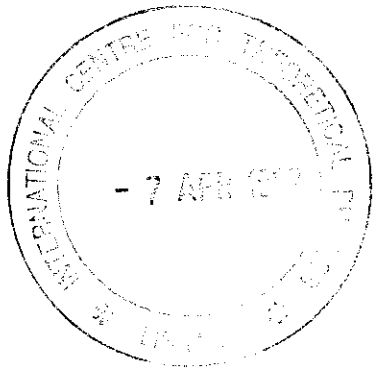
0 000 000 046073 K

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

SMR/941 - 1

"Third ICTP/WMO International Workshop on Tropical Limited Area Modelling "
21 October - 1 November 1996



"Dynamical Aspects of Regional Modelling"

M. NAGATA
Japan Meteorological Agency
Tokyo
Japan

Please note: These are preliminary notes intended for internal distribution only.

STRADA COSTIERA 11 TELEX 460392

Dynamical Aspects of Regional Modeling

Masashi Nagata, JMA

14:00 - 15:00 Mon. 21 October 1996

<< Tropical LAM Workshop in Trieste, Italy >>

1. Introduction

Since regional modeling to be discussed here occupies a portion of the atmospheric modeling field, its basic principle has no essential difference from that of the other portion of the atmospheric modeling field, i.e., global modeling. They are both modeling of physical stetes of the atmosphere based on physical laws governing the atmospheric motion and state. Since the physical laws are given in the form of non-linear partial differential equations, their solutions are calculated numerically. This is why the weather prediction is called numerical weather prediction.

(Discretization)

The most common way of solving the equations is to find approximate expressions defined using values at grid points or amplitudes of base functions at various wave numbers. The discretization is done both in the horizontal and in the vertical. Most widely used in the horizontal discretization are a finite difference method and a spectral method. For the vertical discretization, a multi-layer modeling (finite difference method) is commonly used. We call the model using finite difference method in the horizontal discretization a grid point model while the one using spectral method a spectral model. For the time differencing, finite difference method is widely used.

Finite difference method has been used for the horizontal discretization since the very beginning of the atmospheric modeling. Compared to it, the spectral method appeared later in the numerical weather prediction (NWP) field. It was first successfully applied to global NWP in 1970s. However, it had not been applied to regional NWP until mid to late 1980s when the Japan Meteorological Agency (JMA) first succeeded in putting spectral regional models into operation. The much later application of spectral method to regional NWP than that to global NWP was mainly due to the difficulty encountered in handling time dependent lateral boundary conditions which will contradict boundary conditions pertaining to the orthogonal base functions for expansion.

There have been proposed many kinds of finite difference schemes with different characters and different degrees of accuracy for the time differencing as well as for the space differencing (grid point model). On the other hand, there is much smaller room in choosing discretization schemes for the spectral space differencing (spectral model). This is because once you choose a particular set of base functions for a spectral model that means you have already determined the whole space differencing and the lateral boundary conditions as well which are inherent to the base functions.

2. Finite difference method

<Accuracy, consistency, convergence, stability>

In the grid point model (in the finite difference method), space derivatives must be approximated using finite differences. The way how they are approximated determine the degree (order) of accuracy. For example, if $\frac{\partial u}{\partial x}$ is approximated with an uncentered forward difference as:

$$\frac{u_{i+1} - u_i}{\Delta x}, \tag{2-1}$$

its order of accuracy is one, since using the Tailor series:

$$u_{i+1} = u_i + \left(\frac{\partial u}{\partial x}\right)_i \Delta x + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_i \Delta x^2 + \cdots, \tag{2-2}$$

it is expanded as:

$$\frac{\partial u}{\partial x} \approx \frac{u_{i+1} - u_i}{\Delta x} = \left(\frac{\partial u}{\partial x}\right)_i + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_i \Delta x + \cdots, \tag{2-3}$$

and the approximation contains errors including a term with Δx to the 1st order. If we adopt a centered difference scheme as:

$$\frac{u_{i+1} - u_{i-1}}{2\Delta x}, \tag{2-4}$$

the order of accuracy is two.

When we wish to increase the order of accuracy of the approximate expressions, we need to increase the number of terms involved in the approximation, which requires more computation in actual time integration. Thus there is a kind of trade-off relationship between the accuracy you want and the computer resources you need to attain such accuracy.

Consistency, which means that the finite difference form of equations approaches the differencial equations when both space and time intervals approach zero, is an basic and important characteristics which an approximate scheme should have. A necesary condition for consistency is that the finite difference scheme has at least the order of accuracy of 1.

Convergency of the numerical solution means that errors associated with the use of approximate expressions approach zero as the grid intervals are reduced to zero. Consistency of a scheme does not necessarily guarantee its convergent character. To guarantee the convergency of a scheme which is adopting a conventional Eulerian type of scheme of time integration, the condition for a simple advective equation:

$$c\,\Delta t \le \Delta x, \tag{2-5}$$

where c is the largest propagation speed of disturbance in the model, is necessary.

The last but most important characteristics which an approximate scheme must have is the stability of the numerical solution. Actually we often examine the boundedness of the error with the Von Neumann method that analyzes the linearized version of the non-linear equations using a single harmonic solution as a representative of all harmonic components of the solution. In applying the method we use an amplification factor defined by:

$$u^{n+1} \equiv u^n \lambda, \tag{2-6}$$

and consider that $|\lambda| \le 1$ guarantees the necessary condition for stability of the solution. If we apply the method to the simple advection equation,

$$c\,\Delta t \le \Delta x \tag{2-7}$$

is again the necessary condition for stability of the solution, which is the same as that for the necessary condition for convergency.

The same discussions above except for those of accuracy are applicable to a spectral model. [In a spectral model, accuracy associated with the spatial discretization is solely determined by the truncation of the base functions we have chosen.]

<Time differencing schemes>

We consider the equation:

$$\frac{du}{dt} = f(u,t), \quad u = u(t). \tag{2-8}$$

There can be also a wide variety of schemes for time differencing. However, less effect has been devoted to using a complicated higher-order schemes, compared to in the case of space differencing. One of the major reasons for this may be the fact that there is usually insufficient information about the initial conditions, which may become larger sources of errors, which probably waste the effort. Another reason may be that errors associated with spatial differencing are much larger than those due to the time differencing. This, of course, does not mean that we do not need to care about the properties of time differencing schemes.

- Number of time levels of the scheme

Schemes using values at two time levels n and $n+1$, such as the Euler (forward) scheme:

$$u^{n+1} = u^n + \Delta t\, f^n \tag{2-9}$$

are called 'two-time-level scheme', while those using values at three time levels $n-1$, n and $n+1$ are called 'three time-level scheme'. To define time derivatives of dependent variables, values at two time levels are necessary and sufficient. Therefore, if we use a three time-level scheme, such as the leap-frog scheme:

$$u^{n+1} = u^{n-1} + 2\Delta t\, f^n, \tag{2-10}$$

an extra (computational) mode will appear besides the physically meaningful mode. When the extra mode has an amplification factor greater than unity, we have to adopt some treatment suppressing the mode, such as the time filtering by Asselin (1972).

- Explicit and implicit schemes

When we approximate the right-hand side of (2-8): f without using values at future time level $n+1$, the scheme is called explicit scheme, while when we approximate f using values at future time level $n+1$, the scheme is called implicit scheme. In general when we use an implicit scheme, we have to solve the whole time difference equations simultaneously.

The properties including stability of the time differencing schemes depend on the form of the equation. The trapezoidal implicit scheme:

$$u^{n+1} = u^n + \Delta t\,(\tfrac{1}{2} f^n + \tfrac{1}{2} f^{n+1}) \tag{2-11}$$

is always neutral when it is applied to the oscillation equation. This property is utilized in the widely used semi-implicit scheme integrating high-frequency gravity modes with a much longer time step than the time constraint for explicit schemes. The stabiltiy of a scheme for a particular form of equation does not necessarily guarantee the stability of the scheme for other forms of equations. For example, both of the two modes in the leap-frog scheme are stable and neutral in the oscillation equation if the stability condition:

$$\left|\frac{c\Delta t}{\Delta x}\right| < 1 \tag{2-12}$$

is satisfied, while the computational mode is unstable for any Δt in the friction equation. Therefore, it is not suitable for numerical integration of the friction equation while it is widely used for that of the oscillation equation such as the advection equation.

For the friction terms, we use stable schemes, such as the backward implicit scheme, while we use the leap-frog scheme for the advective terms, forming a combination of time integration schemes.

<Phase error and comutational dispersion>

When we use a second-order centered differece scheme for the advective term, in

the advection equation:

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \tag{2-13}$$

waves propagate with the phase speed which is a function of the wave number k:

$$c^* = c\frac{\sin k\Delta x}{k\Delta x}. \tag{2-14}$$

Thus, the finite differencing in space has produced a dispersion of waves (computational dispersion), which nature is not included in the original differential equation. As $k\Delta x$ increases from zero (longest wave length), the phase speed c^* decreases monotonically from c to become zero at the shortest wave length of $2\Delta x$, where $k\Delta x = \pi$. This means that the shortest resolvable wave is stationary.

There are two problems in using the finite difference scheme. One is the retardation of advective processes. The other is the false dispersion, which will result in a deformation of disturbances, especially those of smaller scales, such as tropical cyclones, fronts, and shear lines, which are important for regional (mesoscale) weather predictions and are major targets of limited-area models.

A sprectral model can escape from the phase error and the computational dispersion arising from the spatial discretization. Meanwhile it has a problem in representing very sharp gradients like step functions, which may produce the Gibbs phenomenon in the discretization.

<Aliasing error and non-linear instability>

Since the atmospheric equations are generally non-linear, waves with wave number greater than that resolved can be generated at every time step. These waves are projected onto the waves that can be resolved. This is the aliasing error, which can produce false change of energy spectrum, especially around the marginally resolved wave number:

$$k_{max} = \frac{\pi}{\Delta x} \tag{2-15}$$

The false energy change can cause a blowup of the model (non-linear instability). The Arakawa Jacobian succeeded in solving the instability problem in an elegant way, considering the conservation properties of kinetic energy and enstrophy (square of vorticity). We do not need to use any artificial dissipation scheme to prevent the instability when we use the Arakawa Jacobian.

In a spectral model, we do not need to care about the same instability problem if numbers of transform grid points are determined so that all the aliased waves are projected outside the wave number range which can be represented.

<Integral properties conservation schemes>

Numerical atmospheric models are encouragingly advised to adopt schemes which have properties conserving physically important integrals, such as total energy, potential temperature, and enstrophy when simplified to a two-dimentional non-divergent model.

<Horizontal distribution of variables on the grid in light of simulation of geostrophic adjustment>

There are several possible schemes of distribution of prognostic variables over the horizontal grid in a grid point model. Properties in the simulation of geostrophic adjustment are different from scheme to scheme. Schemes which have a property as close to that of the differential equation as possible are good. Arakawa and Lamb (1977) recommended that the Arakawa C grid be used for atmospheric modeling from the view point of the simulation of geostrophic adjustment.

<Vertical distribution of variables on the grid in light of simulation of baroclinic waves>

There are a few possible schemes of distribution of prognostic variables over the vertical grid in both grid point model and spectral model. Arakawa and Moothi (1988) and Arakawa and Konor (1996) discussed that the Charney-Phillips grid on which temperature is defined at the same level as vertical velocity and stagggerd from horizontal velocity is more appropriate than the Lorentz grid on which temperature is staggered from vertical velocity and defined at the same level as horizontal velocity in terms of simulating baroclinic waves, although the former loses conservation property of the volume integral of potential temerature while the latter can keep it. That is because the latter scheme includes a computational mode arising from the averaging of potential temperature in the vertical differencing and making a stationary zigzag vertical distribution of potential temperature and an excess degree of freedom in potential vorticity, leading to spurious potential vorticity disturbances resulting in a growth of small-scale features.

3. Spectral model

In a spectral model, we prepare a series of base functions and use them to express fields of dependent variables of the modeled atmosphere. There are two conditions for the base functions. First, the functions are orthogonal to each other so that we can easily and uniquely decompose the fields onto the wave number space. Second, the functions satisfy the boundary conditions. One major advantage of prepresenting fields with analytical functions is that once a field is expressed on the wave number space, spatial derivatives are automatically given in analytical forms of the base functions, which do not include any approximation. Other advantages of the spectral model are:

- Several integral constraints, such as conservation of mass, energy and potential temperature, are held without any special treatment like those in the finite difference

schemes of the grid point method.

- There is not either phase retardation, associated false dispersion, or non-linear instability without any special treatment.
- The coding of numerical calculations can be very simple since there is no need to approximate the horizontal space derivatives with finite differences.
- A smaller number of waves than that of grid points can express spatial variations of physical quantities with the same accuracy.
- The semi-implicit scheme is much more efficiently used since the Helmholtz equation can be solved directly and easily because the base functions are the eigenfunctions of the Laplacian operator. This property works to save much computation.

<Wave number truncation>

Since the atmospheric equations are non-linear, the expansion base functions interact with each other. Waves with wave number n and m interact to produce waves with wave number $n+m$ and $n-m$ at every time step. We have to specify an upper limit of the wave number (truncation wave number) beyond which waves are not handled in the spectral model. Once we determine the truncation wave number, waves with wave number larger than this are automatically eliminated from the model.

<Expansion functions>

For global models, the trigonometric functions are used for expansion in the longitudinal (zonal) direction and the associated Legendre functions in the latitudinal (meridional) direction, which compose the spherical harmonics and form a set of solutions of the Laplace equation on the sphere. For a regional model, double Fourier series are used, which are also a set of solutions of the same equation. These expansion functions have the following advantages:

- Transformation and inverse transformation between physical space and wave-number space can be done economically with the FFT method.
- The semi-implicit time integration scheme:

$$\frac{u^{n+1} - u^{n-1}}{2\Delta t} = i\omega_R u^n + \frac{i\omega_G(u^{n+1} + u^{n-1})}{2}, \tag{3-1}$$

where ω_R is the frequency of a representative Rossby mode and ω_G is that of a representative gravity-wave mode,

can be used in a very economical way, since the scheme is rearranged into the Helmholtz equation including the Laplacian operator.

<Interaction coefficient method and transform method>

We consider a one-dimensional non-linear equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + a\frac{\partial u}{\partial x} = 0 \tag{3-2}$$

where a is a constant.

The second term and the third term on the left-hand side is a non-linear and a linear term, respectively. To solve this equation by the spectral method, first we transform it into the spectral (wave-number) space using:

$$u(x) = \sum_m U(m)e^{imx} \tag{3-3}$$

$$U(m) = \frac{1}{2\pi}\int_0^{2\pi} u(x)e^{-imx}dx \tag{3-4}$$

$$\frac{\partial u}{\partial x} = \sum_m imU(m)e^{imx} \tag{3-5}$$

$$\frac{\partial u}{\partial t} = \sum_m \frac{\partial U(m)}{\partial t}e^{imx} \tag{3-6}$$

as:

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial u}{\partial t}e^{-imx}dx + \frac{1}{2\pi}\int_0^{2\pi} u\frac{\partial u}{\partial x}e^{-imx}dx + \frac{1}{2\pi}\int_0^{2\pi} a\frac{\partial u}{\partial x}e^{-imx}dx = 0, \tag{3-7}$$

for wave number m:

$$\frac{\partial U(m)}{\partial t} + \sum_k ikU(k)\frac{1}{2\pi}\int_0^{2\pi} u(x)e^{-i(m-k)x}dx + imaU(m) = 0 \tag{3-8}$$

$$\frac{\partial U(m)}{\partial t} + \sum_k ikU(k)U(m-k) + imaU(m) = 0 \tag{3-9}$$

This is the equation in the spectral form (in the wave number space).

In this form, $U(m)$ is independent of space (x), dependent of time (t). Therefore, it is free from the error originating from space differencing. To solve the equation (3-9), we simply approximate the time derivative with a suitable time differencing scheme and perform the time integration with a proper time step.

The second term on the left-hand side of (3-9) shows that two waves with wave

number k and m - k, whose sum makes m, interact and produce a wave with wave number m. The number of combination of such waves is quite large. On the other hand, the linear third term on the left-hand side of (3-9) is composed of a single wave with wave number m and has no interaction with waves with different wave numbers.

In the case of one-dimensional model we can easily calculate the nonlinear term. However, once we apply spectral method (interaction coefficient method) of this type to a two-dimensional case, the situation changes drastically. The transformation and inverse transformation are expressed with double summation and double integration. If we derive the spectral form of the two-dimensional non-linear equation, the non-linear term is written in a form including tetra-summation. Due to this, the number of combination of waves which compose a particular wave is enormously larger than that in a one-dimensional case. Thus, the number of calculations required by the interaction coefficient method is so huge that this type of spectral method has not yet been applied to numerical weather prediction.

In 1970, a new method was developed which dramatically decreases the number of calculations of the non-linear terms but preserves the accuracy of the calculations of the terms. That is the transform method. With this method, the non-linear second term on the left-hand side of (3-7) is not transformed into the spectral space but left unchanged:

$$\frac{\partial U(m)}{\partial t} + \frac{1}{2\pi}\int_{0}^{2\pi} u\frac{\partial u}{\partial x}e^{-imx}dx + imaU(m) = 0 \qquad (3\text{-}10)$$

To calculate the second term on the left-hand side, first we calculate u and $\frac{\partial u}{\partial x}$ on appropriate grid points by making inverse transformation:

$$u(x) = \sum_{m=-M}^{M} U(m)e^{imx} \qquad (3\text{-}11)$$

$$\frac{\partial u}{\partial x} = \sum_{m=-M}^{M} imU(m)e^{imx} \qquad (3\text{-}12)$$

then they are multiplied at the grid points and finally integrated after being multiplied by e^{-imx} with respect to x from 0 to 2π.

Thus, the transform method can ingeneously escape from huge calculations associated with interactions among waves and has been used quite successfully in the atmospheric modeling including numerical weather prediction.

<Sufficient number of grid points to avoid aliasing>

We must introduce sufficient number of grid points for transformation to avoid

aliasing error arising from the non-linear term. The maximum wave number which can be expressed by I grid points is I/2 and a wave whose wave number m* exceeds I/2 is misrepresented as a wave with wave number

$$2 \cdot \frac{I}{2} - m^*. \qquad (3\text{-}13)$$

To avoid the aliasing error, this wave must have a wave number greater than the truncation wave number M,

$$2 \cdot \frac{I}{2} - m^* > M \qquad (3\text{-}14)$$

Here m* may have wave number 2M from the quadratic non-linear term. Therefore,

$$I > 3M \qquad (3\text{-}15)$$

is the sufficient condition for avoiding the aliasing error arising from the quadratic non-linear term.

4. A regional spectral model

Finite difference scheme had been used for dynamical part of most limited-area models, while the spectral method has been used widely in global models in many NWP centers. Although the spectral method has several advantages, it had not been applied to limited-area NWP models with time-dependent lateral boundary conditions owing to a difficulty in properly formulating the lateral boundary conditions. Familiar basic function series, such as the double Fourier series, do not satisfy a prescribed time-dependent lateral boundary condition. To apply the spectral method to regional NWP, we have to find some treatment method which meets the following two requirements simultaneously:

- using base function series which satisfy idealized boundary conditions and
- incorporating prescribed time-dependent lateral boundary conditions.

In this section some essential points of the JMA regional spectral model is described.

4.1 Spectral representation and prognostic equations for a free-slip wall boundary condition

Using the double Fourier series, dependent variables are projected onto the spectral (wave-number) space as:

$$U_{o\,kl} = \varepsilon_k\varepsilon_l\int_{0}^{\pi}\int u^* \sin k\hat{x}\cos l\hat{y}\, d\hat{x}\, d\hat{y} \qquad (4\text{-}1)$$

$$V_{o\,kl} = \varepsilon_k \varepsilon_l \int_0^{\pi}\int_0^{\pi} v^* \cos k\hat{x} \sin l\hat{y}\, d\hat{x}\, d\hat{y} \tag{4-2}$$

$$S_{o\,kl} = \varepsilon_k \varepsilon_l \int_0^{\pi}\int_0^{\pi} T \cos k\hat{x} \cos l\hat{y}\, d\hat{x}\, d\hat{y} \tag{4-3}$$

$$\Pi_{o\,kl} = \varepsilon_k \varepsilon_l \int_0^{\pi}\int_0^{\pi} \pi^* \cos k\hat{x} \cos l\hat{y}\, d\hat{x}\, d\hat{y}, \tag{4-4}$$

where $\hat{x}$ and $\hat{y}$ are horizontal coordinates normalized with the lengths of the forecast domain L_x and L_y:

$$\hat{x} = \frac{\pi x}{L_x}, \qquad \hat{y} = \frac{\pi y}{L_y} \tag{4-5}$$

and k and l are the integer wave number in x and y direction, respectively, and ε_n is a normalized factor defined as:

$$\varepsilon_n = \begin{cases} \frac{1}{\pi} & \text{for } n = 0 \\ \frac{2}{\pi} & \text{for } n = 1,2,3, \text{ - - -} \end{cases} \tag{4-6}$$

The above representation satisfies a free slip wall boundary conditions. The bases within the four sets of double Fourier series are orthogonal to each other in the rectangular forecast domain and each set satisfies different lateral boundary conditions. All the bases of every set are eigenfunctions of the Lalacian operator in the rectangular domain on a conformal map projection. This property provides a great advantage to us in solving the semi-implicit time integration, as shown later.

We use matrix expression of dependent variables a and A in the physical (grid point) space and the spectral (wave number) space, respectively. The transformation and the inverse transformation is simply written as:

$$\underline{A}_o = F_o \underline{a}, \qquad \underline{a}_o = F_o^{-1} \underline{A}_o \tag{4-7}$$

where the subscript o denotes the component which satisfies the wall boundary conditions.

Thus, F_o eliminates non-orthogonal components which may exist in a. Actually the orthogonal winds have no components across the boundary and the orthogonal T and π^* have no gradients normal to the boundary.

If we define differential operators in the spectral space as:

$$F_0^{-1} D_x \underline{A}_0 \equiv \frac{\partial \underline{a}_0}{\partial x}, \qquad F_0^{-1} D_y \underline{A}_0 \equiv \frac{\partial \underline{a}_0}{\partial y}, \tag{4-8}$$

we have

$$D_x U_{o\,kl} = \frac{\pi}{L_x} k U_{o\,kl}, \qquad D_y U_{o\,kl} = -\frac{\pi}{L_y} l U_{o\,kl}. \tag{4-9}$$

Using these relations we can convert the prognostic equation in the physical space

$$\frac{\partial}{\partial t}\underline{a} = a_t\left(\underline{a}, \frac{\partial}{\partial x}\underline{a}, \frac{\partial}{\partial y}\underline{a}, \frac{\partial}{\partial \sigma}\underline{a}\right) \tag{4-10}$$

into the one in the spectral space for the wall boundary conditions.

$$\frac{\partial}{\partial t}\underline{A}_o = F_o a_t\left(F_o^{-1}\underline{A}_o, F_o^{-1} D_x \underline{A}_o, F_o^{-1} D_y \underline{A}_o, \frac{\partial}{\partial \sigma} F_o^{-1}\underline{A}_o\right) \tag{4-11}$$

4.2 Spectral representation and prognostic equation with a time-dependent lateral boundary condition

The spectral model with the wall boundary conditions have no flexibility to allow time-dependent lateral boundary conditions, which are essential to actual limited-area modeling. To treat the difficulty, we introduce non-orthogonal additional bases for four sets of wall-boundary orthogonal bases. Thus, the new sets of bases consists of the wall-boundary orthogonal bases and the additional non-orthogonal bases. Specifically for the Fourier sine series in a one-dimensional case, we introduce cosine functions of small wave numbers:

$$f_N(\hat{x}) = F_0 \cos 0\hat{x} + F_{-1} \cos 1\hat{x}, \quad F_N = (F_0, F_{-1})^T \tag{4-12}$$

to represent the non-orthogonal component $f_N(\hat{x})$ with its values at the lateral boundaries:

$$f_N(0) = f(0), \quad f_N(\pi) = f(\pi) \tag{4-13}$$

Using this, the spectral representation of $f(\hat{x})$: F is

$$F = F_N + F_o = F_N + F_o(f(\hat{x}) - f_N(\hat{x})), \tag{4-14}$$

which is the modified Fourier sine series that includes additional non-orthogonal components. From the definition of f_N,

$$f_N(0) = F_0 + F_{-1}$$

$$f_N(\pi) = F_0 - F_{-1} \tag{4-15}$$

The lateral boundary conditions are pojected on to the spectral space as:

$$F_0 = \frac{1}{2}(f_N(0) + f_N(\pi))$$

$$F_{-1} = \frac{1}{2}(f_N(0) - f_N(\pi)) \tag{4-16}$$

For a modified Fourier cosine expansion, we introduce additional bases specifying non-zero $\frac{\partial}{\partial x}$ at the boundary:

$$g_N(\hat{x}) = G_{-1} \sin 1\hat{x} + G_{-2} \sin 2\hat{x}, \quad G_N = (G_{-1}, G_{-2})^t \tag{4-17}$$

to represent the non-orthogonal component $g(\hat{x})$ with its first derivative at the lateral boundary:

$$\frac{\partial g_N}{\partial x}(0) = \frac{\partial g}{\partial x}(0), \quad \frac{\partial g_N}{\partial x}(\pi) = \frac{\partial g}{\partial x}(\pi). \tag{4-18}$$

Using this, the spectral representation of $g(\hat{x})$: G is

$$G = G_N + G_o = G_N + F_o(g(\hat{x}) - g_N(\hat{x})) \tag{4-19}$$

which is the modified Fourier cosine series that includes additional non-orthogonal components. From the definition of g_N,

$$\frac{\partial g_N}{\partial \hat{x}}(0) = G_{-1}\cos 0 + G_{-2}\cos 2\cdot 0 = G_{-1} + 2G_{-2}$$

$$\frac{\partial g_N}{\partial \hat{x}}(\pi) = G_{-1}\cos \pi + G_{-2}\cos 2\cdot\pi = -G_{-1} + 2G_{-2} \tag{4-20}$$

The lateral boundary conditions are pojected on to the spectral space as:

$$G_{-1} = \frac{1}{2}(\frac{\partial g_N}{\partial \hat{x}}(0) - \frac{\partial g_N}{\partial \hat{x}}(\pi))$$

$$G_{-2} = \frac{1}{4}(\frac{\partial g_N}{\partial \hat{x}}(0) + \frac{\partial g_N}{\partial \hat{x}}(\pi)) \tag{4-21}$$

There are a wide variety of functional forms to be used for additional bases. From the practical point of view, our choice of small wave-number sinusoidal functions has the following advantages:

- all operations and calculations can be done in the wave-number space
- the Fast Fourier Transform (FFT) algorithm is efficiently used
- the model structure can be simplified resulting in computational economy.

For the two-dimensional case, the same procedure applies except for double Fourier series. The correspondence of dependent variables to the modified Fourier series is:

c - c* : T, π*
c - s* : v*
s - c* : u*
s - s* : q (4-22)

We specify the following boundary conditions:

values u* at $\hat{x} = 0, \pi$
values v* at $\hat{y} = 0, \pi$
values q at the lateral boundaries
gradients of T and π* normal to the boundaries.

Other quantities, such as u* at $\hat{y} = 0, \pi$, are predicted in the model.

To reduce the spurious solution arising from the ill-posedness, we apply the boundary relaxation technique to a frame region with some width along the lateral boundaries.

The spectral prognostic equation can be derived from the definition of the modified Fourier series:

$$\underline{A}_o(t) = F_o(\underline{a}(t) - F_N^{-1}\tilde{\underline{A}}_N(t)), \tag{4-23}$$

where $\tilde{\underline{A}}_N(t)$ is the prescribed values at the lateral boundaries and F is the modified Fourier transform operator. By differenciating this with respect to time, we have

$$\frac{\partial}{\partial t}\underline{A}_o(t) = F_o(\frac{\partial}{\partial t}\underline{a}(t) - F_N^{-1}\frac{\partial}{\partial t}\tilde{\underline{A}}_N(t)). \tag{4-24}$$

Using the original prognostic equations, this can be written as:

$$\frac{\partial}{\partial t}\underline{A}_0(t) = F_0(\underline{a}_t(F^{-1}\underline{A}, F^{-1}D_x\underline{A}, F^{-1}D_y\underline{A}, \frac{\partial}{\partial\sigma}F^{-1}\underline{A})) - F_0F_N^{-1}\frac{\partial}{\partial t}\tilde{\underline{A}}_N(t) \qquad (4\text{-}25)$$

Note here that the effects of the time-dependent non-orthogonal lateral boundary conditions are incorporated into the prognostic equation as a convolution. Thus, we have obtained the prognostic equations of the orthogonal components with the time-dependent lateral boundary conditions included.

4.3 Aliasing-free transform grid and wave number truncation

We use the transform method to evaluate the non-linear terms of the prognostic equations. To eliminate aliasing error, the number of grid points has to be large: Since the size of the full domain from $\hat{x} = 0$ to $\hat{x} = 2\pi$ is $2(I-1)$ in number of grid points, the maximum wave number represented by the grid is $\frac{2(I-1)}{2}$. Meanwhile the quadratic non-linear terms, such as $m^2u^*\frac{\partial u^*}{\partial x}$ will yield waves with maximum wave number:

$$2K + K_m \qquad (4\text{-}26)$$

where K is the truncation wave number for the prognostic variables, K_m is the one for m^2 (square of the map factor). To escape from aliasing, aliased waves must be projected outside the truncation wave number as:

$$2 \cdot \frac{2(I-1)}{2} - (2K + K_m) > K.$$

max. wave number represented by grids — wave number of waves yielded by non-linear term — truncation wave number

<----------> wave number of aliased wave

Thus,

$$I > \frac{1}{2}(3K + K_m) + 1 \qquad (4\text{-}27)$$

is the condition for alias-free transform grid.

4.4 Boundary relaxation

We use a Laplacian-type (red spectrum) and/or a Newtonian-type (white spectrum) boundary relaxation to suppress spurious solutions arising from the ill-posedness:

$$\frac{\partial}{\partial t}\underline{a} = a_t(\text{- - - -}) + \alpha\nabla^2(\underline{a}-\tilde{\underline{a}}) - \beta(\underline{a}-\tilde{\underline{a}}),$$

$$\frac{\partial}{\partial t}\underline{A}_0 = \text{- - - -} + F_0(\alpha\nabla^2F^{-1}(\underline{A}-\tilde{\underline{A}}) - \beta F^{-1}(\underline{A}-\tilde{\underline{A}})). \qquad (4\text{-}28)$$

4.5 Integration procedure

(1) Starting with grid-point data $a^*(t)$, first transform them into the spectral form $\underline{A}(t)$, using the prescribed lateral boundary condition $\tilde{\underline{A}}_n(t)$, and then apply the wave number truncation on $\underline{A}_0(t)$.

(2) Calculate $-F_0F_N^{-1}\frac{\partial}{\partial t}\tilde{\underline{A}}_N$ from coarse-mesh-model predicted fields and store them at appropriate time intervals.

(3) Calculate the grid-point values of $\underline{a}$, $\frac{\partial\underline{a}}{\partial x}$ and $\frac{\partial\underline{a}}{\partial y}$ by applying the F^{-1} operator on $\underline{A}$, $D_x\underline{A}$ and $D_y\underline{A}$.

(4) Calculate time tendencies of the prognostic variables $\frac{\partial\underline{a}}{\partial t}$ using $\underline{a}$, $\frac{\partial\underline{a}}{\partial x}$ and $\frac{\partial\underline{a}}{\partial y}$, the latter two of which can be obtained analytically.

(5) Substitute the results of (2) and (4) into the spectral prognostic equation (4-25) to calculate $\frac{\partial}{\partial t}A_0$.

(6) Calculate the next time-step values of $A_0(t+\Delta t)$ with a time integration scheme.

(7) Go back to (3).

Note that physical processes parameterized are incorporated to $\frac{\partial\underline{a}}{\partial t}$ in (4).

The most important feature is that the space derivatives $\frac{\partial\underline{a}}{\partial x}$ and $\frac{\partial\underline{a}}{\partial y}$ are calculated analytically without any approximation.

4.6 Semi-implicit time integration scheme

We apply the semi-implicit time integration scheme which combines the leap-frog scheme for Rossby-wave modes and the trapezoidal implicit scheme for gravity-wave modes:

$$\frac{Q^{n+1} - Q^{n-1}}{2\Delta t} = i\omega_R Q^n + \frac{i\omega_G(Q^{n+1} + Q^{n-1})}{2}. \qquad (4\text{-}29)$$

To simplify expressions, we use the following notations:

$$Q_t = Q(t)$$

$$\bar{Q}^t = \frac{1}{2}(Q_{t+\Delta t} + Q_{t-\Delta t})$$

$$\delta_t Q = \frac{1}{2\Delta t}(Q_{t+\Delta t} - Q_{t-\Delta t})$$

$$\rightarrow \quad \bar{Q}^t = \Delta t\, \delta_t Q + Q_{t-\Delta t} \tag{4-30}$$

and

$$V^* = u^*\mathbf{i} + v^*\mathbf{j}$$

$$D = \nabla \cdot V^* = \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y}$$

$$\zeta = \mathbf{k} \cdot \nabla \times V^* = \frac{\partial v^*}{\partial x} - \frac{\partial u^*}{\partial y} \tag{4-31}$$

The formulations of the scheme are:

$$\delta_t D + \nabla \cdot (\overline{PGF})^t = \nabla \cdot (\frac{\partial V^*}{\partial t})_t + \nabla \cdot (PGF)_t,$$

where $(PGF) \equiv \nabla\phi + RT_0 \nabla\pi^*$,

$$\delta_t T + (\overline{TVA})^t = (\frac{\partial T}{\partial t})_t + (TVA)_t,$$

where $(TVA) = \sigma^\kappa \dot{\sigma} \frac{\partial}{\partial \sigma}(T_0 \sigma^{-\kappa}) - \frac{RT_0}{C_p}(\frac{\partial \pi^*}{\partial t})_D$,

$$\delta_t \pi^* + (\overline{DIV})^t = (\frac{\partial \pi^*}{\partial t})_t + (DIV)_t,$$

where $(DIV) \equiv m_0^2 \int_0^1 D d\sigma$,

$$(\frac{\partial \pi^*}{\partial t})_D \equiv m_0^2 \int_0^1 -D d\sigma = -(DIV),$$

$$\frac{\partial \dot{\sigma}_D}{\partial \sigma} \equiv -m_0^2 D - (\frac{\partial \pi^*}{\partial t})_D . \tag{4-32}$$

Here T_0, which is a function of σ only, is the reference temperature profile and m_0^2 (constant) is a reference square of map factor. By introducing these reference quantities, we can linearize the gravity-wave terms and integrate them with the trapezoidal implicit scheme.

The lateral boundary relaxation terms are included in $\frac{\partial Q}{\partial t}_t$ using $(Q - \bar{Q})$ at $t - \Delta t$ (forward scheme) to avoid the linear computational stability.

Introducing the matrix notation for the vertical differencing schemes of the hydrostatic equation and the thermodynamic equation:

$$\underline{\Phi} = \mathbf{P}\, \underline{T}$$

$$(\underline{TVA}) = \mathbf{H}\, \underline{D}, \tag{4-33}$$

where $\underline{Q} = (Q_1, - - - , Q_K)^T$, $k = 1 - K$ is the index for the vertical level. The constant matrices $\mathbf{P}$ and $\mathbf{H}$ are uniquely determined from the vertical difference schemes. Using these expressions, the prognostic equations become:

$$\delta_t \underline{D} + \Delta t\, \mathbf{P}\, \nabla^2 \delta_t \underline{T} + \Delta t\, R\underline{T}_0 \nabla^2 \delta_t \pi^* = \underline{ND}$$

$$\delta_t \underline{T} + \Delta t\, \mathbf{H}\, \delta_t \underline{D} = \underline{NT}$$

$$\delta_t \pi^* + \Delta t\, m_0^2 \underline{\Delta\sigma}^T \delta_t \underline{D} = N\pi^*, \tag{4-34}$$

where $\underline{ND}$, $\underline{NT}$ and $N\pi^*$ include all other terms than the linearized gravity-wave terms as:

$$\underline{ND} = (\frac{\partial}{\partial t} \nabla \cdot V^*)_t + \mathbf{P}\, \nabla^2 (\underline{T}_t - \underline{T}_{t-\Delta t}) + R\underline{T}_0 \nabla^2 (\pi^*_t - \pi^*_{t-\Delta t})$$

$$\underline{NT} = (\frac{\partial}{\partial t} \underline{T})_t + \mathbf{H}\, (\underline{D}_t - \underline{D}_{t-\Delta t})$$

$$N\pi^* = (\frac{\partial}{\partial t} \pi^*)_t + m_0^2 \underline{\Delta\sigma}^T (\underline{D}_t - \underline{D}_{t-\Delta t}) . \tag{4-35}$$

Eliminating $\delta_t\underline{T}$ and $\delta_t\pi^*$, we obtain the Helmholtz-type equation with respect to $\delta_t\underline{D}$.

$$- \Delta t^2(\mathbf{PH} + m_0^2 R\underline{T}_0\Delta\sigma^T)\nabla^2\delta_t\underline{D} + \delta_t\underline{D}$$

$$= \underline{ND} - \Delta t\, \mathbf{P}\, \nabla^2(\underline{NT}) - \Delta t\, R\underline{T}_0\nabla^2(N\pi^*). \tag{4-36}$$

Up to this point, the formulation of the scheme has no difference between the grid point model and the spectral model.

Then we transform the Helmholtz equation into the spectral space by calculating the spectral expression of $(\frac{\partial Q}{\partial t})_t$ using the spectral prognostic equation with u* and v* converted to D and ζ by applying the D_x and D_y operators. We do not need to take into account the non-orthogonal bases in this conversion.

$$\{\mathbf{I} + \Delta t^2(k^{*2}+l^{*2})\,\mathbf{C}\}\delta_t\underline{D}_{kl} = (\underline{NL})_{kl}. \tag{4-37}$$

where $\nabla^2 Q_{kl} = -(k^{*2}+l^{*2})Q_{kl}$ has been used and $(\underline{NL})_{kl}$ representing the spectral form of the right-hand side of the Helmholtz equation and $\mathbf{I}$ is the unit matrix. The matrix $\mathbf{C}$ is a constant matrix:

$$\mathbf{C} = \mathbf{PH} + m_0^2 R\underline{T}_0\Delta\sigma^T. \tag{4-38}$$

We have prepared the matrix $\mathbf{E}$ which consists of the eigenvectors of the matrix :

$$\mathbf{E}^{-1}\mathbf{CE} = \begin{bmatrix} c_1^2 & & 0 \\ & c_i^2 & \\ 0 & & c_K^2 \end{bmatrix}, \quad \mathbf{E}^{-1}\mathbf{E} = \mathbf{E}\mathbf{E}^{-1} = \mathbf{I}. \tag{4-39}$$

Here, c_i is the phase speed of the gravity-wave mode of the i-th vertical mode. Then the equation is written as:

$$\mathbf{E}^{-1}[\Delta t^2(k^{*2}+l^{*2})\,\mathbf{C} + \mathbf{I}]\mathbf{E}\mathbf{E}^{-1}\delta_t\underline{D}_{kl} = \mathbf{E}^{-1}(\underline{NL})_{kl}. \tag{4-40}$$

The equation can now be solved as:

$$\delta_t\underline{D}_{kl} = \mathbf{E}\begin{bmatrix} (\omega_1^2+1)^{-1} & & 0 \\ & (\omega_i^2+1)^{-1} & \\ 0 & & (\omega_K^2+1)^{-1} \end{bmatrix}\mathbf{E}^{-1}(\underline{NL})_{kl}. \tag{4-41}$$

where $\omega_i^2 = (c_i\Delta t)^2(k^{*2}+l^{*2})$ is the frequency normalized by Δt of the i-th vertical mode with wave number k and l.

Using $\delta_t\underline{D}_{kl}$, we can obtain $\delta_t\underline{T}_{kl}$ and $\delta_t\pi^*_{kl}$ and then $\delta_t u^*_{kl}$ and $\delta_t v^*_{kl}$. Finally, future values of all the prognostic variables $Q_{t+\Delta t}$ are obtained by:

$$Q_{t+\Delta t} = 2\Delta t\cdot\delta_t Q + Q_{t-\Delta t}. \tag{4-42}$$

5. Recent developments of regional models

<Hybrid sigma-p vertical coordinate>

To reduce the error in the calculation of pressure gradient force in the middle and upper troposphere and the stratosphere, a hybrid sigma-p vertical coordinate has been developed and put into use in replacement of the sigma coordinate system both in global and regional models.

<Regional spectral model using spectral nesting>

Another framework of regional spectral model was developed by Juang and Kanamitsu (1994) and has been used operationally in the U.S. NCEP. The basic idea is that the regional model only predicts perturbations from the large-scale field predicted by a coarse-mesh model.

<Semi-Lagrangian time integration>

To relax the limit of time interval for stable integration, semi-Lagrangian scheme has been developed and used in an increasing number of NWP centers. The time changes of prognostic variables due to the advection process are evaluated by searching upstream points from where air parcels are supposed to be advected to the relevant grid points in a time interval. Although some interpolation enters the calculation, which might become additional source of errors, the stability criterion of the time interval has been relaxed considerably. This has lead to a large reduction of computation time.

One problem might be the lack of conservation properties for physical integrals, although usually it is not a serious problem for short-range NWP. Efforts are now being devoted to developing models with conservation properties.

<Regionally enhanced resolution and adaptive grid>

To focus on a particular meso-scale phenomenon with limited computation resources, there have been developed models with regionally enhanced resolution or those with temporally adaptive grid. The triply-nested model for operational hurricane prediction at NCEP, which was developed in GFDL, is an excellent example.

References

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of motion: Two-dimensional incompressible flow. Part I. *J. Comput. Phys.*, 1, 119-143.

Arakawa, A., 1972: Design of the UCLA General Circulation Model. Technical Report No. 7, Dept. of Met., Univ. of California, Los Angels, 116pp.

Arakawa, A. and C. S. Konor, 1996: Vertical differencing of the primitive equations based on the Charney-Phillips grid in hybrid σ-θ vertical coordinates. Mon. Wea. Rev., 124, 511-528.

Arakawa, A. and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. Methods in Computational Physics, 17, Academic Press, 173 265, 337pp.

Arakawa, A. and S. Moothi, 1988: Broclinic instability in vertically discrete systems. *J. Atmos. Sci.*, 45, 1688-1707.

Juang, H.-M. and M. Kanamitsu, 1994: The NMC nested regional spectral model. *Mon. Wea. Rev.*, 122, 3-26.

Tatsumi, Y., 1986: A spectral limited area model with time dependent lateral boundary conditions and its application to a multi-level primitive equation model. *J. Meteor. Soc. Japan*, 64, 637-664.