity; \( D \), the divergence of the horizontal velocity; \( T \), the virtual temperature; \( \phi \), the geopotential; \( \mathbf{F} = \mathbf{d}x/dt \), the vertical sigma velocity; and \( \mathcal{Q} = \ln P \). The gas constant is \( R \) and \( \kappa = R/c_p \), where \( c_p \) is the specific heat of air at constant pressure.

The forcing functions \( F_t, F_D, \) and \( \mathcal{Q} \) represent parameterized physical processes. Also

\[
E = \frac{u^2 + v^2}{2},
\]

(6)

where \( u \) is the eastward velocity and \( v \) is the northward velocity. The operator \( \alpha \) is defined as

\[
\alpha(g, h) = \frac{1}{a(1 - \mu^2)} \frac{\partial g}{\partial \lambda} + \frac{1}{a} \frac{\partial h}{\partial \mu},
\]

where \( a \) is the planetary radius, \( \lambda \) is longitude, and \( \mu = \sin(\text{latitude}) \). The definitions of \( A \) and \( B \) are

\[
A = (\hat{\psi} + 2\Omega \mu) U + \frac{\partial V}{\partial \sigma} + \frac{RT}{a} (1 - \mu^2) \frac{\partial q}{\partial \mu},
\]

(7)

\[
B = (\hat{\psi} + 2\Omega \mu) V - \frac{\partial U}{\partial \sigma} - \frac{RT \partial q}{a} \frac{\partial \mu}{\partial \lambda},
\]

(8)

where \( \Omega \) is the planetary rotational frequency, \( U = u(1 - \mu^2)^{1/2} \), and \( V = v(1 - \mu^2)^{1/2} \). The total time derivative is

\[
\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\partial}{\partial \sigma} \left( \frac{U}{1 - \mu^2} \frac{\partial \mathbf{\phi}}{\partial \mu} \right) + \frac{\partial}{\partial \mu} \left( \frac{\partial \mathbf{\phi}}{\partial \mu} \right) + \mathbf{\nabla} \cdot \mathbf{\nabla},
\]

(9)

where \( \mathbf{\nabla} = (u, v) \), and \( t \) is time. The formulas \( \mathbf{T} = \alpha(V, -U) \) and \( D = \alpha(U, V) \) relate vorticity and divergence to the horizontal velocities. In implementing these equations in a GCM, a prognostic equation for \( \mathbf{q} \) and a diagnostic equation for \( \mathbf{\phi} \) are derived by vertically integrating (4) using the boundary conditions \( \hat{\mathbf{q}} = 0 \) at \( \sigma = 0 \) and \( \sigma = 1 \):

\[
\frac{d\mathbf{q}}{dt} = -\int_0^1 (\mathbf{v} \cdot \nabla \mathbf{q} + D) d\mathbf{\sigma}
\]

(10)

and

\[
\hat{\mathbf{q}} = \frac{\sigma}{\sigma} \int_0^1 (\mathbf{v} \cdot \nabla \mathbf{q} + D) d\sigma - \int_0^1 (\mathbf{v} \cdot \nabla \mathbf{q} + D) d\sigma.
\]

(11)

b. Vertical discretization

The vertical discretization is accomplished by defining \( \hat{\sigma}_1 = 0 \) and a set of interface spacings \( \Delta \kappa \) such that

\[
\sum_{k=1}^{K} \Delta_k = 1.
\]

Then \( \hat{\sigma}_k = \hat{\sigma}_{k-1} + \Delta_k \) and \( \hat{\sigma}_{K+1} = 1 \). The vertical \( \sigma \) velocity at \( \hat{\sigma}_k \) is taken to be \( \hat{\sigma}_{k-1} \), and \( \hat{\sigma}_0 = \hat{\sigma}_K = 0 \). The other variables are defined at staggered levels \( \sigma_k \), where \( 1 \leq k \leq K \), and \( \hat{\sigma}_k < \sigma_k < \hat{\sigma}_{k+1} \), and are denoted as \( \hat{\sigma}_k \), \( D_k \), \( T_k \), and \( \mathbf{\phi}_k \). Except for vertically differenced terms, vertical discretization is accomplished by replacing the variables \( \hat{\mathbf{\phi}} \), \( D \), \( T \), and \( \mathbf{\phi} \) and forcings \( F_t, F_D, \) and \( \mathcal{Q} \) in (1)–(4) by their subscripted counterparts. Similarly, the continuous variables in (6)–(9) are subscripted.

Vertical advection terms (those containing \( \hat{\sigma} \)) in (3), (7), and (8) are replaced by the quadratic conserving finite difference form

\[
\left( f(\hat{\sigma}) \frac{\partial \mathbf{X}}{\partial \sigma} \right)_k = \frac{f(\hat{\sigma}_k)}{2\Delta_k} \left[ \hat{\sigma}_k(X_{k+1} - X_k) + \hat{\sigma}_{k-1}(X_k - X_{k-1}) \right],
\]

where terms multiplied by \( \hat{\sigma}_0 \) or \( \hat{\sigma}_K \) are taken to be zero. The term \( \hat{\mathbf{\sigma}} \partial \mathbf{\phi} / \partial \sigma \) in (4) is written as \( (\hat{\sigma}_k - \hat{\sigma}_{k-1}) / \Delta_k \) at level \( k \).

Requiring energy conservation leads to the finite difference form of the hydrostatic equation:

\[
\hat{\mathbf{\phi}}_k - \hat{\mathbf{\phi}}_{k-1} = \frac{c_p}{2} \left\{ T_k \left( \left( \frac{\hat{\sigma}_k}{\hat{\sigma}_{k-1}} \right)^s - 1 \right) + T_{k-1} \left( 1 - \left( \frac{\hat{\sigma}_k}{\hat{\sigma}_{k-1}} \right)^s \right) \right\}
\]

\[
k = 2, \ldots, K
\]

and

\[
\sum_{j=1}^{K} \Delta_j \phi_j = \hat{\phi}_{K+1} + R \sum_{j=1}^{K} \Delta_j T_j
\]

(12)

\[
(13)
\]

where \( \hat{\phi}_{K+1} \) is the surface geopotential.

For time integration, \( \hat{\mathbf{\phi}}_k \) is eliminated in favor of the vertically discretized analogs of (10) and (11). Then \( \hat{\sigma}_k \) depends on values of \( \hat{\mathbf{\phi}}_j \) and \( D_j \) for all \( j \leq k \) as well as \( \sigma \). Also, (12) and (13) are used to eliminate \( \mathbf{\phi}_k \) from the vector whose \( k \)th element is \( \phi_k \), in terms of the vector \( \mathbf{T} \) whose \( k \)th element is \( T_k \) and the surface geopotential. The result is that \( \phi_k \) is a linear combination of the values \( T_j \) at all levels and the surface geopotential. The
prognostic system with variables $\xi_k, D_k, T_k,$ and $q$ is strongly coupled in the vertical, as is the linearized version of these equations.

c. Steady linearized equations

The steady linear model is found by taking $\partial / \partial t = 0$, defining each dependent variable and forcing as $X = \dot{X}$ + $X'$, substituting in the discretized equations, and neglecting products second order or higher in the primed quantities. The basic state represented by $\bar{X}$ is some type of ensemble mean such that $\dot{X} = \bar{X}$, $\dot{X}' = 0$, and $\ddot{X}/\partial t = 0$. The basic state is supposed to be a solution to the steady nonlinear equations with forcing $\bar{F}_f$, $\bar{F}_d$, and $\bar{Q}$.

In order for the matrix governing the linear system to have the desired quasi-block-tridiagonal structure, the complete system (1)-(5) is used; $\bar{\sigma}$ and $\bar{\phi}$ are not eliminated from (1)-(3) using (4) and (5) as in the prognostic model. Additionally, we note that the hydrostatic equation system can be made tridiagonal by the use of variables $\phi_k$ (the interface geopotentials) and in place of $\phi_k$ (the layer geopotentials). The relation

$$\phi_k = RT_k + \frac{1}{\Delta_k} (\bar{\phi}_{k+1} \bar{\sigma}_{k+1} - \bar{\phi}_k \bar{\sigma}_k)$$  \hspace{1cm} (14)$$

is found in Sela (A.18) and leads to (13) when summed from $k = 1$ to $K$. Substituting (14) into (12) yields the desired tridiagonal system. The appropriate form of the hydrostatic equation for the linear model will depend on the form used in the GCM.

The steady linearized vertically sparse system is

$$\alpha'(A_k' + B_k') = F_k'$$ 

$$-\alpha'(B_k' - A_k') + \nabla^2 \left[ E_k' + RT_k' \right] + \frac{1}{\Delta_k} \left( \bar{\sigma}_{k+1} \bar{T}_{k+1} - \bar{\sigma}_k \bar{T}_k \right) = F_d_k'$$  \hspace{1cm} (17)$$

\[ \bar{v}_k \cdot \nabla T_k' + \bar{v}_k' \cdot \nabla \bar{T}_k' + \frac{1}{2\Delta_k} \left[ \bar{\sigma}_k \left( \frac{\sigma_k}{\sigma_{k+1}} \right) ^* \bar{T}_{k+1} - \bar{T}_k' \right] \]

$$+ \bar{\sigma}_{k-1} \left[ T_k' - \left( \frac{\sigma_k}{\sigma_{k-1}} \right) ^* \bar{T}_{k-1} \right]$$

$$+ \bar{\sigma}_{k} \left[ \left( \frac{\sigma_k}{\sigma_{k+1}} \right) ^* \bar{T}_{k+1} - \bar{T}_k \right]$$

$$+ \bar{\sigma}'_{k-1} \left[ \bar{T}_k - \left( \frac{\sigma_k}{\sigma_{k-1}} \right) ^* \bar{T}_{k-1} \right]$$

$$- \kappa T_k (\bar{v}_k \cdot \nabla q' + \bar{v}_k' \cdot \nabla \bar{q}) - \kappa T_k' (\bar{v}_k \cdot \nabla \bar{q}) = \frac{1}{\bar{c_p}} Q'_k$$  \hspace{1cm} (18)$$

$$\bar{v}_k \cdot \nabla q' + \bar{v}_k' \cdot \nabla \bar{q} + D_k' + \frac{\bar{\sigma}'_{k-1} - \bar{\sigma}_{k-1}}{\Delta_k} = F_q'$$  \hspace{1cm} (19)$$

3. Solution algorithm

The solution algorithm is essentially that of Lindzen and Kuo (1969), modified to take into account the influence of $q$ at all levels.

Using a tilde to denote a transpose, define

$$\tilde{\xi}_k = (\xi'_k, D'_k, T'_k, \bar{\phi}_k, \bar{\sigma}_k) \hspace{0.5cm} k = 1, \ldots, K - 1$$

$$\tilde{I}_k = (\bar{\xi}_k, D'_k, T'_k, \phi'_k, q')$$  \hspace{1cm} (25)$$

In the case of spectral horizontal discretization, the subvectors of $f_k$ contain the spectral coefficients of the variables at level $k$ in some order.

The system of Eqs. (16)-(24) can then be written

$$A_1 f_1 + B_1 f_2 + D_1 f_K = F_1$$  \hspace{1cm} (26)$$
\[ A_k f_{k-1} + B_k f_k + C_k f_{k+1} + D_k f_k = F_k, \]
\[ k = 2, \ldots, K - 2 \] (27)
\[ A_{K-1} f_{K-2} + B_{K-1} f_{K-1} + C_{K-1} f_{K-1} = F_{K-1} \] (28)
\[ A_k f_{k-1} + B_k f_k = F_k \] (29)

where
\[ \tilde{F}_k = (F'_\xi, F'_\eta, Q'_\xi, 0, F'_\zeta). \] (30)

A solution of the form
\[ f_k = \alpha_k f_{k+1} + \beta_k + \gamma_k f_k, \quad k = 1, \ldots, K - 2 \] (31)
\[ f_{K-1} = \alpha_{K-1} f_k + \beta_{K-1} \] (32)
is assumed. First, the \( \alpha_k, \beta_k, \) and \( \gamma_k \) are calculated. For \( k = 1, (26) \) and (31) give
\[ \alpha_1 = -A_1^{-1}B_1, \quad \beta_1 = A_1^{-1}F_1, \]
\[ \gamma_1 = -A_1^{-1}D_1. \] (33)

The remaining \( \alpha_k, \beta_k, \gamma_k \) are determined recursively from (27) and (31) for \( k = 2, \ldots, K - 2 \) and are given by
\[ \alpha_k = -(A_k \alpha_{k-1} + B_k)^{-1}C_k \]
\[ \beta_k = (A_k \alpha_{k-1} + B_k)^{-1}(F_k - A_k \beta_{k-1}) \]
\[ \gamma_k = -(A_k \alpha_{k-1} + B_k)^{-1}(A_k \gamma_{k-1} + D_k). \] (34)

It follows from (28) and (32) that
\[ \alpha_{K-1} = -(A_{K-1} \alpha_{K-2} + B_{K-1})^{-1}(A_{K-1} \gamma_{K-2} + C_{K-1}) \]
\[ \beta_{K-1} = (A_{K-1} \alpha_{K-2} + B_{K-1})^{-1}(F_{K-1} - A_{K-1} \beta_{K-2}). \] (35)

Finally
\[ f_k = (A_k \alpha_{K-1} + B_k)^{-1}(F_k - A_k \beta_{K-1}) \] (36)
is found from (29) and (32). Given \( f_K, \) (32) is used to determine \( f_{K-1}, \) and then (31) is applied to complete the solution.

4. Implementation

The new method described here should only be applied to appropriate problems. If the model equations are second order finite differences in the horizontal, a sparse matrix algorithm could be designed to diagonalize the system in either a horizontal or vertical direction. The direction with the largest number of grid points should be chosen; longitude in a typical grid point GCM. For semispectral discretization in the horizontal, a tridiagonal matrix cannot be obtained for zonally asymmetric basic states by structuring the problem by zonal wavenumber, but sparse systems can be designed by organizing the equations along either the vertical or meridional directions, using the one that is better resolved. In the case of fully spectral horizontal discretization, only the vertical direction is available for developing a quasi-tridiagonal system. A scheme derived along the lines described above, where details may differ depending on the vertical differencing scheme, is then an alternative to the one described by Hoskins and Karoly (1981).

In the case of the steady linear model with fully spectral horizontal differencing, the choice of algorithm will be based on relative efficiency [a discussion of the spectral technique and appropriate references are given in Sela (1980)]. Operation counts provide a good estimate of relative performance. If there are \( N \) spectral coefficients for each variable in the horizontal and \( L \) vertical levels, the Hoskins and Karoly (1981) scheme (method 1) requires approximately \([N(3L + 1)]^3/3 \sim 9(NL)^3\) operations to solve the system, since there are three variables (\( \xi, D, T \)) defined at \( L \) levels and one variable (\( q \)) defined at one level. The scheme derived here (method 2) requires five variables with \( N \) coefficients at \( L \) levels. Assuming that matrix inversion and multiplication each take \( O(5N^3) \) operations, the number of operations required to solve the expanded system is \( O([5L(5N)^3]) \). Advantage can be taken of the structures of the \( A_k \) and \( C_k \) matrices to significantly reduce this estimate. These structures and those of \( B_k \) and \( D_k \) are schematically

\[
A_k = \begin{bmatrix}
\times & \times & 0 & 0 & \times \\
\times & \times & 0 & 0 & \times \\
0 & 0 & x & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\] (37)

\[
B_k = \begin{bmatrix}
\times & \times & \times & 0 & \times \\
\times & \times & \times & \times & \times \\
0 & 0 & 0 & 0 & 0 \\
\times & \times & 0 & 0 & 0 \\
\end{bmatrix}
\] (38)

\[
C_k = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\] (39)

\[
D_k = \begin{bmatrix}
0 & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times \\
\end{bmatrix}
\] (40)

where each subblock is \( N \times N \). \( \times \) denotes a full subblock, \( \backslash \) a diagonal subblock, and \( 0 \) denotes a subblock will all elements zero. The matrix multiplications involving \( A_k \) and \( C_k \) can be programmed to avoid multiplication by zero for the \( \backslash \) and \( 0 \) subblocks and reduce the operation count to \( O(3L(5N)^3) \). Then the
ratio of the number of operations required for method 1 to the number required for method 2 is approximately $3L^2/125$. The two schemes are comparably efficient for $L \sim 6.5$. For many more than seven levels, method 2 is more efficient. For example for $L = 18$, method 2 is a factor of 8 more efficient than method 1.

Also, as the total number of variables becomes large, the time required to calculate the matrix elements becomes an important consideration. Using the crude estimate that the number of operations required for computation of the matrices is proportional to the total number of elements minus the number of known zero elements (if the transform technique is used as in method 1), the number of operations for matrix computation for method 1 is $O((3NL)^2)$. For method 2 the result is $O(30N^2L)$, as can be seen from (37)–(40). Then the ratio of method 1 to method 2 operations for matrix generation is about $L/3$. Additionally, method 2 does not require storage of the full matrix, since in the application of the solution algorithm only the parts of the matrix at the level being reduced are required to generate the factors $\alpha_k, \beta_k, \gamma_k$, and only factors need be saved $[O(50N^2L)$ numbers].

We note that although the equations governing the method 1 and method 2 spectral systems are formally equivalent for infinite spectral truncation, the way the transform technique is commonly applied in spectral GCMs leads to systems that are not equivalent. In practice, the GCMs are not consistent in the philosophy of applying the transform technique to exactly evaluate the dynamical terms [those that correspond to the ones on the left-hand side of (1)–(4)] to the spectral truncation of the input fields. This inconsistency occurs in two areas. First, the term involving $q$ in (3) involves a triple product of spectral series, while the Gaussian grid is usually chosen for exact evaluation of quadratic or lower order terms only. Second, evaluation of $\hat{r}$ using (11) is performed on the Gaussian grid, effectively producing $\hat{r}$ with twice the number of zonal and meridional harmonics of the input prognostic fields. Again, the Gaussian grid that would be necessary for exact evaluation of the vertical advection terms in (1)–(3) is that for triple products of variables with the truncation of the prognostic variables, while the grid used is only quadratically accurate.

On the other hand, a spectral truncation for $\hat{r}$ that is the same as that of the prognostic variables has been assumed in implementing method 2 [and in (25)]. To be consistent on the quadratic Gaussian grid, the GCM implementation of the transform technique should take the same approach. This would involve spectrally truncating the $\hat{r}$ obtained from (10) and (11), transforming this spectrally truncated field back to the quadratic Gaussian grid, and only then computing the tendencies in (1)–(3). The total derivative of $q$ in (3) should also be replaced by $-(D + \partial\hat{r}/\partial\sigma)$ using (4).

5. Discussion

We have applied our method for solution of the steady linearized primitive equations to a model with 18 vertical levels and spectral horizontal discretization, using a zonally symmetric basic state. The matrices were computed using the interaction coefficient approach. For rhomboidal 7 (R7) spectral truncation the solutions found using this method were compared to those found using the Hoskins and Karoly approach. Despite the problem of different effective spectral truncations of $\hat{r}$ discussed above, the two solutions were very similar. The solutions using the Hoskins and Karoly method were computed on a CRAY X-MP/48, while those using our scheme were found on an IBM 4381. Eight byte precision for real numbers was necessary on the IBM. Solutions for R40 have also been found on the IBM, although at this resolution a supercomputer becomes convenient. The method developed here has not yet been applied to a zonally asymmetric basic state linear model.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grant ATM-8512524 and Grant ATM-8719167.

REFERENCES


