Tellus (1996), 48A, 41–56
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

Orthogonal expansion functions for meridional profiles

By INGEMAR HOLMSTRÖM, *St. Papoul, 31450 Odars, France*

(Manuscript received 13 August 1993; in final form 5 December 1994)

ABSTRACT

In order to find an efficient spectral representation for the variables in a non-linear partial differential equation, one may require that the expanding functions should satisfy the equation, not exactly as in the linear case, but instead as well as possible in the least square sense. This variational approach is here applied to the non-linear potential vorticity equation for a zonal flow and finite amplitude perturbations in a beta-plane channel. The solutions are required to be neutral. For zonal wave numbers 2–10 primary modes are found for which the zonal wind has a typical jet-like profile. Its form varies very little with the zonal scale and it is shown that a short wave will have the same meridional profile as a much longer wave. The longest waves are quasi-stationary as is usually observed in the atmosphere. Shorter waves move east with phase velocities which are higher than those obtained with the Rossby formula. They agree, however, well with observations.

For higher orthogonal modes a mathematical modification is made in the equation for the primary modes so that a complete Sturm-Liouville orthogonal system is obtained for all modes. These higher modes also show little dependence on zonal wave number and one may therefore question if it is necessary to have separate expansion functions in the meridional direction for each zonal wave number. For the beta-plane model this necessity was not confirmed in an expansion of the meridional profile of the perturbation Jacobian, where a comparison also was made with Fourier expansions. In this case the functions determined here are found to give a much faster convergence.

1. Introduction

The use of spectral models in numerical weather prediction naturally poses the problem of the best choice of the expanding functions. Spherical harmonics were directly indicated by the geometry of the earth but were not related to the model equations. A more compatible representation is obtained with the normal modes, Kasahara (1976) and references, which during the past two decades have replaced the spherical harmonics in spectral models, where they also are used for data assimilation. The normal modes are obtained as eigensolutions to the shallow water equation, perturbed about a state of rest, so that all non-linear terms are excluded. The atmosphere, however, with differential heating, jet stream and developing instabilities is very far from a state of rest and one may therefore question the efficiency of the normal modes for expansion of atmospheric or computed data.

In Tanaka and Kung (1989) it is estimated that expansions with 7 vertical normal modes and 19 meridional normal modes represent about 90 per cent of the total energy of the atmosphere. This may be compared with expansions with the aid of empirical orthogonal functions (EOF's), where it is known that 4 to 5 functions are sufficient for representation of more than 95 per cent of the variance in the vertical or meridional directions, Holmström (1963), Bradley and Wiin-Nielsen (1968). EOF's are not model related but a perfect or nearly perfect model should evidently produce data that give the same EOF's and reductions of variances as those obtained from observed data. Therefore one may take the EOF expansion as an indication of what it should be possible to obtain with "normal modes" determined in a different way.

In Daley (1978) it was stated that "ideally, one would like to obtain the normal modes from the fully non-linear equations" but it was added that

this was, of course, not possible. In a linear partial differential equation, where the variables can be separated, the eigenfunctions satisfy the equations exactly. In nonlinear equations, where the variables cannot in the same way be separated, one may still introduce separate expansions in the three coordinate directions and require that these functions should satisfy the equation, not exactly, but as well as possible. This leads to a variational problem as was shown in Holmström (1977), where an attempt was made to determine vertical structure functions for the atmosphere. There, the results were acceptable in the troposphere only after relaxing a, probably erroneous, lower boundary condition. In the stratosphere the computations gave very large and unrealistic vertical shears with corresponding large north-south temperature gradients. The reason for this failure will be discussed in a later paper where the same problem will be taken up again.

In the present paper, we shall use a similar approach to the simpler problem of finding orthogonal functions for series expansion of meridional profiles. For simplicity, we shall here consider a quasi-geostrophic beta-plane model with walls at 30°N and 70°N and a Cartesian coordinates system. The results will therefore be of limited interest for direct applications, but they will be used later when the problem of vertical structure functions is considered. However, even with this restriction it is found that this simple model in many respects has the same or similar characteristics as the atmosphere. Standard notations will be used. The term wave number is ambiguous and we shall use k in the trigonometric expressions for the waves in the zonal direction and N for the number of waves around the hemisphere.

2. A first separation

The quasi-geostrophic potential vorticity equation is (Holton, 1992),

$$M(\psi)=\frac{\partial D^2\psi}{\partial t}+J(\psi, D^2\psi)+\beta\frac{\partial\psi}{\partial x}=0 \qquad (2.1)$$

where M is the nonlinear operator, J the Jacobian and D^2 defined by

$$D^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+f_0^2\frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial}{\partial p}\right)$$

Here f_0 is the constant coriolis parameter and σ the static stability, taken to be a function of pressure only. We are in this paper primarily interested in expansion functions for meridional profiles, but since meridional and vertical scales are interrelated, it is necessary to include the vertical structures. We now separate the independent variables x and t from the variables y and p by taking

$$\psi(x, y, p, t)=A_1(y, p)+\alpha(x, t)\,A_2(y, p)+\psi_3(x, y, p, t) \qquad (2.2)$$

where $A_1(y, p)$ provides for a zonal flow. In the second term we take the finite amplitude, $\alpha(x, t)$, to be of the form $\alpha(x-ct)$ in order to have a traveling perturbation with phase velocity c. Finally $\psi_3(x, y, p, t)$ is a residual. Then, if at a certain time, $t=0$, the field is entirely determined by $A_1+\alpha A_2$, i.e. if $\psi_3=0$ everywhere initially, the equation (2.1) at $t=0$ reduces to

$$M(A_1+\alpha A_2)=-\frac{\partial}{\partial t}(D^2\psi_3) \qquad (2.3)$$

Here, the right-hand side is a measure of the rate of loss of potential vorticity that the field $A_1+\alpha A_2$ suffers at initial time. We now wish to determine A_1, α and A_2 so that this rate of loss is minimized in the least square sense. If L is the length of the channel, W its width and P the surface pressure, taken to be 1000 hPa, we wish to have

$$I=\frac{1}{LWP}\int_0^P\int_0^W\int_0^L[M(A_1+\alpha A_2)]^2\times dx\,dy\,dp=\min$$

Inserting $\psi=A_1+\alpha A_2$ in (2.1), the integral I may be written

$$I=\frac{1}{LWP}\int_0^L\int_0^W\int_0^P\left\{\frac{\partial\alpha}{\partial x}\left[(U_1-c)D_1^2A_2-A_2(D_1^2U_1-\beta)\right]+\frac{\partial^3\alpha}{\partial x^3}A_2(D_1^2U_1-\beta)+\alpha\frac{\partial\alpha}{\partial x}\left(A_2D_1^2\frac{\partial A_2}{\partial y}-\frac{\partial A_2}{\partial y}D_1^2A_2\right)+A_2\frac{\partial A_2}{\partial y}\left(\frac{\partial\alpha}{\partial x}\frac{\partial^2\alpha}{\partial x^2}-\alpha\frac{\partial^3\alpha}{\partial x^3}\right)\right\}^2 dx\,dy\,dp$$

where now

$$D_1^2 = \frac{\partial^2}{\partial y^2} + f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial}{\partial p}\right)$$

and where we have also introduced $U_1(y, p) = -\partial A_1/\partial y$. The variation of this integral with respect to A_1, α and A_2 is of course possible even if it would require considerable effort. But fortunately the problem can be considerably simplified. The last term in the integrand will vanish exactly if we take $\partial^2\alpha/\partial x^2 = -k^2\alpha$, where k is the wave number. This reduction of the value of the integral confirms that trigonometric functions are efficient for expansion in the zonal direction. The variational problem is in this case reduced to finding the functions U_1 and A_2 that minimize the integral. It is convenient here to take

$$\alpha(x - ct) = -\frac{1}{k}\cos k(x - ct), \tag{2.4}$$

whereby the integral reduces to

$$I = \frac{1}{LWP}\int_0^L \int_0^W \int_0^P \Big\{ [(U_1 - c) D_1^2 A_2 - A_2(D_1^2 U_1 - \beta) - k^2 A_2 (U_1 - c)] \sin k(x - ct) + \frac{1}{2k}\left(A_2 D_1^2 \frac{\partial A_2}{\partial y} - \frac{\partial A_2}{\partial y} D_1^2 A_2\right) \times \sin 2k(x - ct) \Big\}^2 \, dx \, dy \, dp.$$

We take here $k = 2\pi/L$, square the expression in I and carry out the integration in the x-direction. Since $\sin k(x - ct)$ and $\sin 2k(x - ct)$ are orthogonal over L, the integral splits into two separate parts, I_1 and I_2, where

$$I_1 = \frac{1}{2WP}\int_0^W \int_0^P [(U_1 - c) D_1^2 A_2 - A_2(D_1^2 U_1 - \beta) - k^2 A_2 (U_1 - c)]^2 \, dy \, dp, \tag{2.5}$$

$$I_2 = \frac{1}{8k^2 WP}\int_0^W \int_0^P \left[A_2 D_1^2 \frac{\partial A_2}{\partial y} - \frac{\partial A_2}{\partial y} D_1^2 A_2 \right]^2 dy \, dp. \tag{2.6}$$

In I_1 we find inside the square brackets the same terms as in the linearized perturbation vorticity equation so that the integral will vanish exactly if U_1 and A_2 satisfy this equation. The variational problem is then reduced to minimizing the second integral, I_2. We could here choose $D_1^2 A_2 = -m^2 A_2$ which would make $I_2 = 0$, but this is the Craig (1945) or Neamtan (1946) solution and not interesting. Instead we shall here accept that the integral I_2 does not vanish and later show that it may be minimized by a suitable choice of the zonal wave number. Since we have derived the perturbation equation *without the assumption of small perturbations*, we shall take the value of I_2 as a measure, indicating the extent to which the linearized equation may replace the nonlinear potential vorticity equation in the case of finite amplitudes.

3. The neutral solution

Equating to zero the expression inside square brackets in I_1 in (2.4), we obtain the perturbation equation for U_1 and A_2

$$(U_1 - c) D_1^2 A_2 = A_2[k^2(U_1 - c) + D_1^2 U_1 - \beta]. \tag{3.1}$$

For a given positive U_1 and for eastward traveling waves, this equation has a singularity at points where $U_1 = c$ and would therefore require complex A_2 and c for a solution. It is, however, not very interesting to have a series expansion in terms of functions that already in the linearized case start to grow or fade away, and we are therefore in the first instance only interested in neutral solutions to (3.1). That the atmosphere should be close to a neutral state was suggested by Lorenz (1967). A further and decisive reason for this choice is that the perturbation vorticity equation (3.1) only gives one relation between the two unknown, dependent variables, U_1 and A_2. A second relation or condition is therefore required in order to make the system determinate. This desired relation is provided by the condition that the solution to (3.1) should be neutral. This is obtained and the singularity removed if we simply take

$$A_2 = \alpha_0 U_1, \tag{3.2}$$

where α_0 is a constant, which determines the amplitude of the perturbation. It will be discussed later in Subsection 4.3. With (3.2), we obtain from (3.1) the following equation for the zonal flow:

$$\frac{\partial^2 U_1}{\partial y^2} + f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial U_1}{\partial p}\right) = \left(k^2 + \frac{\beta}{c}\right) U_1 - \frac{k^2}{c} U_1^2, \tag{3.3}$$

where U_1 has the same lateral boundary condition as A_2 or $U_1 = 0$ for $y = 0, W$. The vertical boundary conditions will be discussed later. From (2.2), (2.5) and (3.2), it follows that the perturbation streamfunction is

$$\psi' = -\frac{\alpha_0}{k} U_1 \cos k(x - ct),$$

with the rotational wind components

$$u' = \frac{\alpha_0}{k}\frac{dU_1}{dy} \cos k(x - ct),$$
$$v' = \alpha_0 U_1 \sin k(x - ct). \tag{3.4}$$

One may note that in the derivation of normal modes from the shallow water equation about a state of rest, meridional profiles for zonal wave number zero cannot directly be obtained Kasahara (1980), Tanaka and Kasahara, (1992). In contrast, including non-linearity, the solution to eq. (3.3) not only gives the profile of the zonal flow, but also defines what we shall call a *primary finite amplitude perturbation* without which the zonal flow would be undetermined. Together, they give a meandering basic flow. With (3.2) and (3.3), the integral I_2 in (2.6) may be considerably simplified. It is easily shown that the integral reduces to

$$I_2 = \frac{\alpha_0^4 k^2}{8c^2 WP} \int_0^W \int_0^P \left(U_1^2 \frac{\partial U_1}{\partial y}\right)^2 dp\, dy. \tag{3.5}$$

One interesting property of eq. (3.3) is that it gives identical solutions for the profile $U_1(y, p)$ for two different wave numbers and two corresponding phase velocities, provided boundary conditions are similar. If the wave numbers are k_1 and k_2 and the phase velocities c_1 and c_2, the condition is:

$$k_1^2 + \frac{\beta}{c_1} = k_2^2 + \frac{\beta}{c_2}, \qquad \frac{k_1^2}{c_1} = \frac{k_2^2}{c_2}, \tag{3.6}$$

which gives the following relations for determining k_2 and c_2 when k_1 and c_1 are prescribed or known.

$$k_2^2 = \frac{\beta}{c_1}, \qquad c_2 = \frac{\beta}{k_1^2}. \tag{3.7}$$

We note here, that this duality does not exist on the f-plane and that the first relation does not permit a negative value for c_1. From (3.7), it also follows that we may have long neutral slow waves with the same meridional and vertical structure as short waves with a much larger phase velocity. This lack of isotropy is thought to be the reason for the difficulty found in Bradley and Wiin-Nielsen (1968), Baer (1974) to establish a clear interdependence between zonal scale and vertical or meridional scales from observed data.

We may also note that, if $U(y, p)$ is a solution to (3.3), then $\mu U(y, p)$, where μ is a positive constant, is also a solution but with a different wave number k_3 and phase velocity, c_3. An expression for c_3 in terms of c_1 and c_2 is obtained if, in the second eq. (3.6), we multiply the right-hand side by μ. This gives

$$c_3 = \frac{\mu(c_1 + c_2)}{2} \pm \frac{1}{2}\sqrt{\mu^2(c_1 + c_2)^2 - 4\mu c_1 c_2}, \tag{3.8}$$

where we must have

$$\mu \geqslant \frac{4c_1 c_2}{(c_1 + c_2)^2}$$

in order for c_3 to be positive and the corresponding k_3 to be real. Even with this restriction, it is evident that similar profiles will exist in this beta-plane model which all can be described by one and the same normalized function. To what extent this holds for the atmosphere is not exactly known but an indication is certainly given in Bradley and Wiin-Nielsen (1968), where, for each zonal wave number separate vertical and meridional EOF expansions were made and where the functions were found to be almost independent of zonal wave number. Their conclusion was that the verti-

cal structure for geopotential could be represented by three universal characteristic patterns and the latitudinal structure for all zonal wave numbers by four characteristic patterns. We shall return to this question and to eq. (3.8) when values for the phase velocities c_1 and c_2 have been determined.

4. The second separation

4.1. The equation

In (3.3), the variables y and p cannot be separated so, if we introduce $U_1(y, p) = U(y)\,F(p)$, the equation will not be satisfied. Here $F(p)$ is a normalized function with

$$\frac{1}{P}\int_0^P F^2(p)\,\mathrm{d}p = 1,$$

where P is taken to be 995 hPa and where, in applications, we start the integration from 5 hPa in order to exclude the upper stratosphere. The same type of normalization will be used throughout this paper. For shortness, we now put

$$a = k^2 + \frac{\beta}{c}, \qquad b = \frac{k^2}{c},$$
$$F_\mathrm{d} = f_0^2 \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1}{\sigma}\frac{\mathrm{d}F}{\mathrm{d}p}\right). \tag{4.1}$$

With the separation of the variables y and p in (3.3) the residue is

$$R(y, p) = F\frac{\mathrm{d}^2U}{\mathrm{d}y^2} + UF_\mathrm{d} - aUF + bU^2F^2,$$

and we now wish to determine $U(y)$ and $F(p)$ so that the integral

$$I = \frac{1}{WP}\int_0^W \int_0^P R^2\,\mathrm{d}p\,\mathrm{d}y$$

is minimized. The variation with respect to U is straightforward and details are given in the Appendix. It gives the fourth order differential equation

$$\frac{\mathrm{d}^4U}{\mathrm{d}y^4} + p_1\frac{\mathrm{d}^2U}{\mathrm{d}y^2} + p_2\left(\frac{\mathrm{d}^2U^2}{\mathrm{d}y^2} + 2U\frac{\mathrm{d}^2U}{\mathrm{d}y^2}\right) + p_3U + p_4U^2 + p_5U^3 = 0, \tag{4.2}$$

where

$$p_1 = 2(\overline{FF_\mathrm{d}} - a); \qquad p_2 = b\overline{F^3};$$
$$p_3 = \overline{F_\mathrm{d}F_\mathrm{d}} - 2a\overline{FF_\mathrm{d}} + a^2; \tag{4.3}$$
$$p_4 = 3(b\overline{F^2F_\mathrm{d}} - ab\overline{F^3}); \qquad p_5 = 2b^2\overline{F^4},$$

and where the overbar indicates a vertical average. The variation also gave a so-called natural boundary condition with $\mathrm{d}U/\mathrm{d}y = 0$ for $y = 0$, W.

Ideally one would like to solve eq. (3.3) as it stands in order to obtain a two-dimensional function but this turned out to be too difficult due to the presence of the static stability and its rapid variation in the stratosphere, see Holmström (1977). For the same reason, the variation with respect to $F(p)$ and its computation is complicated and lengthy. It will therefore be left to a later paper. In order to obtain solutions to (4.2), we shall here instead determine approximate values for the vertically averaged terms of the coefficients $p_1 - p_5$ in (4.3).

4.2. The temperature field

For the function $F(p)$ in (4.3) it would be natural to take the lowest vertical EOFs since these functions give the most rapid convergence. They are, however, not easy to use since they are known only at a limited number of levels so that vertical derivatives are difficult to evaluate. We shall therefore determine these functions in a different way, using a simplified vertical temperature distribution. If T_0 is the surface temperature, g_0 the tropospheric vertical temperature gradient, T_1 an approximate tropopause temperature and g_1 the vertical temperature gradient in the stratosphere, then the hyperbola

$$[T(z) - (T_0 - g_0z)] \times [T(z) - (T_1 - g_1z)] = M \tag{4.4}$$

gives a realistic vertical temperature distribution. The value of the constant M determines how close $T(z)$ will approach the two asymptotes. Together with the hydrostatic approximation, the equation will give the temperature or height as a function of pressure. As a test we show with open circles in Fig. 1 the temperatures for the basic state used by Tanaka et al. (1986) together with the curve marked T, computed from (4.4) with $M = 5$. For clarity, the circles and the curve have been shifted

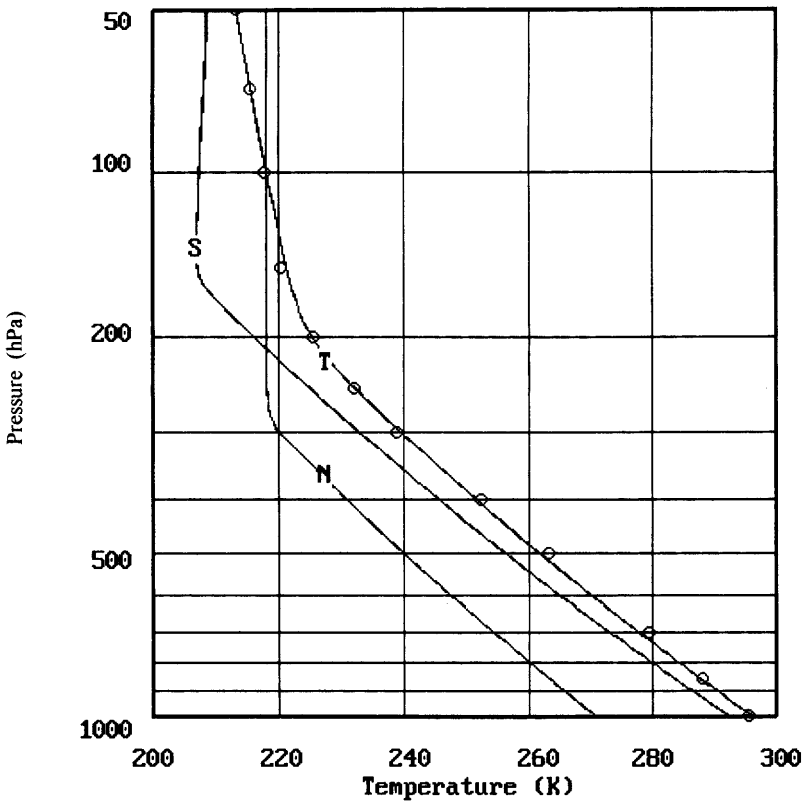

Fig. 1. Basic state temperature (circles) from Tanaka et al. (1986) and computed with (4.4), curve T. Temperatures at southern wall, curve S and at northern wall, curve N, computed with (4.4).

everywhere 5° to the right. It is seen that a close fit can be obtained. The two other curves show the temperatures at the southern (S) and the northern (N) walls to be used in the present paper. Data for eq. (4.4) are given in Table 1 together with h_0, the height of the 1000 hPa surface, which has to be given in order to specify a mean surface zonal wind.

These data have been determined to correspond to average winter conditions. With the height computed as a function of pressure, the geostrophic approximation gives the mean zonal wind. Its normalized profile, $F_1(p)$, is shown as the curve marked F1 in Fig. 2. The small circles show for comparison the first vertical EOF mode for wind (u and v combined), computed from Holmström (1963). Other EOF expansions of the wind vector up to 10 or 15 hPa (unpublished) confirm the tendency for the first mode wind to vanish at the upper level. A small stratospheric temperature gradient is therefore included at the southern wall in order to make $F_1(p)$ vanish at 5 hPa. It is seen that the computed profile very nearly approximates the first EOF mode. The differences lie well within the variations of the EOFs when computed from different data samples of wind observations, and we may therefore use the computed function $F_1(p)$ for calculating the required vertical averages in (4.3). The result is given in Table 2.

Table 1. *Data for temperature calculation with eq. (4.4)*

	T_0 (°C)	h_0 (m)	T_1 (°C)	g_0 (°C/km)	g_1 (°C/km)
30° N	20	160	−71	6.5	−0.33
70° N	−2	60	−55	6.0	0.0

We shall later need also a second vertical mode and for this purpose, we take advantage of the

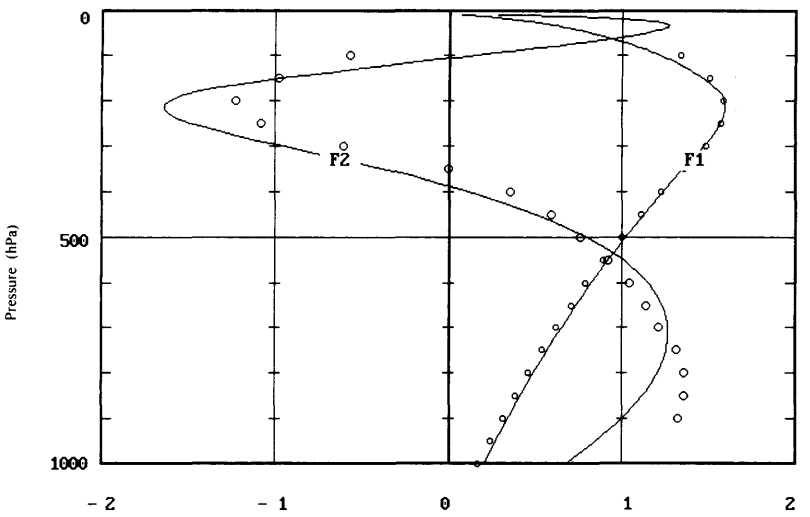

Fig. 2. Normalized vertical profiles for the first two computed EOFs, compared with EOFs given in Holmström (1963), circles.

known dominance of the first mode. Since it represents 85 to 90% of the total atmospheric variance, the second mode should primarily be generated by the non-linear interaction of the first mode with itself. If the flow has the dominant vertical profile $F_1(p)$ then the Jacobian in the vorticity equation will have the vertical profile $F_1^2(p)$ and we should therefore approximately have

$$F_1^2(p) = \mu_1 F_1(p) + \mu_2 F_2(p)$$

where we require that the normalized second mode $F_2(p)$ should be orthogonal to $F_1(p)$. Using this orthogonality and the normalization we obtain with the values in Table 2

$$F_1^2(p) = 1.244 F_1(p) - 0.323 F_2(p). \qquad (4.5)$$

This equation shows that the first mode by nonlinear interaction reproduces to a large extent its own vertical structure but now for a different zonal wavelength. With continued nonlinear interactions this same vertical profile will therefore be found over the whole zonal wave spectrum. This is clearly evident in the map for the first mode height coefficients shown in Holmström (1963), where short as well as very long waves are present.

The profile of the second mode, $F_2(p)$, is shown in Fig. 2 as the curve marked F2. The circles represent the second EOF mode for wind, again taken from Holmström (1963). It is seen that there is a great similarity between the two curves but differences exist. If they are due to the influence of neglected nonlinear terms in the vorticity equation or to other effects is not known. However, since the variability of the second EOF mode depending on sample is larger than in the first, we shall later take the curve F2 as an acceptable approximation for the determination of vertical averages.

4.3. Results

Eq. (4.2) has been solved numerically, mostly on an 80-point grid with the fourth and second derivatives in their standard, finite difference analogue. In the following discussion, we shall use

Table 2. *Values of vertical averages for the first vertical mode*

$\overline{F^3}$	$\overline{F^4}$	$\overline{FF_d}$	$\overline{F^2F_d}$	$\overline{F_d^2}$
1.244	1.653	-1.780×10^{-12}	-2.049×10^{-12}	5.559×10^{-24}

the number N of waves around the hemisphere, since the wave number k has no immediate interpretation. N is prescribed and taken to be $N=2, 3, ..., 10$. Also prescribed is an over the channel averaged zonal velocity U_m, defined by

$$U_m = \frac{1}{W}\int_0^W U(y)\,dy.$$

It represents the meridional forcing and is here chosen to be either 10 or 15 ms^{-1}. Since $F_1(p)$ is close to 1 at 500 hPa, this corresponds geostrophically to a height difference at 500 hPa of approximately 500 or 750 m between the latitudes 30° and 70°, representing winter conditions in the Northern and Southern Hemispheres, respectively, see Figs. 3.1 and 3.4 in Palmén and Newton (1969).

In order to satisfy the boundary conditions, we note that for a small U, the linearized part of (4.2) has the solution

$$U = p_6 \sin(q_1 y)\sinh(q_2 y),$$

where p_6 is an undetermined constant of integration and the values of q_1 and q_2 depend only on the unknown phase velocity c once the zonal wave number and the vertical structure are prescribed. In order to solve (4.2) a first guess is made on the amplitude p_6 and the phase velocity c and the program then modifies these values until the boundary condition at $y = W$ is satisfied and the mean value of $U(y)$ is equal to U_m, both to an accuracy of 10^{-5} ms^{-1}.

Meridional profiles of $U(y)$ computed from (4.2) are shown in Fig. 3 for $N=2$, 3 and 4 and $U_m = 15$ ms^{-1} and in Fig. 4 for $N=5$ and $U_m = 10$ and 15 ms^{-1}. All solutions have a more or less pronounced jet-like profile, which would have to be established by a preceding convergence of momentum transport. Once the profile has been established, the momentum transport vanishes, as is seen from (3.5), which gives the rotational v-component as a sine function and the u-component as a cosine function. From the thermodynamic equation, using the geostrophic assumption, it is easily shown that the components of the divergent part of the wind do not contribute either to a momentum transport.

For the values of U_m used here, hemispheric wave number $N=1$ (not shown) gives a very pronounced jet but an unrealistic, very high maximum velocity and an unrealistic meridional shear well exceeding the coriolis parameter. This is also the case for $N=2$ (Fig. 3) for part of the profile. In order to see if this is due to the neglect of

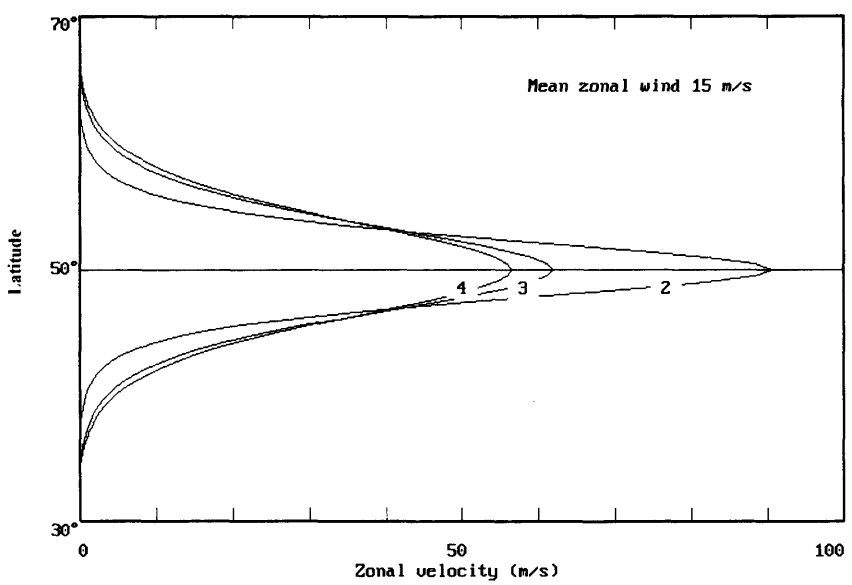

Fig. 3. Meridional profiles of zonal wind for $U_m = 15$ ms^{-1} and $N=2$, 3 and 4.

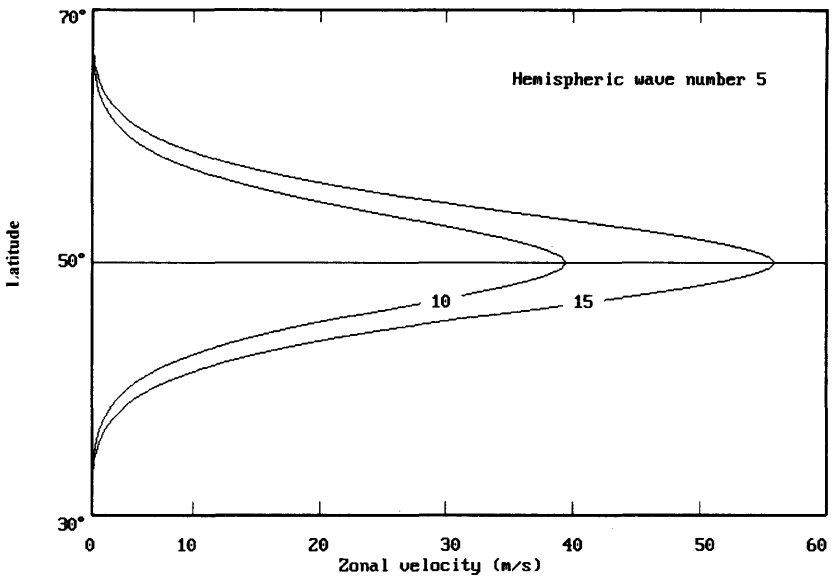

Fig. 4. Meridional profiles of zonal wind for $N = 5$ and $U_m = 10$ and 15 ms^{-1}.

dU/dy in comparison with f_0 in the divergence term in the definition of the potential vorticity equation (Holton, 1992), a test computation has been made including this term. The shear south of the jet is then reduced and instead increased to the north of the jet. But at the same time, the jet is moved to the northern half of the channel with very little wind in the southern half, clearly indicating that other neglected terms, probably friction and possibly also the latitudinal variation of β, should also be considered. This has not been done here since it would considerably complicate the computations, which have all been made without dU/dy in the divergence term.

Fig. 3 may give the impression that the lowest wave numbers give the most pronounced jet. This is however not the case due to the relations (3.6). For the present values of U_m, the profiles for $N = 5$ or 6 are the flattest and for $N = 8$, one obtains almost the same profile as for $N = 4$. It is again seen that the zonal scale has little influence on the meridional scale for the neutral solution.

Computed phase velocities, c_1, are given in Table 3 for $U_m = 10$ and in Table 4 for $U_m = 15$ ms^{-1}. N_1 is the prescribed hemispheric number of waves and N_2 and the phase velocity c_2 are computed from eqs. (3.6). Also given are the maximum zonal velocity, U_{max}, and the mean zonal kinetic

Table 3. *Prescribed wave number N_1 and computed phase velocity c_1 in ms^{-1} for $U_m = 10$ ms^{-1}*

N_1	2	3	4	5	6	7	8	9	10
c_1	0.4	1.6	3.9	6.6	9.6	12.5	15.5	18.3	21.1
N_2	25.1	12.2	7.9	6.0	5.0	4.4	4.0	3.6	3.4
c_2	61.7	27.4	15.4	9.9	6.8	5.0	4.0	3.0	2.5
c_R	−10.0	−4.2	−0.1	2.6	4.4	5.7	6.6	7.2	7.7
U_{max}	82.4	48.3	40.5	38.7	38.6	39.4	40.5	42.0	43.7
K	27.65	16.36	13.85	13.28	13.26	13.50	13.86	14.32	14.87

Corresponding second wave number N_2 and phase velocity c_2 computed from (3.6). The Rossby phase velocity, c_R, is computed from eq. (3.9). U_{max} is maximum velocity in ms^{-1}, and K, the average zonal kinetic energy in 10^5 Jm^{-2}.

Tellus A/48/1-4

Table 4. *Same as Table 3 but with* $U_m = 15\ ms^{-1}$

N_1	2	3	4	5	6	7	8	9	10
c_1	0.8	3.1	6.7	10.9	15.2	19.6	23.8	28.0	32.1
c_R	−5.0	0.8	4.9	7.6	9.4	10.7	11.6	12.2	12.7
N_2	17.3	8.8	6.0	4.7	4.0	3.5	3.2	2.9	2.8
U_{max}	90.7	61.9	56.5	55.7	56.4	57.9	59.9	62.3	64.9
K	45.84	31.72	29.16	28.78	29.13	29.84	30.77	31.90	33.17

The second phase velocity is not given; it is the same as in Table 1.

energy, K. For comparison we have also included the Rossby-Haurwitz phase velocity c_R, computed from

$$c_R = U_m - \frac{\beta}{k^2 + m^2}. \tag{3.9}$$

The difference between the phase velocities c_1 and Rossby phase velocities, c_R in Tables 3 and 4, is indeed very great. For the longest waves, we find no negative phase velocities, instead these waves tend to be quasi-stationary as is usually observed in the atmosphere. Topographic influence, as suggested by Charney and Eliassen (1949), is therefore not required in order to keep these long neutral waves stationary. On the other hand the influence from the earth's surface may well be the factor that determines a preferred longitudinal position of the waves. The influence of topography is discussed in Hansen et al. (1989), where it was found that its role in the general circulation may be more significant for the traveling large-scale waves (wave number $\leqslant 5$) than for standing waves.

In the Tables 3 and 4 it is seen that the shorter waves move much faster than the Rossby waves, for $N_1 \geqslant 7$ even faster than the mean zonal wind, U_m. In view of these differences, a verification has been made using the power spectral density diagrams given in Hansen et al. (1989) for the traveling variance determined by the Hyashi approach. Their Figs. 2a and 7a are reproduced here in Figs. 5 and 6 with the addition of plus signs (+) indicating the periods corresponding to the phase velocities in Tables 3 and 4. Fig. 5 shows the result for the Northern Hemisphere winter, where periods correspond to $U_m = 10\ ms^{-1}$ and

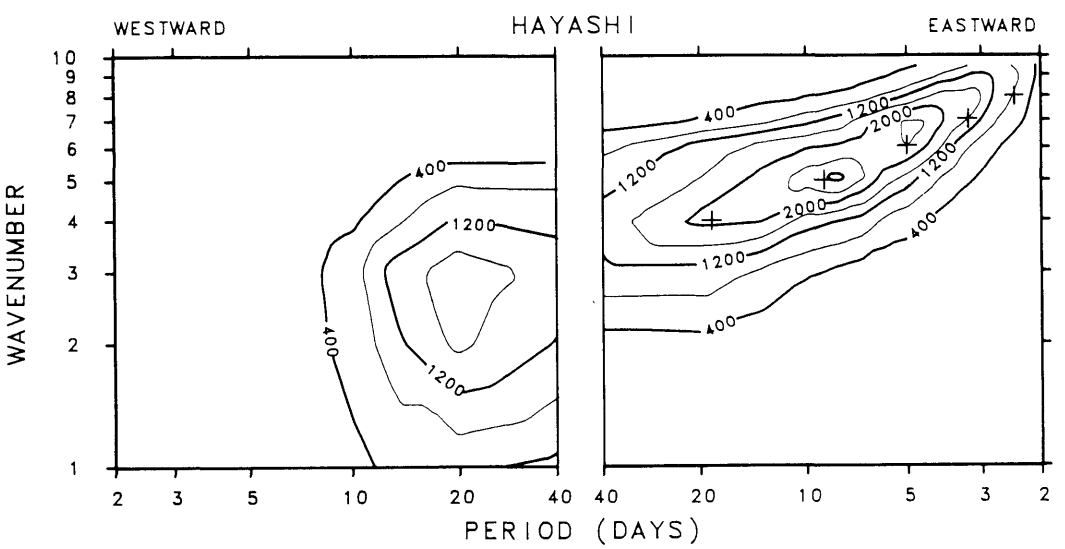

Fig. 5. Power spectral density of NH winter 500 mb geopotential height averaged between 40°N–60°N, reproduced from Hansen et al. (1989), with addition of plus signs giving computed periods corresponding to $U_m = 10\ ms^{-1}$.

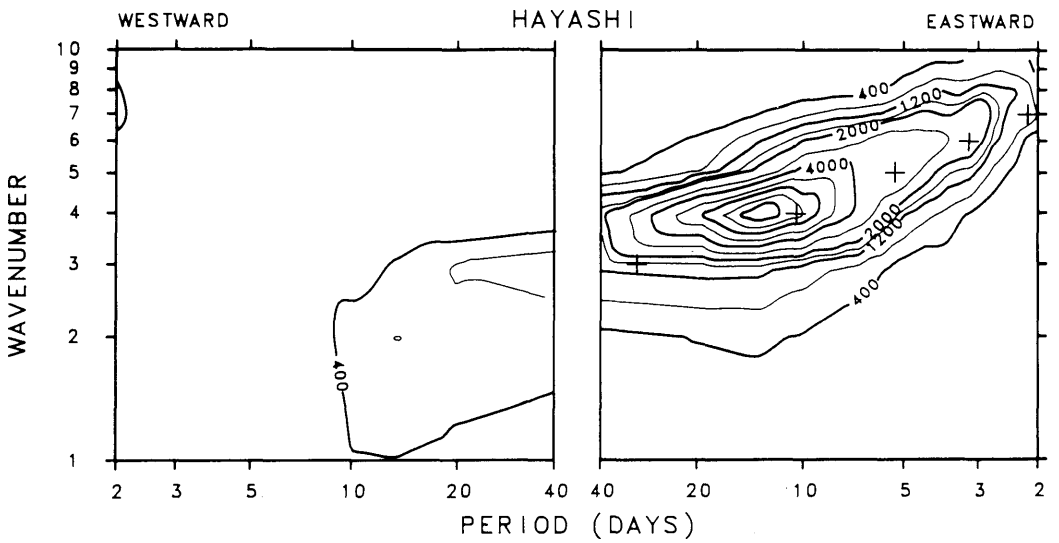

Fig. 6. Same as Fig. 5 but for SH winter and with $U_m = 15\ \mathrm{ms}^{-1}$.

Fig. 6 for Southern Hemisphere winter with $U_m = 15\ \mathrm{ms}^{-1}$. It is seen that computed periods fit surprisingly well the observed data.

For completeness, it is mentioned that for a given mean velocity, U_m, the 2nd phase velocity, c_2, when plotted against N_2, falls on the same curve as c_1 plotted against N_1.

One feature of particular interest in the Figs. shown by Hansen et al. (1989) is that the maximum power spectral density, which here is taken to correspond to the most frequent circulation pattern, is found at hemispheric wave number 4 in the SH winter and at hemispheric wave number 5 in the NH winter. This could indicate that the thermal forcing, higher in the SH, determines the preferred wavelength in much the same way as in rotating annulus experiments. A similar tendency may be seen in the computed zonal kinetic energy, where a minimum is found at $N_1 = 6$ for $U_m = 10\ \mathrm{ms}^{-1}$ and at $N_1 = 5$ for $U_m = 15\ \mathrm{ms}^{-1}$. Computations made with $U_m = 5$ and $20\ \mathrm{ms}^{-1}$ confirm that an increase in the thermal forcing results in an increase of the wave length of minimum zonal kinetic energy. Changing U_m in small increments, we find the minimum kinetic energy at $N = 4.5, 4.9, 5.5$ and 6.8 for $U_m = 20$, 15, 10 and $5\ \mathrm{ms}^{-1}$ respectively. On the assumption that perturbations are initially sufficiently small for their kinetic energy to be negligible in comparison with the zonal kinetic energy, it is difficult to see why the flow would be established at a higher kinetic energy level than these minima. The wave numbers for the minima are here one unit higher than those found in Hansen et al. but the tendency is the same. The possibility that the preferred wavelength should depend on the meridional temperature difference or on a zonal index is only indicated here and will be further investigated separately.

The perturbation kinetic energy depends on the value of the coefficient α_0 in (3.4). Under the assumption that the amplitude in the north-south direction of the central streamline is the same for all wave numbers, we find minimum perturbation kinetic energy at $N = 4$ for $U_m = 15$ and between 4 and 5 for $U_m = 10\ \mathrm{ms}^{-1}$. The tendency for a minimum at intermediate wavelengths is also found in the rms values of the rate of loss of vorticity, which has values of the order 10^{-9} at $N = 4$ or 5. Since the rms of vorticity is of the order 10^{-5}, this indicates that under suitable initial conditions (a neutral wave) the linearized equation will rather correctly describe the evolution of the flow for a considerable number of timesteps.

With the values of c_1 and c_2 given in Table 4 with $U_m = 15\ \mathrm{ms}^{-1}$ we may now use (3.7) and (3.8) to calculate phase velocities and wave numbers for waves that have the same normalized meridional profile. As an example we give in Table

Table 6

N	2	3	4	5	6	7	8	9
U_m	31.7	15.0	9.3	6.8	5.7	5.6	7.1	20.5

6 the wave numbers, N, and the average zonal velocities U_m that have the same normalized zonal profile.

It is seen that in the case of the neutral solution in a beta-plane model, the same normalized profile is found at short and long waves and also that the meridional forcing is important in the determination of this profile. This effect cannot be taken into consideration in normal modes determined as perturbations on a basic state of rest.

5. Higher modes

5.1. *The equation*

In order to determine higher modes a possible procedure would be to consider small secondary perturbations on one of the zonal flows determined in the previous section, disregarding wave-wave interaction. The equation to be used would then be (3.1) with $U_1(y, p)$ given. With c different from c_1 and c_2, the equation will have complex solutions with profiles in the vertical and meridional directions that will depend on the three parameters, k_1, U_m and the new perturbation wave number k. Since, furthermore, different vertical profiles will have to be considered, this approach would give a very large number of different orthogonal sets for expansion of meridional profiles.

In order to avoid this multitude we shall proceed in a different way. Noting that eq. (3.3) may be written

$$\frac{\partial^2 U_1}{\partial y^2} + f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial U_1}{\partial p}\right) = U_1(a - bU_1), \tag{5.1}$$

we may mathematically define a function $A(y, p)$ by the equation

$$\frac{\partial^2 A}{\partial y^2} + f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial A}{\partial p}\right) = A(\lambda - bU_1) \tag{5.2}$$

where λ is an eigenvalue to be determined and b and U_1 are the same as in (5.1). It should be stressed that this is a purely mathematical definition since, in order to satisfy boundary conditions and admit a complete set of solutions, λ has to be negative, except for $\lambda = a$, which is (5.1). From (3.1), it then follows that this would correspond to a complex k, which is contrary to our assumptions. It is easily shown by standard methods that (5.2) gives orthogonal eigenfunctions, which also are orthogonal to U_1, provided boundary conditions are similar.

An alternative to (5.2) is to retain $(a - bU_1)$ in the parenthesis as in (5.1) and multiply the right-hand side by $-\lambda$. The solutions would then instead be orthogonal with respect to the weighting function $(a - bU)$. This has not been tested.

Since it is not possible to separate the variables in eq. (5.2) we proceed in the same way as before. Taking

$$A(y, p) = B_n(y)\, F_n(p),$$

where we take $F_n(p)$ to be normalized, we require that each term in the series should satisfy (5.2) as well as possible in the least square sense. The variation of the corresponding integral with respect to B_n is similar to the procedure given in the Appendix and gives the equation

$$\frac{d^4 B_n}{dy^4} + q_{1n}\frac{d^2 B_n}{dy^2} + q_{2n}\left(\frac{d^2 B_n U}{dy^2} + U\frac{d^2 B_n}{dy^2}\right) + q_{3n}B_n + q_{4n}B_n U + q_{5n}B_n U^2 = 0, \tag{5.3}$$

where

$$q_{1n} = 2(\overline{F_n F_{nd}} - \lambda); \qquad q_{2n} = b\overline{F_1 F_n^2};$$
$$q_{3n} = \overline{F_{nd}^2} - 2\lambda\overline{F_n F_{nd}} + \lambda^2;$$
$$q_{4n} = 2b(\overline{F_1 F_n F_{nd}} - \lambda\overline{F_1 F_n^2}); \qquad q_{5n} = b^2\overline{F_1^2 F_n^2}$$

and where F_1 is the vertical mode of the zonal flow, F_n the vertical mode of the perturbation and

$$F_{nd} = f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial F_n}{\partial p}\right).$$

Also here the variation gives the natural boundary condition $dB_n/dy = 0$ at $y = 0, W$. Values of the vertical averages are given in Table 5, using the same F_1 and F_2 as before.

Table 5. *Vertical averages for the second mode*

$\overline{F_1F_2^2}$	$\overline{F_1^2F_2^2}$	$\overline{F_2F_{2d}}$	$\overline{F_1F_2F_{2d}}$	$\overline{F_{2d}^2}$
0.947	1.126	-1.043×10^{-11}	-8.915×10^{-12}	1.571×10^{-22}

5.2. Results

In order to solve (5.3) we need to prescribe the zonal flow which is defined by the wave number N of the primary perturbation and by the mean velocity, U_m. We also need to specify the vertical mode, F_n, of the secondary perturbation and we shall refer to the solution by the notation $B(N, U_m, n)$. Three cases have been computed. First $B(5, 15, 1)$, which represents the least pronounced jet, and second and third $B(7, 10, 1)$ and $B(7, 10, 2)$. The jet is here more pronounced but not extreme. The first three modes of $B(5, 15, 1)$ are shown in Fig. 8 together with the primary mode which here is normalized. Fig. 9 shows the same for the case $B(7, 10, 2)$. It is seen that the modes are not identical but very similar and a natural question is therefore if it could be possible to use only one set of orthogonal functions for the expansions in the meridional direction. In order to test this, expansions have been made of the odd function $U^2\,dU/dy$ from (3.5) computed with $N=5$, $U_m=15$, using the three different sets of odd orthogonal functions. A comparison is also made with an odd sine-function expansion since, for large negative λ, the solution to (5.3) will approach sine-functions. The result is given in Table 7.

It is seen that the three sets of eigenfunctions determined from eq. (5.3) are much more efficient than the odd sine-functions for which it takes 15 terms to arrive at 99.9% explained variance. It is also seen that the functions $B(7, 10, 2)$ are most efficient. It is a set determined with a different zonal wave number, different mean velocity and a higher vertical mode.

In a second test, we have expanded the meridional profile of the most concentrated jet, $U_m=10$, $N=2$, using the eigenfunctions computed for the least concentrated jet, $U_m=15$, $N=5$. With the 80-point grid, it is found that a 99% variance reduction is obtained with only 4 terms and 99.9% with 6 terms.

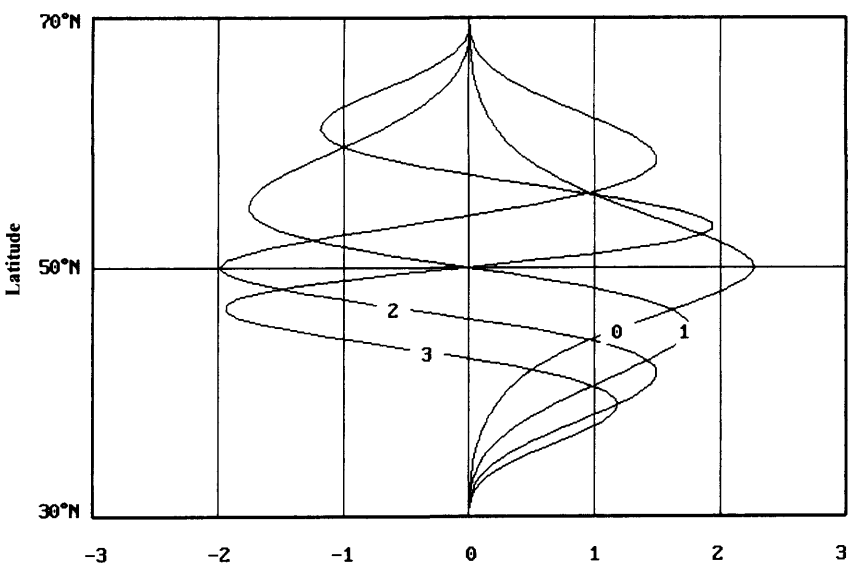

Fig. 7. First three higher modes for $B(5, 15, 1)$ together with the normalized primary mode.

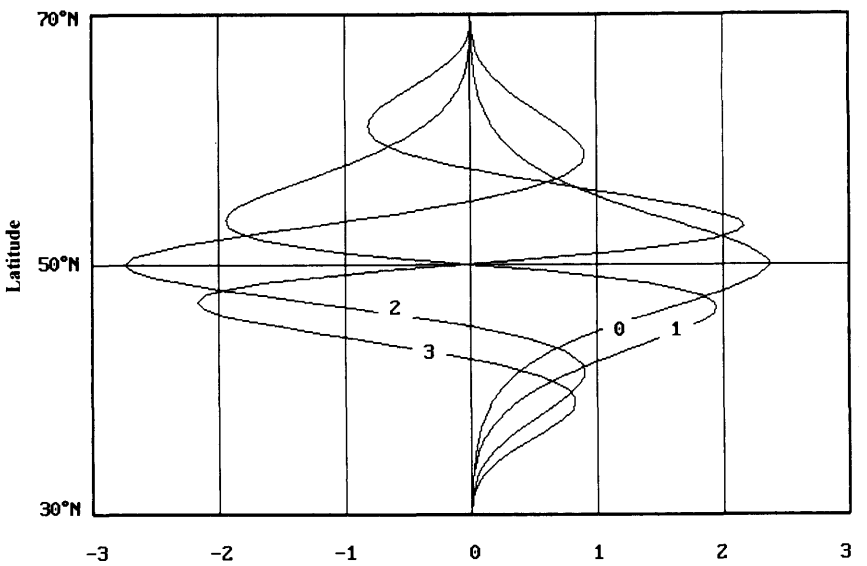

Fig. 8. Same as Fig. 7 but for $B(7, 10, 2)$.

6. Summary and conclusion

The normal modes used in present day spectral primitive equation models for computation and data assimilation are derived as eigenfunctions to the shallow water equation, perturbed about a state of rest. These eigenfunctions satisfy the linearized equation exactly. Since the atmosphere is certainly not in a state of rest, one may question the efficiency of these normal modes for expanding atmospheric or computed data. An attempt has therefore been made to determine the equivalence of normal modes from the non-linear equations, using the condition that these modes should satisfy the equations, not exactly as in the linear case, but as well as possible in the least square sense.

This variational approach has here been applied to the potential vorticity equation in a beta-plane Cartesian coordinate system with walls at 30° N and 70° N. At first the variables x and t have been separated from y and p. It was found that the variational integral could be considerably simplified by choosing trigonometric expansion functions in the zonal direction. With this, the integral splits into two parts, the first having as squared integrand the terms of the perturbation potential vorticity equation and thus vanishing exactly if the perturbation streamfunction and the zonal flow satisfy this equation. By this approach the perturbation equation is derived without any assumption concerning the amplitude of the perturbations and is therefore applied to finite amplitudes waves. The second integral, which is found to have a minimum for a suitable choice of the zonal wave number, gives also a measure of how well the linearized equation approximates the non-linear evolution. It was found that the linearized equation could well be used for a considerable time provided suitable initial conditions exist.

Table 7. *Variance explained by M first odd functions in expanding* $U^2\, dU/dy$

M	1	2	3	4	5	6	7
$B(5, 15, 1)$	50.1	78.0	91.8	97.2	99.1	99.8	100.0
$B(7, 10, 1)$	56.7	82.5	94.1	98.3	99.6	100.0	100.0
$B(7, 10, 2)$	76.7	89.0	96.3	99.2	100.0	100.0	100.0
sine	11.8	40.3	68.8	86.6	94.8	98.0	99.1

The perturbation vorticity equation obtained in this way gives only one relation between the zonal flow and the perturbations. In order to have a closed system a second relation or equation is required. It is found that this is provided simply by the requirement that the perturbations should be

neutral. With this condition a second-order partial differential equation is obtained that determines the zonal wind $U_1(y, p)$ as a function of wave number and meridional forcing. The perturbation streamfunction has the same vertical and meridional profiles as the zonal wind. In $U_1(y, p)$ the variables are again separated into $U(y)\,F(p)$ and the same variational method is used to determine the meridional profile, $U(y)$, of the zonal wind. This gives a fourth-order ordinary but non-linear differential equation. The vertical function $F(p)$ is taken to be either the first or the second vertical structure function, which are here determined from computed temperature profiles.

Basic solutions to the fourth-order equation are computed for hemispheric wave numbers 2–10 and for a forcing corresponding approximately to a height difference between the southern and northern walls of 500 and 750 m at 500 hPa. These correspond respectively to Northern and the Southern Hemisphere winter conditions. The zonal wind is found to have a typical jet-like meridional profile. Together with the perturbation wind components it gives a realistic, meandering westerly flow with waves moving only east. For different wave numbers, the meridional profiles of the zonal wind show a great similarity and are for instance for hemispheric wave numbers 4 and 8 almost identical. In this model a clear scale interdependence therefore does not exist between the zonal scale and the meridional or vertical scales.

Computed phase velocities show that the longest waves are stationary as is generally observed in the atmosphere. A topographical influence is therefore not necessary for reducing their phase velocity. For shorter waves, computed phase velocities are much greater than those obtained for Rossby waves. They agree, however, surprisingly well with observations.

For a meridional forcing corresponding to a mean geostrophic zonal wind of 5, 10, 15 and 20 ms^{-1} the kinetic energy of the zonal wind varies considerably with the prescribed zonal wave number. It is very high for the shortest and the longest waves and has a minimum at an intermediate wave length. On the hypothesis that, for a given forcing, the flow would not establish itself at a higher kinetic energy level than this minimum, we find here an interdependence between forcing and wavelength which seems to correspond to what is found in rotating annulus experiments. An increase in the meridional forcing increases the wavelength with minimum kinetic energy.

In order to determine higher modes for the expanding functions a natural method would be to consider perturbations on one of the zonal flows obtained previously. However, in this case the perturbation equation would require a complex solution and with all the parameters involved, one would end up with a very great number of expanding functions. A simple mathematical definition of an orthogonal set has therefore been used for determining higher orthogonal modes for expansion of meridional profiles. These have been tested in an expansion of the meridional structure of the perturbation Jacobian. It was found that the functions so defined give a much faster convergence than a Fourier expansion and furthermore, that different sets based on different zonal wave numbers for the primary perturbation performed almost equally well. In this model it would therefore seem unnecessary to have a separate set of meridional expansion functions for each zonal wave number.

To what extent these results apply to the atmosphere is not known and further research is certainly required. Here valuable information could be obtained from data by expanding global data of for instance observed geopotential into EOFs of the form $\Phi(\lambda, \varphi, p, t) = \sum A_n(\lambda, t)\,F_n(\varphi, p)$, where λ and φ are longitude and latitude.

In conclusion it may be said, that the variational approach to non-linear equations has been tested and found to be efficient, when it comes to separate the variables and determine alternatives to the normal modes. Furthermore, that in spite of differences in geometry and boundary conditions, the neutral solutions to the perturbation potential vorticity equation have a number of characteristic properties that correspond to what is found in the atmosphere and in rotating annulus experiments, and therefore is theoretically interesting. It will therefore be further investigated, in particular with regard to the implication and possible explanation of the minimum kinetic energy condition. Here work has already been done and will soon be reported. Other questions that have been left open are the application of the method to westward moving waves, the determination of vertical structure functions and finally the generalization to primitive equations on a spherical earth. These will be considered in later papers.

7. Acknowledgment

The author is indebted to two unknown reviewers for their criticism and suggestions. These resulted in a revision of the paper. Thanks are also due to Dr. S. Bodin for carefully reading and commenting on the revised version.

8. Appendix

The variation of the integral with respect to U is

$$\delta I_U = \frac{2}{WP}\int_0^W \int_0^P \left[RF\,\delta\frac{\mathrm{d}^2U}{\mathrm{d}y^2} + R\,\delta U(F_\mathrm{d} - aF + 2bUF^2)\right] \mathrm{d}y\,\mathrm{d}p.$$

The first term in the brackets integrated twice by parts gives

$$\int_0^W \int_0^P FR\,\delta\frac{\mathrm{d}^2U}{\mathrm{d}y^2} = \int_0^P \left[FR\,\delta\frac{\mathrm{d}U}{\mathrm{d}y}\right]_0^W \mathrm{d}p - \int_0^P \left[F\frac{\partial R}{\partial y}\delta U\right]_0^W \mathrm{d}p + \int_0^P \int_0^W F\frac{\partial^2 R}{\partial y^2}\delta U\,\mathrm{d}y\,\mathrm{d}p.$$

The first term on the rhs will vanish if we add the boundary condition $\mathrm{d}U/\mathrm{d}y = 0$ at $y = 0, W$. The second term vanishes since $U = 0$ at $y = 0, W$. The variation of I depends now only on the arbitrary δU, so it remains only to carry out the multiplications and the integrations with respect to p in order to obtain (4.2).

REFERENCES

Baer, F. 1974. *Vertical structure of hemispheric temperature spectral components.* Report DM-13, Department of Meteorology, University of Stockholm, Sweden.

Bradley, J. H. S. and Wiin-Nielsen, A. 1968. On the transient part of atmospheric planetary waves. *Tellus* **20**, 533–544.

Craig, R. A. 1945. A solution to the nonlinear vorticity equation for atmospheric motion. *J. Meteor.* **2**, 175–178.

Daley, R. 1978. Variational non-linear normal mode initialization. *Tellus* **30**, 201–218.

Hansen, A. R., Sutera, A. and Venne, D. 1989. An examination of mid latitude power spectra: evidence for standing variance and the signature of El Nino. *Tellus* **41A**, 371–384.

Holmström, I. 1963. On a method for parametric representation of the state of the atmosphere. *Tellus* **15**, 127–149.

Holmström, I. 1977. Optimization of atmospheric models. *Tellus* **29**, 415–421.

Holton, J. R. 1992. *An introduction to dynamic meteorology*, 3rd edition. Academic Press, 390 pp.

Kasahara, A. 1976. Normal modes of ultralong waves in the atmosphere. *Mon. Wea. Rev.* **104**, 669–690.

Kasahara, A. 1980. Effect of zonal flows on the free oscillations of a barotropic atmosphere. *J. Atmos. Sci.* **37**, 917–929.

Lorenz, E. N. 1967. *The nature and the theory of the general circulation of the atmosphere.* World Meteorological Organization. Geneva, Switzerland. 142 pp.

Neamtan, S. M. 1946. The motion of harmonic waves in the atmosphere. *J. Meteor.* **3**, 53–56.

Palmén, E. and Newton, C. W. 1969. *Atmospheric circulation systems.* Academic Press, New-York.

Tanaka, L. H., Kung, E. C. and Baker, W. E. 1986. Energetics analysis of the observed and simulated general circulation using three-dimensional normal mode expansion. *Tellus* **38A**, 412–428.

Tanaka, L. H. and Kung, E. C. 1989. A study of low-frequency unstable planetary waves in realistic zonal and zonally varying basic states. *Tellus* **41A**, 179–199.

Tanaka, L. H. and Kasahara, A. 1992. On the normal modes of Laplace's tidal equation for zonal wave number zero. *Tellus* **44A**, 18–32.