CHAPTER 11 Spectral Methods

11.1 Introduction

Assume that $u(x, t)$ is real and integrable. If the domain is periodic, with period L, we can express $u(x, t)$ exactly by a Fourier series expansion:

$$u(x, t) = \sum_{k=-\infty}^{\infty} \hat{u}_k(t) e^{ikx} . \tag{11.1}$$

The complex coefficients $u_k(t)$ can be evaluated using

$$\hat{u}_k(t) = \frac{1}{L} \int_{x-L/2}^{x+L/2} u(x, t) e^{-ikx} \, dx . \tag{11.2}$$

Recall that the proof of (11.1) and (11.2) involves use of the orthogonality condition

$$\frac{1}{L} \int_{x-L/2}^{x+L/2} e^{-ikx} e^{ilx} \, dx = \delta_{kl} , \tag{11.3}$$

where

$$\delta_{kl} \equiv \begin{cases} 1, k = l \\ 0, k \neq l \end{cases} \tag{11.4}$$

is the Kronecker delta.

From (11.1), we see that the x-derivative of u satisfies

$$\frac{\partial u}{\partial x}(x, t) = \sum_{k=-\infty}^{\infty} ik\hat{u}_k(t)e^{ikx} . \tag{11.5}$$

Inspection of (11.5) shows that $\frac{\partial u}{\partial x}$ does not have a contribution from u_0; the reason for this should be clear.

A spectral model uses equations similar to (11.1), (11.2), and (11.5), but with a finite set of wave numbers, and with x defined on a finite mesh:

$$u(x_j, t) \cong \sum_{k=-n}^{n} \hat{u}_k(t)e^{ikx_j} , \tag{11.6}$$

$$\hat{u}_k(t) \cong \frac{1}{M}\sum_{j=1}^{M} u(x_j, t)e^{-ikx_j}, -n \le k \le n , \tag{11.7}$$

$$\frac{\partial u}{\partial x}(x_j, t) \cong \sum_{k=-n}^{n} ik\hat{u}_k(t)e^{ikx_j} . \tag{11.8}$$

Note that we have used "approximately equal signs" in (11.6) - (11.8). In (11.7) we sum over a grid with M points. In the following discussion, we assume that the value of n is chosen by the user. The value of M, corresponding to a given value of n, is discussed below.

Substitution of (11.6) into (11.7) gives

$$\hat{u}_k(t) = \frac{1}{M}\sum_{j=1}^{M}\left\{\left[\sum_{l=-n}^{n} \hat{u}_l(t)e^{ilx_j}\right]e^{-ikx_j}\right\}, -n \le k \le n . \tag{11.9}$$

This is of course a rather circular substitution, but the result serves to clarify some basic ideas. If expanded, each term on the right-hand side of (11.9) involves the product of two wave numbers, l and k, each of which lies in the range $-n$ to n. The range for wave number l is explicitly spelled out in the inner sum on the right-hand side of (11.9); the range for wave number k is understood because, as indicated, we wish to evaluate the left-hand side of (11.9) for k in the range -n to n. Because each term on the right-hand side of (11.9) involves the product of two Fourier modes with wave numbers in the range

$-n$ to n, each term includes wave numbers up to $\pm 2n$. We therefore need $2n+1$ complex coefficients, i.e. $2n+1$ values of the $u_k(t)$.

Because u is real, it must be true that $\hat{u}_{-k} = \hat{u}_k^*$, where the * denotes the conjugate. To see why this is so, consider the $+k$ and $-k$ contributions to the sum in (11.6):

$$\begin{aligned} T_k(x_j) &\equiv \hat{u}_k(t) e^{ikx_j} + \hat{u}_{-k}(t) e^{-ikx_j} \\ &\equiv R_k e^{i\theta} e^{ikx_j} + R_{-k} e^{i\mu} e^{-ikx_j} . \end{aligned} \tag{11.10}$$

where $R_k e^{i\theta} \equiv \hat{u}_k(t)$ and $R_{-k} e^{i\mu} \equiv \hat{u}_{-k}(t)$, and R_k and R_{-k} are real and non-negative. By linear independence, our assumption that $u(x_j, t)$ for all x_j is real implies that the imaginary part of Tkx_j must be zero, for all x_j. It follows that

$$R_k \sin(\theta + kx_j) + R_{-k} \sin(\mu - kx_j) = 0 \quad \text{for all } x_j . \tag{11.11}$$

The only way to satisfy this for all x_j is to set

$$\theta + kx_j = -(\mu - kx_j) = -\mu + kx_j \text{ , which means that } \theta = -\mu , \tag{11.12}$$

and

$$R_k = R_{-k} . \tag{11.13}$$

Eqs. (11.12) and (11.13) imply that

$$\hat{u}_{-k} = \hat{u}_k^* , \tag{11.14}$$

as was to be demonstrated.

Eq. (11.14) implies that u_k and u_{-k} together involve only two distinct real numbers. In addition, it follows from (11.14) that u_0 is real. Therefore, the $2n+1$ complex values of u_k embody the equivalent of only $2n+1$ distinct real numbers. The Fourier representation up to wave number n is thus equivalent to representing the real function $u(x, t)$ on $2n+1$ grid points, in the sense that the information content is the same. We conclude that, in order to use a grid of M points to represent the amplitudes and phases of all waves up to $k = \pm n$, we need $M \geq 2n+1$; we can use more than $2n+1$ points, but not fewer.

As a very simple example, a highly truncated Fourier representation of u including just wave numbers zero and one is equivalent to a grid-point representation of u using 3 grid points. The real values of u assigned at the three grid points suffice to compute the coefficient of wave number zero (the mean value of u) and the phase and amplitude (or "sine and cosine coefficients") of wave number one.

Substituting (11.7) into (11.8) gives

$$\frac{\partial u}{\partial x}(x_l, t) \cong \sum_{k=-n}^{n} \left[\frac{ik}{M} \sum_{j=1}^{M} u(x_j, t) e^{-ikx_j} \right] e^{ikx_l} \ . \tag{11.15}$$

Reversing the order of summation leads to

$$\frac{\partial u}{\partial x}(x_l, t) \cong \sum_{j=1}^{M} \alpha_j^l u(x_j, t) \ , \tag{11.16}$$

where

$$\alpha_j^l \equiv \frac{i}{M} \sum_{k=-n}^{n} k e^{ik(x_l - x_j)} \ . \tag{11.17}$$

The point of this little derivation is that (11.16) can be interpreted as a finite-difference approximation - a special one involving *all* grid points in the domain. From this point of view, spectral models can be regarded as a class of finite-difference models.

Now consider the one-dimensional advection equation with a constant current, c:

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \ . \tag{11.18}$$

Substituting (11.6) and (11.8) into (11.18) gives

$$\sum_{k=-n}^{n} \frac{d\hat{u}_k}{dt} e^{ikx} + c \sum_{k=-n}^{n} ik\hat{u}_k e^{ikx} = 0 \ . \tag{11.19}$$

By linear independence, we obtain

$$\frac{d\hat{u}_k}{dt} + ikc\hat{u}_k = 0 \ \text{ for } -n \le k \le n \ . \tag{11.20}$$

Note that $\frac{d\hat{u}_0}{dt}$ will be equal to zero; the interpretation of this should be clear. We can use (11.20) to predict $u_k(t)$. When we need to know $u(x, t)$, we can get it from (11.6).

Compare (11.20) with

$$\frac{d\hat{u}_k}{dt} + ikc\left(\frac{\sin k\Delta x}{k\Delta x}\right)\hat{u}_k = 0 , \tag{11.21}$$

which, as we have seen, is obtained by using centered second-order space differencing. The spectral method gives the *exact* advection speed (for each Fourier mode), while the finite difference method gives a slower value. Similarly, spectral methods give the *exact* phase speeds for linear waves, while finite difference methods generally underestimate the phase speeds. Keep in mind, however, that the spectral solution is not really exact, because only a finite number of modes are kept.

To evaluate the horizontal pressure gradient force, it is necessary to take horizontal derivatives of the terrain height. Suppose that we have continents and oceans, as schematically shown in Fig. 11.1. In a spectral model the terrain heights will have to be expanded and

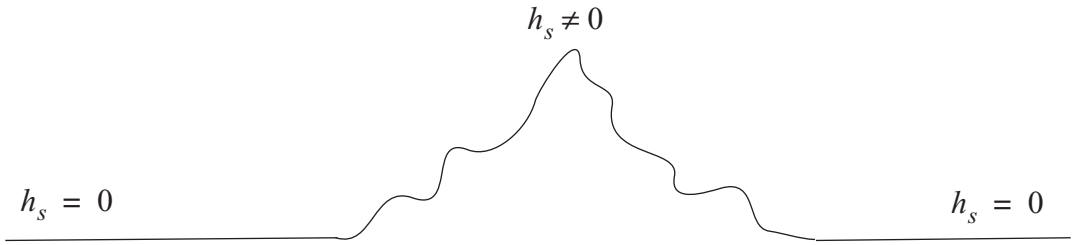

Figure 11.1: The Earth is bumpy.

truncated, like all of the other variables. Truncation leads to "bumpy" oceans. Recently some new approaches have been suggested to alleviate this problem (Boutelou, 1995; Holzer, 1996; Lindberg and Broccoli, 1996).

Another strength of spectral methods is that they make it very easy to solve boundary value problems. As an example, consider

$$\nabla^2 u = f(x, y) , \tag{11.22}$$

as a problem to determine u for given $f(x, y)$. In one dimension, (11.22) becomes

$$\frac{d^2u}{dx^2} = f(x) \, . \tag{11.23}$$

We assume periodic boundary conditions and expand both u and f as Fourier series in x, following (11.1). Then (11.23) becomes

$$\sum_{k=-n}^{n} (-k^2)\hat{u}_k e^{ikx} = \sum_{k=-n}^{n} \hat{f}_k e^{ikx} \, . \tag{11.24}$$

Equating coefficients of e^{ikx} , we find that

$$\hat{u}_k = \frac{-\hat{f}_k}{k^2} \quad \text{for} \quad -n \le k \le n \quad (\text{unless } k = 0) \, . \tag{11.25}$$

Eq. (11.25) can be used to obtain $\hat{u}_k$, for $k = 1, n$. Then $u(x)$ can be constructed using (11.1). This completes the solution of (11.23), apart from the application of an additional boundary condition to determine u_0 . The solution is exact *for the modes that are included*; it is approximate because not all modes are included.

Now consider a nonlinear problem, such as

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} \, , \tag{11.26}$$

again with a periodic domain. Substitution gives

$$\sum_{k=-n}^{n} \frac{d\hat{u}_k}{dt} e^{ilx} = -\left(\sum_{l=-n}^{n} \hat{u}_l e^{ilx} \right) \left(\sum_{m=-n}^{n} im\hat{u}_m e^{imx} \right) . \tag{11.27}$$

Our goal is to predict $u_k(t)$ for k in the range $-n$ to n . The right-hand-side of (11.27) involves products of the form

$$e^{ilx} e^{imx} \, , \tag{11.28}$$

where l and m are in the range $-n$ to n . These products can generate "new" wave numbers, some of which lie outside the range $-n$ to n . Those that lie outside this range

are simply neglected, i.e., they are not included when we evaluate and make use of the left-hand side of (11.27).

For a given Fourier mode, (11.27) implies that

$$\frac{d\hat{u}_k}{dt} = \sum_{l=-\alpha}^{\alpha} \sum_{m=-\alpha}^{\alpha} [\hat{u}_l \hat{u}_m e^{i(l+m)x}] e^{-ikx}, \; -u \le k \le n \tag{11.29}$$

Here we must choose α large enough so that we pick up all possible combinations of l and m that lie in the range $-n$ to n. See Fig. 11.2. The circled X 's in the figure denote

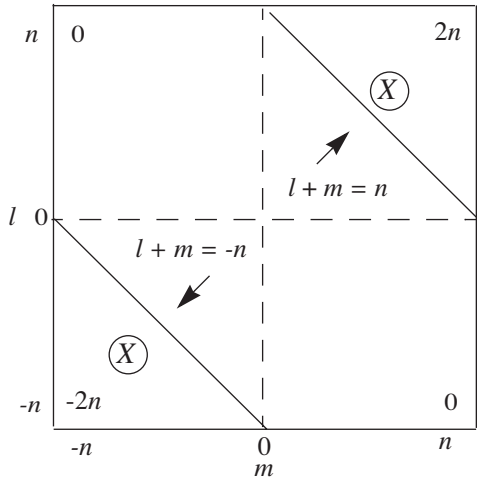

Figure 11.2: Table of $l + m$., showing which (l, m) pairs can contribute to wave numbers in the range $-n$ to n. The pairs in the triangular regions marked by X's do not contribute.

excluded triangular regions. The number of points in each region is

$$1 + 2 + 3 + \ldots + (n-1) = \frac{n(n-1)}{2}. \tag{11.30}$$

The number of points *retained* is

$$\begin{aligned} &(2n+1)^2 - 2\left[\frac{n(n-1)}{2}\right] \\ &= (4n^2 + 4n + 1) - (n^2 - n) \\ &= 3n^2 + 5n + 1 \end{aligned} \tag{11.31}$$

This is the number of terms that must be evaluated in (11.27). The number of terms in the product of sums on the right-hand-side of (11.27) is of order n^2, i.e. it grows very rapidly as n increases. The amount of computation grows rapidly as n increases, and of course the problem is "twice as hard" in two dimensions. At first, this poor scaling with problem size appeared to make spectral methods prohibitively expensive for nonlinear (i.e. realistic) problems.

A way around this practical difficulty was proposed by Orszag, and independently by Eliassen et al., both in 1970. They suggested a "transform method" in which (11.1) and (11.5) are used to evaluate u and $\frac{\partial u}{\partial x}$ on a grid. Sufficient grid points are use to allow the *exact* representation, for wave numbers in the range -*n* to *n*, of *quadratic* nonlinearities like $u\frac{\partial u}{\partial x}$. Of course, here "exact" means "exact up to wave number *n*." Because the solution is exact for wave numbers up to *n*, there is no error for those wave numbers, and in particular, *there is no aliasing error*. Therefore, a model of this type is not subject to aliasing instability arising from quadratic terms like $u\frac{\partial u}{\partial x}$. Aliasing can still arise, however, from "cubic" or higher-order nonlinearities.

To investigate the transform method, we proceed as follows. By analogy with (11.7), we can write

$$\left(u\widehat{\frac{\partial u}{\partial x}}\right)_k = \frac{1}{M}\sum_{j=1}^{M}\left[u(x_j)\frac{\partial u}{\partial x}(x_j)e^{-ikx_j}\right], -n \le k \le n \ . \tag{11.32}$$

Here the hat on $\left(u\widehat{\frac{\partial u}{\partial x}}\right)_k$ indicates that the entire quantity is represented in wave-number space rather than grid space. Now use (11.6) and (11.8) to express $u(x_j)$ and $\frac{\partial u}{\partial x}(x_j)$ in terms of Fourier series:

$$\left(u\widehat{\frac{\partial u}{\partial x}}\right)_k = \frac{1}{M}\sum_{j=1}^{M}\left[\left(\sum_{l=-n}^{n}\hat{u}_l e^{ilx_j}\right)\left(\sum_{m=-n}^{n} im\hat{u}_m e^{imx_j}\right)e^{-ikx_j}\right], -n \le k \le n \ . \tag{11.33}$$

Eq. (11.33) is analogous to (11.9). When expanded, each term on the right-hand side of (11.33) involves the product of three Fourier modes (*k, l* and *m)*, and therefore includes zonal wave numbers up to $\pm 3n$. We need $3n+1$ complex coefficients to encompass

wave numbers up to $\pm 3n$, and because $u\frac{\partial u}{\partial x}$ is real these $3n+1$ complex coefficients actually correspond to $3n+1$ independent real numbers. We therefore need

$$M \geq 3n+1 \tag{11.34}$$

grid points to represent $u\frac{\partial u}{\partial x}$ exactly, up to wave number n.

In summary, the transform method to solve (11.26) works as follows:

1) Initialize the spectral coefficients u_k, for $-n \leq k \leq n$.

2) Evaluate both u and $\frac{\partial u}{\partial x}$ on a grid with M points, where $M \geq 3n+1$. Here $\frac{\partial u}{\partial x}$ is computed using the spectral method, i.e. Eq. (11.8).

3) Form $u\frac{\partial u}{\partial x}$ on the grid, by multiplication.

4) Transform $u\frac{\partial u}{\partial x}$ back into wave- number space, for $-n \leq k \leq n$.

5) Predict new values of the u_k, using

$$\frac{d\hat{u}_k}{dt} = -\left(u\frac{\hat{\partial u}}{\partial x}\right)_k .$$

1) Return to step 2, and repeat this cycle as many times as desired.

Note that the grid-point representation of u contains more information $(3n+1 \text{ real values})$ than the spectral representation $(2n+1 \text{ real values})$. For this one-dimensional problem the ratio is approximately 3/2. The additional information embodied in the grid-point representation is thrown away in step 4 above, when we transform from the grid back into wave-number space. It is not "remembered" from one time step to the next. In effect, we throw away about 1/3 of the information that is represented on the grid. This is the price that we pay to avoid aliasing due to quadratic nonlinearities.

11.2 Spectral methods on the sphere

Spectral methods on the sphere were first advocated by Silberman (1954). A function F that is defined on the sphere can be represented by

$$F(\lambda, \varphi) = \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} F_n^m Y_n^m(\lambda, \varphi) , \tag{11.35}$$

where the

$$Y_n^m(\lambda, \phi) = e^{im\lambda} P_n^m(\sin\varphi) \tag{11.36}$$

are spherical harmonics, and the $P_n^m(\sin\varphi)$ are the associated Legendre functions of the first kind, which happen to be polynomials, satisfying

$$P_n^m(x) = \frac{(2n)!}{2^n n!(n-m)!}(1-x^2)^{\frac{m}{2}} \left[x^{n-m} - \frac{(n-m)(n-m-1)}{2(2n-1)} x^{n-m-2} \right.$$

$$\left. + \frac{(n-m)(n-m-1)(n-m-2)(n-m-3)}{2 \cdot 4(2n-1)(2n-3)} x^{n-m-4} - \ldots \right] , \tag{11.37}$$

Here m is the zonal wave number and n is the "meridional nodal number." As discussed in the Appendix on spherical harmonics, we must require that $n \geq m$. The spherical harmonics Y_n^m are the eigenfunctions of the Laplacian on the sphere:

$$\nabla^2 Y_n^m = \frac{-n(n+1)}{a^2} Y_n^m . \tag{11.38}$$

Here a is the radius of the sphere. See the Appendix for further discussion.

We can approximate F by a truncated sum:

$$\bar{F} = \sum_{m=-M}^{M} \sum_{n=|m|}^{N(m)} F_n^m Y_n^m . \tag{11.39}$$

Here the overbar indicates that F is an approximation to F. In (11.39), the sum over m from $-M$ to M ensures that F is real. The choice of $N(m)$ is discussed below. For smooth F, $\bar{F}$ converges to F very quickly. Only a few terms are needed to obtain a good representation.

Why should we expand our variables in terms of the eigenfunctions of the Laplacian on the sphere? The Fourier representation discussed earlier is also based on the eigenfunctions of the Laplacian, in just one dimension, i.e. sines and cosines. What is so special about the Laplacian operator? There are infinitely many differential operators, so why choose the Laplacian? A justification is that:

- the Laplacian consumes scalars and returns scalars, unlike, for example, the gradient, the curl, or the divergence;
- the Laplacian can be defined without reference to any coordinate system;
- the Laplacian is isotropic, i.e. it does not favor any particular direction on the sphere;
- the Laplacian is simple.

How should we choose $N(m)$? This is the problem of truncation. The two best-known possibilities are *triangular truncation* and *rhomboidal truncation*:

$$\text{Rhomboidal: } N - |m| = M = \text{constant} \tag{11.40}$$

$$\text{Triangular: } N = M = \text{constant}, \text{ or } N - |m| = M - |m| . \tag{11.41}$$

These are illustrated in Fig. 11.3. As shown in Fig. 11.4, triangular truncation represents the *observed* kinetic energy spectrum more accurately, with a small number of terms, than does rhomboidal truncation (Baer, 1972). The thin lines in Fig. 11.4 show the modes kept with triangular truncation. With rhomboidal truncation the thin lines would be horizontal. The thick lines show the observed kinetic energy percentage in each component. For example, we might want to truncate so that we keep all modes with $\geq 0.01\,\%$ of the kinetic energy, and discard all others. Triangular truncation can do that.

In addition, triangular truncation has the beautiful property that it is not tied to a coordinate system. Here is what this means: In order to actually perform a spherical harmonic transform, it is necessary to adopt a spherical coordinate system (λ, φ) . There are of course infinitely many such systems, which differ in the orientations of their poles. There is no reason in principle that the coordinates have to be chosen in the conventional way, so that the poles of the coordinate system coincide with the Earth's poles of rotation. The choice of a particular spherical coordinate system is, therefore, somewhat arbitrary. Suppose that we choose two different spherical coordinate systems (tilted with respect to one another in an arbitrary way), perform a triangularly truncated expansion in both, then plot the results. It can be shown that the two maps will be identical, i.e.

$$F(\lambda_1, \varphi_1) = F(\lambda_2, \varphi_2) , \tag{11.42}$$

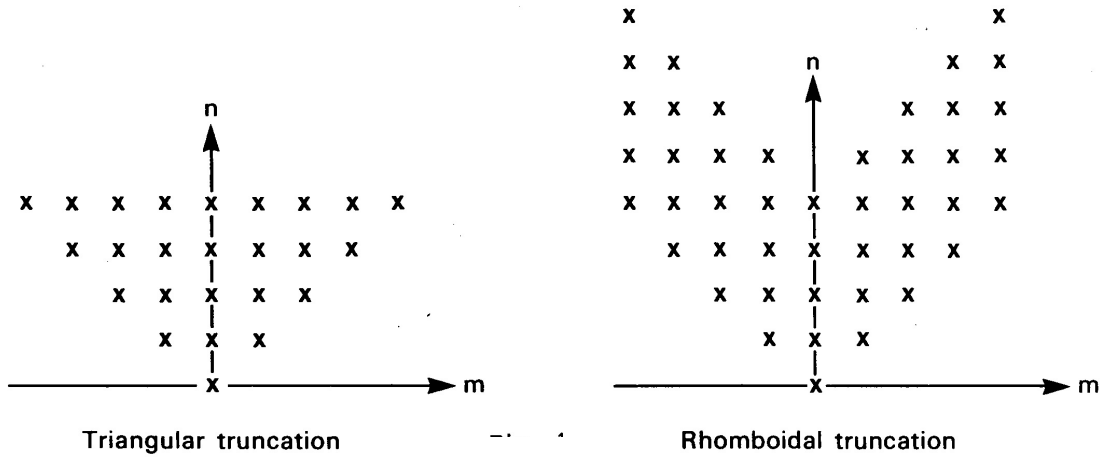

Figure 11.3: Rhomboidal and triangular truncation. From Jarraud and Simmons (1983).

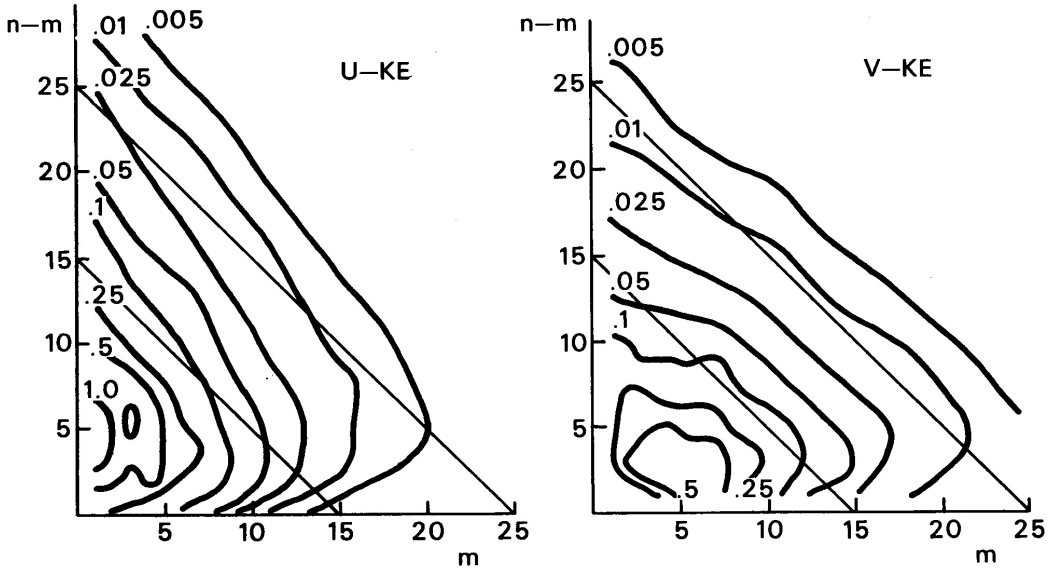

Figure 11.4: Percentage of total kinetic energy in each spectral component. From Jarraud and Simmons (1983) based on Baer (1972).

where the subscripts indicate alternative spherical coordinate systems. This means that the arbitrary orientations of the spherical coordinate systems used have no effect

whatsoever on the results obtained. The coordinate system used "disappears" at the end. Triangular truncation is very widely used today, in part because of this nice property.

In order to use (11.39), we need a "spherical harmonic transform," analogous to a Fourier transform. From (11.36), we see that a spherical harmonic transform is equivalent to the combination of a Fourier transform and a Legendre transform. The Legendre transform is formulated using a method called "Gaussian quadrature." The idea is as follows. Suppose that we are given a function $f(x)$ defined on the interval $-1 \le x \le 1$, and we wish to evaluate

$$I = \int_{-1}^{1} f(x)dx , \tag{11.43}$$

by a numerical method. If f is defined at a finite number of x, denoted by x_i, then

$$I \cong \sum_{i=1}^{N} f(x_i) w_i , \tag{11.44}$$

where the w_i are "weights." Now suppose that $f(x)$ is a weighted sum of Legendre polynomials. Gauss showed that in that case (11.44) gives the *exact* value of I, provided that the x_i are chosen to be the roots of the highest Legendre polynomial used. In other words, we can use (11.44) to evaluate the integral (11.43) exactly, provided that we choose the latitudes so that they are the roots of the highest Legendre polynomial used. These latitudes can be found by a variety of iterative methods. The Gaussian quadrature algorithm is used to perform the Legendre transform.

With *either* triangular *or* rhomboidal truncation, choosing M fixes the expansion; hence the expressions "R15" or "T106." The numeral is the value of M. The numbers of complex coefficients needed are

$$f_R = (M+1)^2 + M^2 + M , \tag{11.45}$$

and

$$f_T = (M+1)^2 , \tag{11.46}$$

respectively.

With the transform method described earlier, the number of grid points needed to avoid aliasing of quadratic nonlinearities exceeds the number of degrees of freedom in the spectral representation. The number of grid points around a latitude circle must be

$\geq 3M+1$. The number of latitude circles must be $\geq \frac{(3M+1)}{2}$ for triangular truncation, and so the total number of grid points needed is $\geq \frac{(3M+1)^2}{2}$. Referring back to (11.46), we see that for large M the grid representation uses about 2.25 times as many equivalent real numbers as the triangularly truncated spectral representation. A similar conclusion holds for rhomboidal truncation. Computing the physics on the fine grid is standard procedure, but wasteful.

In summary, the spectral transform method as it is applied to global models works as follows.

First, we choose a spectral truncation, e.g. T42. Then we identify the number of grid points needed in the longitudinal and latitudinal directions, perhaps with a view to avoiding aliasing due to quadratic nonlinearities. Next, we identify the highest degree Legendre polynomial needed with the chosen spectral truncation, and find the latitudes where the roots of that polynomial occur. These are called the "Gaussian latitudes." At this point, we can set up our "Gaussian grid."

The horizontal derivatives are evaluated in the spectral domain, essentially through "multiplication by wave number." When we transform from the spectral domain to the grid, we combine an inverse fast Fourier transform with an inverse Legendre transform. The nonlinear terms and the model physics are computed on the grid. Then we use the Legendre and Fourier transforms to return to the spectral domain.

The basic logic of this procedure is the same as that described earlier for the simple one-dimensional case.

We have a fast Fourier transform, but no one has yet discovered a "fast Legendre transform," although some recent work points towards one. Lacking a fast Legendre transform, the operation count for a spectral model is of $O(N^3)$, where N is the number of spherical harmonics used. Finite difference methods are, in effect, of $O(N^2)$. This means that spectral models become increasingly expensive, relative to grid-point models, at high resolution.

11.3 The "equivalent grid resolution" of spectral models

Laprise (1992) distinguishes four possible ways to answer the following obvious question: "What is the equivalent grid-spacing of a spectral model?"

1) One might argue that the effective grid spacing of a spectral model is *the average distance between latitudes on the Gaussian grid*. With triangular truncation, this is the same as the spacing between longitudes at the Equator, which is $L_1 = \frac{2\pi a}{3M+1}$. Given the radius of the Earth, and using units of thousands of kilometers, this is equivalent to 13.5/M. For a T31 model (with $M = 31$), we get $L_1 \cong 425\text{km}$. An objection to this measure is that, as

discussed above, much of the information on the Gaussian grid is thrown away when we transform back into spectral space.

2) A second possible measure of resolution is *half the wavelength of the shortest resolved zonal wave at the Equator,* which is $L_2 = \frac{\pi a}{M}$, or about 20/*M* in units of thousands of kilometers. For a T31 model, $L_2 \cong 650\text{km}$.

3) A third method is based on the idea that the spectral coefficients, which are the prognostic variables of the spectral model, can be thought of as a *certain number of real variables per unit area*, distributed over the Earth. A triangularly truncated model has the equivalent of $(M+1)^2$ real coefficients. The corresponding resolution is then $L_3 = \sqrt{\frac{4\pi a^2}{(M+1)^2}} = \frac{2\sqrt{\pi}a}{M+1}$, which works out to about 725 km for a T31 model.

4) A fourth measure of resolution is based on the *equivalent total wave number associated with the Laplacian operator*, for the highest mode. The square of this total wave number is $K^2 = \frac{M(M+1)}{2a^2}$. Suppose that we equate this to the square of the equivalent total wave number on a square grid, i.e. $K^2 = k_x^2 + k_y^2$, and let $k_x = k_y = k$ for simplicity. One half of the corresponding wavelength is $L_4 = \frac{\pi}{k} = \frac{\sqrt{2}\pi a}{M}$, which is equivalent to 28.3/M in units of thousands of kilometers. For a T31 model this gives about 900 km.

These four measures of spectral resolution range over more than a factor of two. The measure that makes a spectral model "look good" is L_1, and so it is not surprising that this is the measure that spectral modelers almost always use when specifying the equivalent grid spacing of their models.

11.4 Semi-implicit time differencing

As we have already discussed in Chapters 5 and 8, gravity waves limit the time step that can be used in a primitive-equation (or shallow water) model. A way to get around this is to use semi-implicit time differencing, in which the "gravity wave terms" of the equations are treated implicitly, while the other terms are treated explicitly. This can be accomplished much more easily in a spectral model than in a finite-difference model.

A detailed discussion of this approach will not be given here, but the basic ideas are as follows. The relevant terms are the pressure-gradient terms of the horizontal equations of motion, and the mass convergence term of the continuity equation. These are the same terms that we focused on in the discussion of the pole problem, in Chapter 8. The terms involve horizontal derivatives of the "height field" and the winds, respectively. Typically the Coriolis terms are also included, so that the waves in question are inertia-gravity waves.

Consider a finite-difference model. If we implicitly difference the gravity-wave terms, the resulting equations will involve the "n+1" time-level values of the heights and the winds at multiple grid points in the horizontal. This means that we must solve simultaneously for the "new" values of the heights and winds. Such problems can be solved, of course, but they can be computationally expensive. For this reason, most finite-difference models do not use semi-implicit time differencing.

In spectral models, on the other hand, we prognose the spectral coefficients of the heights and winds, and so we can apply the gradient and divergence operators simply by multiplying by wave number (roughly speaking). This is a "local" operation in wave-number space, so it is not necessary to solve a system of simultaneous equations.

The use of semi-implicit time differencing allows spectral models to take time steps several times longer than those of (explicit) grid-point models. This is a major advantage in terms of computational speed, which compensates, to some extent, for the expense of the spectral transform.

11.5 Conservation properties and computational stability

Because the spectral transform method prevents aliasing for quadratic nonlinearities, but not cubic nonlinearities, spectral models are formulated so that the highest nonlinearities that appear in the equations (other than in the physical parameterizations) are quadratic. This means that the equations must be written in advective form, rather than flux form. As a result, the models do not exactly conserve anything -- even mass -- for a general, divergent flow.

It can be shown, however, that in the limit of two-dimensional non-divergent flow, spectral models do conserve kinetic energy and enstrophy. Because of this property, they are well behaved computationally. Nevertheless, all spectral models need some artificial diffusive damping to avoid computational instability. In contrast, it is possible to formulate finite-difference models that are very highly conservative and can run indefinitely with no artificial damping at all.

11.6 Moisture advection

The mixing ratio of water vapor is non-negative. We have already discussed the possibility of spurious negative mixing ratios caused by dispersion errors in finite-difference schemes, and we have also discussed the families of finite-difference advection schemes that are "sign-preserving" and do not suffer from this problem.

Spectral models have a very strong tendency to produce negative water vapor mixing ratios (e.g. Williamson and Rasch, 1994). In the global mean, the rate at which "negative water" is produced can be a significant fraction of the globally averaged precipitation rate. Negative water vapor mixing ratios occur not only locally on individual time steps, but even in zonal averages that have been time-averaged over a month.

Because of this disastrous situation, many spectral models are now using non-spectral

methods for advection (e.g. Williamson and Olson, 1994). This means that they are only "partly spectral." When non-spectral methods are used to evaluate the nonlinear advection terms, the motivation for using the high-resolution, non-aliasing grid disappears. Such models can then use a coarser "*linear grid*," with the same number of grid points as the number of independent real coefficients in the spectral representation. This leads to a major savings in computational cost.

11.7 Physical parameterizations

Because most physical parameterizations are highly nonlinear, spectral models evaluate such things as convective heating rates, turbulent exchanges with the Earth's surface, and radiative transfer on their Gaussian grids. The tendencies due to these parameterizations are then applied to the prognostic variables, which are promptly transformed into wave-number space.

Recall that when this transform is done, the spectral representation contains less information than is present on the grid, due to the spectral truncation to avoid aliasing due to quadratic nonlinearities. This means that if the fields were immediately transformed back onto the grid (without any changes due, e.g., to advection), the physics would not "see" the fields that it had just finished with. Instead, it would see spectrally truncated versions of these fields.

For example, suppose that the physics package includes a convective adjustment that is supposed to adjust convectively unstable columns so as to remove the instability. Suppose that on a certain time step this parameterization has done its work, removing all instability as seen on the Gaussian grid. After spectral truncation, some convective instability may re-appear, even though "physically" nothing has happened!

In effect, the spectral truncation that is inserted between the grid domain and the spectral domain prevents the physical parameterizations from doing their work properly. This is a problem for all spectral models. It can be solved by doing the physics on a "linear" grid that has the same number of degrees of freedom as the spectral representation.

11.8 Summary

In summary, the spectral method has both strengths and weaknesses:

Strengths:

- Especially with triangular truncation, it eliminates the "pole problem."
- It gives the exact phase speeds for linear waves and advection by a constant current such as solid-body rotation.
- It converges very rapidly, and gives good results with just a few modes.
- Semi-implicit time-differencing schemes are easily implemented in spectral models.

Weaknesses:

- Spectral models do not exactly conserve anything -- not even mass.
- Partly because of failure to conserve the mass-weighted total energy, artificial damping is needed to maintain computational stability.
- Spectral models have bumpy oceans.
- Because of truncation in the transform method, physical parameterizations do not always have the intended effect.
- Moisture advection does not work well in the spectral domain.
- At high resolution, spectral methods are computationally expensive compared to grid point models.

Problems

1. Write subroutines to compute Fourier transforms and inverse transforms, for arbitrary complex $u(x_j)$. The number of waves to be included in the transform and the number of grid points to be used in the inverse transform should be set through the argument lists of subroutines.

a) Let

$$u(x_j) = 14\cos(k_0 x_j) + 6i\ \cos(k_1 x_j) + 5 \tag{11.47}$$

where

$$\begin{aligned} k_0 &= \frac{2\pi}{L_o}, \quad L_o = \frac{X}{4}, \\ k_1 &= \frac{2\pi}{L_1}, \quad L_1 = \frac{X}{8}. \end{aligned} \tag{11.48}$$

Compute the Fourier coefficients starting from values of x_j on a grid of M points, for each value of M in the angle.

$$2 \le M \le 20 \tag{11.49}$$

Tabulate u_k for $0 \le M \le 8$, and $2 \le M \le 20$. Discuss your results.

b) Repeat for

$$u(x_j) = 5\cos k_0 x_j + 7\sin k_1 x_j + 2\ . \tag{11.50}$$

Use the same values of k_0 and k_1 as given above.