Q. J. R. Meteorol. Soc. (2002), **128**, pp. 679–694

On shallow water potential vorticity inversion by Rossby-number expansions

By A. R. MOHEBALHOJEH*

University of St Andrews, UK

(Received 23 March 2001; revised 29 August 2001)

Summary

For the f-plane shallow water (SW) equations, the basic properties of formal Rossby-number expansions and potential vorticity (PV) inversion by the implied asymptotic balance conditions are re-examined. An indeterminacy inherent in conventional expansions beyond leading order geostrophic balance is highlighted. The indeterminacy gives us the freedom to define, in principle, infinite sets of "balance conditions" relating inertia–gravity modes to Rossby modes. Examples discussed include balance conditions that are obtained when one of the variables, linearised PV, vorticity, or perturbation height is regarded purely as a leading-order geostrophic quantity.

For the sake of conservation of global mass and boundary circulation, SW PV inversion is defined as one that employs the PV anomaly, i.e. the deviation of PV from its domain area average. SW PV inversion by means of various Rossby-number balance conditions and in the sense of Warn *et al.* (1995) are presented. In the latter, the PV anomaly is regarded purely as the leading-order geostrophic quantity and the Rossby formula for SW PV anomaly is satisfied asymptotically. A modification of the inversion procedure of Warn *et al.* satisfying the exact SW equation for the PV anomaly is also presented.

Numerical results for these PV inversion procedures, accurate up to second order in Rossby number, are presented for two SW simulations involving weakly and strongly ageostrophic flows. Also presented are results from PV inversion using the δ–γ hierarchy of balance conditions, which sets the N-th order time derivatives of divergence and ageostrophic vorticity to zero. It is shown that the Warn *et al.* procedure incurs a significant loss of accuracy relative to both the δ–γ hierarchy and other asymptotic procedures like the one that regards the linearised PV as a leading-order geostrophic quantity. The modified Warn *et al.* procedure shows little improvement. The inaccuracy of this procedure stems from a poor leading-order solution for the balanced fields.

Keywords: Potential Vorticity Inversion Balance Condition Modified Expansion

1. Introduction

The work of Warn *et al.* (1995) (hereafter WBSV) identified an asymptotic disordering or nonuniformity of conventional Rossby-number expansions. As a way of circumventing this, WBSV developed the idea of a "modified expansion" where a certain master variable is left unexpanded. Using the potential vorticity (PV) as the master variable, they applied their modified expansion to the f-plane shallow water (SW) equations to derive PV-conserving balanced models, and thus SW PV inversion operators, up to second-order accuracy in Ro. The same procedure has been followed by Vallis (1996), and in a different context by Muraki *et al.* (1999) and Rotunno *et al.* (2000). Details on the modified expansion for other variables like pressure can be found in Bokhove (1997) for the hydrostatic Boussinesq equations.

The modified Rossby-number expansion for PV, hereafter called the "WBSV procedure", has certain features which make it unique among PV inversion methods, if we exclude simple geostrophic inversion. The most important feature is that, at every order, all of the inversion equations are linear. The other feature is that the balance conditions are not *a priori* known. Instead, the balance conditions themselves are generated as part of the WBSV procedure, as explained in detail below. A further feature of this procedure is that the inverted fields satisfy the Rossby equation for PV only asymptotically.

This paper assesses the performance of the WBSV procedure in the wider context of PV inversion by means of Rossby-number expansions and more general non-asymptotic balance conditions. To address the issue, section 2 re-examines conventional

* Corresponding author: Institute of Geophysics, Tehran University, P.O. Box 14155 6466, Tehran, Iran.

Rossby-number expansions and highlights the indeterminacy of such expansions beyond leading-order geostrophy. In Section 3, the indeterminacy is exploited to design PV inversion methods by means of balance conditions obtained by a modified expansion of either linearised PV, vorticity, or perturbation height. Section 4 introduces the WBSV procedure and a modified version of it which recovers the exact Rossby formula for the PV anomaly. Section 5 presents numerical results for various measures of imbalance generated during the evolution of weakly and strongly ageostrophic flows at respectively $|Ro|_{\max} = 0.18,\ 0.88$, where $|Ro|_{\max}$ is the maximum of $|\zeta/f|$ over the domain. In order to probe the optimal measures of imbalance, some results are presented from a hierarchy of balance conditions that sets N-th order time derivatives of divergence and ageostrophic vorticity to zero. Finally, Section 5 gives the main conclusions.

2. Rossby-number-expansion balance conditions

In almost any theoretical treatment of balance, a small parameter is involved, which can be used to derive formal asymptotic expansions. One such expansion is the Rossby-number (Ro) expansion as applied to the shallow water momentum and height equations (see Pedlosky 1987, §3.12). The basic assumptions are that both the Rossby number $Ro = U/fL$ and the Froude number $Fr = U/\sqrt{gH}$ are small and of the same order so that the Burger number $Bu = L_{\mathrm{R}}^2/L^2$ is of order unity. Here, U and L denote appropriate velocity and length scales, respectively, f is the Coriolis parameter taken to be constant (f-plane), g is the acceleration due to gravity, H is the mean free surface height, and $L_{\mathrm{R}} = \sqrt{gH}/f$ is the Rossby deformation radius. For the sake of brevity, following Vallis (1996) we put the small nondimensional parameter Ro in curly brackets where it would appear in the nondimensional equations. A formal Rossby-number expansion can be easily carried out on the vorticity, divergence, and continuity equations as written

$$-f\delta = \{Ro\}\left(\frac{\partial \zeta}{\partial t} + \nabla.(\mathbf{v}\zeta)\right) \tag{1a}$$

$$\gamma = \{Ro\}\left(-2J(u,v) + \nabla.(\mathbf{v}\delta) + \frac{\partial \delta}{\partial t}\right) \tag{1b}$$

$$H\delta = \{Ro\}\left(-\frac{\partial h}{\partial t} - \nabla.(\mathbf{v}h)\right) \tag{1c}$$

where h is the perturbation height, $\mathbf{v}$ is the velocity vector, ζ is the relative vorticity, δ is the divergence, $\gamma = f\zeta - g\nabla^2 h$ is the ageostrophic vorticity multiplied by f, and $J(u,v) = (\partial u/\partial x \partial v/\partial y - \partial u/\partial y \partial v/\partial x)$. Taking the time derivative of γ and using (1a) and (1c) it is easy to arrive at the equation

$$\mathcal{H}\delta = \{Ro\}(f\nabla.(\mathbf{v}\zeta) - g\nabla^2\nabla.(\mathbf{v}h)) + \{Ro^2\}(-2J(u,v)_t + \nabla.(\mathbf{v}\delta)_t + \delta_{tt}) \tag{2}$$

where $\mathcal{H} = gH\nabla^2 - f^2$ is the modified Helmholtz operator. Also, (1a) and (1c) can be combined into an exact equation for the linearised potential vorticity $q_\ell = \zeta - fh/H$,

$$\{Ro\}\left(\frac{\partial q_\ell}{\partial t} + \nabla.(\mathbf{v}q_\ell)\right) = 0\ . \tag{3}$$

The $\{Ro\}$ in front of the left-hand side of (3) has been retained in order to maintain consistency with the asymptotic expansions of the momentum-continuity equations (Pedlosky 1987). Since we are concerned only with the balance conditions and their order of accuracy, retaining or otherwise of $\{Ro\}$ has no significance.

Let us work with the SW equations in the $(q_\ell\ ,\ \delta\ ,\ \gamma)$ representation. Such a representation is appropriate, since q_ℓ and $(\delta\ ,\ \gamma)$ represent the Rossby and inertia–gravity modes of the f-plane SW equations linearised around a rest state. We may thus regard a balance condition as a diagnostic relation between $(\delta\ ,\ \gamma)$ and q_ℓ. The formal Ro expansion of (2), (1b), and (3) order by order provides such diagnostic relations as follows. At $\mathcal{O}(Ro^0)$, we have $\delta = \gamma = 0$, which is simply geostrophic balance, and $q_\ell = q_{\ell 0}$. At $\mathcal{O}(Ro^1)$, we have

$$\begin{aligned} \mathcal{H}\delta_1 &= f\nabla.(\mathbf{v}_0\zeta_0) - g\nabla^2\nabla.(\mathbf{v}_0 h_0) && \text{(4a)} \\ \gamma_1 &= -2J(u_0, v_0) && \text{(4b)} \end{aligned}$$

where the second term on the right hand side of (4a) is zero for the f-plane problem. At this order, the prognostic equation for q_ℓ becomes

$$(\partial/\partial t + \mathbf{v}_0.\nabla)q_{\ell 0} = 0, \tag{5}$$

which is just material conservation of quasigeostrophic (QG) PV. If we stop at $\mathcal{O}(Ro^1)$, the QG relations (4a) and (4b) are not sufficient to relate $(\delta_1\ ,\ \gamma_1)$ to $q_\ell = q_{\ell 0} + q_{\ell 1}$. To determine balance conditions, one has to make a choice for $q_{\ell 1}$. This is exactly the indeterminacy referred to in the introduction. Equivalently, a knowledge of $q_{\ell 0}$ leaves the fields of $q_{\ell 1}$ as well as ζ_1 and h_1 undetermined.

This indeterminacy of QG balance has been explicitly used by Phillips (1960) in the context of initialisation. He considered separately cases where either $h_1 = 0$ or $\zeta_1 = 0$ corresponding to, respectively, taking the observations of height or rotational wind as given to determine the other fields. Also, the QG momentum equations used by Hoskins *et al.* (1978) take the pressure field as given ($h_1 = 0$). Another plausible choice is to set $q_{\ell 1} = 0$. Setting $q_{\ell 1} = 0$ makes QG balance equivalent with the first iteration of the first-order nonlinear normal-mode initialisation introduced by Machenhauer (1977). The equivalence was first shown by Leith (1980) for the f-plane hydrostatic Boussinesq equations. For the f-plane SW equations, the Machenhauer initialisation scheme can be carried out in physical-function space if we invert q_ℓ by means of the balance conditions $\partial\delta/\partial t = 0$ and $\partial\gamma/\partial t = 0$ (see Lynch 1989, Mohebalhojeh and Dritschel 2001, and references therein). The field q_ℓ remains fixed during the iterations involved in the initialisation (inversion), which is equivalent with $q_{\ell 1} = 0$. It should be noted here that it is not our objective to discuss, for example, variational forms of Machenhauer and higher-order schemes whereby $q_{\ell 1}$ changes during initialisation. Having highlighted this indeterminacy, the three choices $q_{\ell 1} = 0$, $\zeta_1 = 0$, and $h_1 = 0$ are simply made to close the equations. The choices are nothing but "modified expansions" for, respectively, the linearised PV, vorticity, and perturbation height.

Let us continue the formal asymptotic procedure to $\mathcal{O}(Ro^2)$. From equations (2) and (1b), we obtain at $\mathcal{O}(Ro^2)$

$$\begin{aligned} \mathcal{H}\delta_2 &= f\nabla.(\mathbf{v}_0\zeta_1 + \mathbf{v}_1\zeta_0) - g\nabla^2\nabla.(\mathbf{v}_0 h_1 + \mathbf{v}_1 h_0) - 2J(u_0, v_0)_t && \text{(6a)} \\ \gamma_2 &= -2(J(u_0, v_1) + J(u_1, v_0)) + \nabla.(\mathbf{v}_0\delta_1) + \delta_{1t} && \text{(6b)} \end{aligned}$$

where the additional equations required to evaluate the lower order time-derivative terms read, after bypassing some manipulations,

$$\begin{aligned} \frac{\partial q_{\ell 0}}{\partial t} &= -\nabla.(\mathbf{v}_0 q_{\ell 0}) \\ &= \frac{g}{f} J(q_{\ell 0}, h_0) && \text{(7a)} \end{aligned}$$

$$\frac{\partial h_0}{\partial t} = -H\delta_1 \tag{7b}$$

$$\begin{aligned}\frac{\partial \delta_1}{\partial t} &= f\mathcal{H}^{-1}\nabla.(\mathbf{v}_0\zeta_0)_t \\ &= \frac{g^2}{f}\mathcal{H}^{-1}(J(h_{0t},\nabla^2 h_0)+J(h_0,\nabla^2 h_{0t})) \\ &= H\mathcal{H}^{-1}(gJ(\zeta_0,\delta_1)-f\nabla.(\mathbf{v}_0\nabla^2\delta_1)) \end{aligned} \tag{7c}$$

$$J(u_0,v_0)_t = \frac{gH}{f}(J(\delta_{1y},v_0)+J(\delta_{1x},u_0))\ . \tag{7d}$$

Eq. (7b) is the $\mathcal{O}(Ro^1)$ accurate continuity equation.

As may be anticipated, while (6a) and (6b) determine $(\delta_2\ ,\ \gamma_2)$ in terms of $q_{\ell 0}$, they leave $q_{\ell 2}$ and thus h_2 and ζ_2 undetermined. The two relations (6a) and (6b) together with (4a), (4b), and with $\delta_0=\gamma_0=0$, have to be closed. Again, we will consider choices $q_{\ell i}=0$, $\zeta_i=0$, and $h_i=0$, for $(i=1,2)$.

3. Shallow-water potential vorticity inversion

It is now well established that the shallow water potential vorticity q,

$$q=\frac{f+\zeta}{H+h} \tag{8}$$

can be inverted to a remarkable accuracy (e.g. McIntyre and Norton 2000, hereafter MN00) for a broad range of atmospheric–oceanic flows. A consideration of boundary conditions is vital in order to arrive at solvable equations. For periodic boundary conditions to be used in the numerical experiments of section 5, the domain area average of the relative vorticity $\overline{\zeta}$ is zero. But re-arranging (8) implies

$$\overline{\zeta}=H\overline{q}+\overline{qh}-f\ . \tag{9}$$

For a given distribution of q and a given set of balance conditions, (9) may give a nonzero value for $\overline{\zeta}$ which is mathematically inconsistent. Equivalently, (9) cannot be satisfied with a zero value for $\overline{\zeta}$. The remedy is to invert the PV anomaly,

$$q-\overline{q}=\frac{f+\zeta}{H+h}-C \tag{10}$$

where C is the domain area average PV of the inverted fields. In this way the requirements of zero boundary circulation as well as conservation of global mass H are met. For boundary conditions other than periodic, of course, we can use either (8) or (10). However, with (8) we can generally expect $\overline{\zeta}$ for the inverted fields to be different from its primitive-equation SW value. Note also that in the conservative limit $\overline{\zeta}$ is time invariant. Therefore, if we would like to conserve both boundary circulation and global mass, (10) should be used instead of (8).

The asymptotic balance conditions can be used to invert the PV anomaly to the required order. In the rest of this section, for brevity we illustrate details of only $\mathcal{O}(Ro)$ inversions for a periodic domain. Other Rossby-number expansion inversions follow the same pattern described.

Truncating ζ and h to $\mathcal{O}(Ro)$, we rewrite (10) as

$$q-\overline{q}=\frac{f+\zeta_0+\zeta_1}{H+h_0+h_1}-C\ . \tag{11}$$

Disregarding C for the moment, in (11) we have four unknowns: ζ_0, ζ_1, h_0, h_1. The three equations $\gamma_0 = 0$, $\gamma_1 = -2J(u_0, v_0)$, and (11) are, here, closed with either $q_{\ell 1} = 0$, $\zeta_1 = 0$, or $h_1 = 0$. Each case is examined in turn.

(i) $q_{\ell 1} = 0$. The equations to be solved are

$$\mathcal{H}h_0 = Hfq_\ell \tag{12a}$$
$$\mathcal{H}h_1 = 2HJ(u_0, v_0) \tag{12b}$$
$$\zeta = (H+h)(q - \overline{q} + C) - f \tag{12c}$$
$$C = \frac{f - \overline{qh}}{H} \ . \tag{12d}$$

Here, (12a) is the expression of geostrophy in $q_\ell = q_{\ell 0}$; (12b) uses $q_{\ell 1} = 0$ in (4b); (12c) is (11) written in terms of ζ; and (12d) ensures $\overline{\zeta} = 0$. Notice that due to the nonlinearity of PV, the resulting system of equations is nonlinear and must be solved iteratively. A solution procedure for (12) can be envisaged as follows. The starting point is a good first guess for h. When a SW solution is available, the first guess can be the actual h itself. When only PV is known, a zero field can simply be chosen. Then, (12d) and (12c) determine C and ζ, respectively. Calculating q_ℓ from a knowledge of ζ and h determines the right-hand side of (12a). The next step is to solve (12a) for h_0, and thus $(u_0\ ,\ v_0)$, the leading-order geostrophic velocities. By $h = h_0 + h_1$, the solution of (12b) for h_1 completes one cycle of the iterative procedure. The iteration can be terminated when a certain convergence criterion for h, e.g. the maximum absolute value of the difference between two successive solutions for h being smaller than a tolerance value, is met. The divergence field can finally be evaluated using $\delta_0 = 0$ and (4a).

For the sake of comparison, it is also useful to write the $\mathcal{O}(Ro)$ balance condition for γ,

$$\gamma = -2g^2H^2J(\mathcal{H}^{-1}\frac{\partial q_\ell}{\partial x}, \mathcal{H}^{-1}\frac{\partial q_\ell}{\partial y}) \tag{13}$$

(ii) $\zeta_1 = 0$. Taking $\zeta = \zeta_0$, $h = h_0 + h_1$, $\gamma_0 = 0$, and (4b), the $\mathcal{O}(Ro)$ balance condition for γ becomes (cf. 13)

$$\gamma = -2J(\frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y}) \tag{14}$$

where ψ is streamfunction. Note that (14) is exactly the Bolin–Charney balance equation. Equations (14), (12c), and (12d) can be used to solved for h, ζ, and C. Since (14) replaces (12a) and (12b), the algorithm described above has to be modified appropriately.

(iii) $h_1 = 0$. Taking $h = h_0$, $\zeta = \zeta_0 + \zeta_1$, $\gamma_0 = 0$, and (4b), the $\mathcal{O}(Ro)$ balance condition for γ becomes (cf. 13 and 14)

$$\gamma = -2\frac{g^2}{f^2}J(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}) \ . \tag{15}$$

The solution procedure described above when $q_{\ell 1} = 0$ can be modified by replacing (12a) and (12b) with (15).

Comparing (i), (ii), and (iii), the following remarks can be made. First, it is only (ii), i.e. the non-expansion of ζ, that leads to exact equations for steady axisymmetric vortices. Second, for $L \ll L_\mathrm{R}$ and $L \gg L_\mathrm{R}$, (i) reduces to (ii) and (iii) respectively*.

* For $L \ll L_\mathrm{R}$, $q_\ell \approx \zeta$ and $\mathcal{H} \approx gH\nabla^2$. For $L \gg L_\mathrm{R}$, $q_\ell \approx -fh/H$ and $\mathcal{H} \approx -f^2$.

4. Modified expansion for potential vorticity (WBSV procedure)

WBSV start from (8) and define a PV anomaly based on the deviation of PV from a background value f/H. Apart from starting from (10) in order to ensure mathematical consistency, the exposition below closely follows the procedure presented in section 4 of WBSV and section 2 of Vallis (1996). Let us first introduce auxiliary fields $\tilde{\zeta} = \zeta/f$ and $\tilde{h} = h/H$. Further, we normalise the PV anomaly by f/H, i.e. $q - \overline{q} = \frac{f}{H}\{Ro\}q'$. Note that $\overline{q'} = 0$.

The modified expansion for PV is obtained by expanding all variables except q'. Crucially, C needs to be expanded to ensure $\overline{\zeta} = 0$ order by order. Using the expansion $C = \frac{f}{H}(1 + \sum_{i=1}\{Ro^i\}C'_i)$, we can write (cf. eqn. 30 of WBSV)

$$q - \overline{q} = \frac{f}{H}\left(\frac{1 + \{Ro\}\tilde{\zeta}}{1 + \{Ro\}\tilde{h}}\right) - C \qquad \Longleftrightarrow \qquad q' = \frac{\tilde{\zeta} - \tilde{h}}{1 + \{Ro\}\tilde{h}} - \sum_{i=1}\{Ro^{i-1}\}C'_i \ . \tag{16}$$

It is helpful to rearrange the expression for q' in (16) as follows,

$$\left(1 + \{Ro\}\sum_{i=0}\{Ro^i\}\tilde{h}_i\right)\left(q' + \sum_{i=1}\{Ro^{i-1}\}C'_i\right) = \sum_{i=0}\{Ro^i\}\left(\tilde{\zeta}_i - \tilde{h}_i\right) \tag{17}$$

From (17), at $\mathcal{O}(Ro^0)$ we have

$$\tilde{\zeta}_0 - \tilde{h}_0 = q' + C'_1 \ . \tag{18}$$

Equating the domain area average values of the two sides of (18) gives $C'_1 = 0$. Together with the geostrophic balance conditions $\delta_0 = \gamma_0 = 0$, (18) determines the whole zero-order fields. At $\mathcal{O}(Ro^1)$, (17) gives

$$\tilde{\zeta}_1 - \tilde{h}_1 = q'\tilde{h}_0 + C'_2 \tag{19}$$

where the consistency of the domain area average values requires

$$C'_2 = -\overline{q'\tilde{h}_0} \ . \tag{20}$$

Together with the $\mathcal{O}(Ro^1)$ relations (4a) and (4b), (19) and (20) determine the whole first-order fields. Particularly revealing at this stage is the $\mathcal{O}(Ro^1)$ accurate balance condition for γ (cf. (13), (14), and (15))

$$\gamma = -2g^2f^2H^2J(\mathcal{H}^{-1}\frac{\partial q'}{\partial x}, \mathcal{H}^{-1}\frac{\partial q'}{\partial y}) \tag{21}$$

which is obtained by inserting the zero-order field $h_0 = f^2H\mathcal{H}^{-1}q'$ in (4b). In other words, regarding the whole q' as a zero-order field, (18) leaves us with the $\mathcal{O}(Ro^1)$ estimate (21) for γ. The balance condition (21) can be used directly to invert the PV anomaly (10). The resulting PV inversion can be considered a "modified WBSV procedure" (hereafter mWBSV) that employs the exact Rossby formula for PV anomaly (11), not the formal asymptotic truncation of (17). This procedure can be obtained by replacing (19) with

$$\tilde{\zeta}_1 - \tilde{h}_1 = q'(\tilde{h}_0 + \tilde{h}_1) + C'_2(1 + \tilde{h}_0 + \tilde{h}_1) \tag{22}$$

and (20) with

$$C'_2 = -\overline{q'(\tilde{h}_0 + \tilde{h}_1)} \ . \tag{23}$$

Note that the presence of $\tilde{h}_1$ on the right-hand side of (23) makes the mWBSV equation (22) nonlinear. It is instructive to pause and compare mWBSV and the first-order PV inversion procedure that sets $q_{\ell 1} = 0$. These two PV inversion procedures differ only in their balance conditions for γ, that is, the relations (21) and (13). It can be seen that (13) is obtained from (21) if we replace q' by q_ℓ/f, i.e. by its linear truncation. Depending on how well q' can be approximated by its linearisation q_ℓ/f, we expect discrepancies between the behaviour of (21) and (13) and thus the results obtained by the two PV inversion procedures.

The general WBSV equations obtained by truncating (17) at $\mathcal{O}(Ro^n)$, $(n > 0)$ read

$$\tilde{\zeta}_n - \tilde{h}_n = q'\tilde{h}_{n-1} + \sum_{i=2}^{n} C'_i \tilde{h}_{n-i} + C'_{n+1} \tag{24a}$$

$$C'_{n+1} = -\overline{q'\tilde{h}_{n-1}}\,. \tag{24b}$$

The fields of h and ζ obtained by the WBSV procedure accurate to $\mathcal{O}(Ro^n)$ involve an $\mathcal{O}(Ro^{n+1})$ error on the PV anomaly,

$$q - \overline{q} = \frac{f + \sum_{i=0}^{n} \zeta_i}{H + \sum_{i=0}^{n} h_i} - \overline{\frac{f + \sum_{i=0}^{n} \zeta_i}{H + \sum_{i=0}^{n} h_i}} + \mathcal{O}(Ro^{n+1})\,. \tag{25}$$

Again to remove this error, one can define a mWBSV procedure by replacing only the $\mathcal{O}(Ro^n)$ equations (24a) and (24b) with

$$\tilde{\zeta}_n - \tilde{h}_n = q'(\tilde{h}_{n-1} + \tilde{h}_n) + \sum_{i=2}^{n} \sum_{j=0}^{i} C'_i \tilde{h}_{n-i+j} + C'_{n+1}(1 + \sum_{i=0}^{n} \tilde{h}_i) \tag{26}$$

$$C'_{n+1} = -\overline{q'(\tilde{h}_{n-1} + \tilde{h}_n)}, \tag{27}$$

while using the WBSV equations for $\mathcal{O}(Ro^0)$ to $\mathcal{O}(Ro^{n-1})$. It should be noted that the WBSV and mWBSV procedures share the same balance conditions. It is only the error in the PV anomaly that distinguishes them.

5. Numerical experiments

The main objective of the numerical results presented below is to give an assessment of the performance of the WBSV procedure against the Rossby-number expansion PV inversion procedures introduced in section 3, particularly the one that leaves q_ℓ unexpanded. Some results using the mWBSV procedure are also presented to assess the importance of the error in the PV anomaly in the WBSV procedure and the accuracy of its balance conditions.

In general, we expect that unsteady vortical flows involve some degree of gravity-wave activity (Ford *et al.* 2000). The accuracy of a set of balance conditions can be measured by its success in disentangling the signature of the freely-propagating gravity waves from that of the PV, i.e. in wave–vortex decomposition. For this purpose, we have probed two non-asymptotic hierarchies of balance conditions for optimal wave–vortex decomposition against which all other balance conditions are compared. The two hierarchies are 'δ balance' ($\partial^{N-1}\delta/\partial t^{N-1} = 0$, $\partial^N \delta/\partial t^N = 0$) and '$\delta$–$\gamma$ balance' ($\partial^N \delta/\partial t^N = 0$, $\partial^N \gamma/\partial t^N = 0$) where $N \geq 1$. Details on these hierarchies can be found in Mohebalhojeh and Dritschel (2001), hereafter MD01. For our purpose, following MN00, we give a

nominal order N to the Nth member of these hierarchies. The orders are nominal because the foregoing hierarchies do not follow a formal asymptotic expansion in Rossby or Froude number. However, these hierarchies are formally valid (for $N \geq 1$) at various orders in two important regimes of flow: the geostrophic scaling employed here and the gradient wind scaling where $Ro = \mathcal{O}(1)$ and $\varepsilon \equiv Fr^2 \ll 1$. This property as well as the possibility of going to arbitrarily high orders make such hierarchies powerful tools to search for optimal balance in a broad range of atmospheric–oceanic flows. Furthermore, numerical evidence shows that they may work well beyond what we may expect from scaling arguments only (MN00).

In what follows, we refer to LPV, VORT, and MASS as the formal asymptotic expansion PV inversion procedures that set, respectively, $q_\ell = q_{\ell 0}$, $\zeta = \zeta_0$, and $h = h_0$. A brief description of the PV inversion procedures tested with the acronyms used is given in Table 1.

TABLE 1. A list of the PV inversion procedures tested and their acronyms and names.

Acronym (name)	Description
WBSV	modified expansion for the PV anomaly and asymptotic truncation of the PV anomaly
mWBSV	modified WBSV at the highest order to recover the exact PV anomaly
LPV	non-expansion of the linearised PV ($q_{\ell i} = 0, i > 0$)
VORT	non-expansion of the vorticity ($\zeta_i = 0, i > 0$)
MASS	non-expansion of the perturbation height ($h_i = 0, i > 0$)
δ–γ balance	N-th order time derivatives of divergence and ageostrophic vorticity set to zero
δ balance	$(N-1)$th and Nth order time derivatives of divergence set to zero

The set-up of the two numerical experiments is the same as the experiment extensively reported in Dritschel *et al.* (1999, hereafter DPM99) from a PV perspective and in Mohebalhojeh and Dritschel (2000, hereafter MD00) from a gravity-wave perspective. The experiments simulate the evolution of a slightly perturbed unstable zonal jet into complex vortical structures involving nonvanishing gravity-wave activity. The main body of the results reported below are obtained with the "contour-advective semi-Lagrangian" (CASL) algorithm described in DPM99, which gives a highly accurate representation of PV. We also present some results from a conventional pseudo-spectral (PS) algorithm, again described in DPM99, to stress the overall insensitivity of the relative performance of the PV inversion procedures to the particular numerical algorithm used.

For our purpose, it suffices to describe the initial PV configuration (see Fig. 1 of DPM99 and MD00). In a doubly-periodic square $[-\pi, \pi] \times [-\pi, \pi]$, the initial PV field is

$$q(x, y, 0) = \bar{q} + Q\mathrm{sgn}(\hat{y})[1 - |\frac{|\hat{y}|}{a} - 1|] \tag{25}$$

for $|\hat{y}(x, y)| < 2a$, and $q = \bar{q}$ otherwise; Q is the amplitude of the PV anomaly, $\bar{q}$ is the mean PV determined by the requirement of zero mean relative vorticity, $2a$ is the distance from the minimum to the maximum potential vorticity, and

$$\hat{y}(x, y) = y + c_m \sin mx + c_n \sin nx \tag{26}$$

is a displaced y coordinate (which preserves the area of differential elements) used to perturb the jet. For the numerical integrations, we have chosen the scalings $H = 1$, $L_R = \sqrt{gH}/f = 0.5$, $a = 0.5$, and $f = 4\pi$, the latter implying that a unit time interval corresponds to a day. The parameters in (26) have been given the values $m = 2$, $n = 3$, $c_2 = -0.1$ and $c_3 = 0.1$. The parameter Q determines the strength of the initial PV anomaly. For the two experiments, we have chosen $Q = (0.25\ ,\ 1.0)f/H$, representing weakly and strongly ageostrophic flows. The resulting PV fields were inverted by means

of the balance conditions $\partial^2\delta/\partial t^2 = \partial^2\delta/\partial t^3 = 0$ (see MN 2000, and MD00) to generate initial states. For the two states, $|Ro|_{\max} = 0.18,\ 0.88$ and $Fr_{\max} = 0.10,\ 0.41$, where $|Ro|_{\max}$ and $Fr_{\max}$ denote the maximum of $|\zeta/f|$ and $[(u^2+v^2)/(g(H+h))]^{1/2}$ over the domain, respectively. The weakly and strongly ageostrophic cases are expected to obey the geostrophic and gradient wind scalings, respectively. Notice that the relevant small parameter for the strongly ageostrophic case is $\varepsilon \equiv Fr^2$, whose maximum value over the domain is 0.17.

The two initial states were integrated forward in time using the CASL and PS algorithms for, respectively, 10 and 40 days. For each flow, the PV inversion procedures mentioned above were carried out at one-day intervals starting from $t = 1$. The spatial resolutions used for both time integration and PV inversion were $n_{\rm g} = 128$ for the CASL algorithm and $n_{\rm g} = 256$ for the PS algorithm, $n_{\rm g}$ being the number of grid points in each direction. A higher resolution was used for PS in order to partially compensate for the adverse effect of diffusion on the representation of PV. The results presented below are from the CASL algorithm, unless stated otherwise.

The performance of the PV inversion procedures is evaluated here by calculating different measures of imbalance (gravity waves). To this end, the PV inversion routines were used to determine the corresponding balanced fields of h, ζ, and δ. By subtracting the balanced fields from their actual SW counterparts the imbalanced fields of h, ζ, and δ were then determined and used to evaluate the state vector $\mathbf{X}_{\rm imb} \equiv (q_\ell\ ,\ \delta\ ,\ \gamma)_{\rm imb}$. Various measures for $\mathbf{X}_{\rm imb}$ are presented, including root-mean-square (rms) values for the fields of $q_{\ell{\rm imb}}$, $\delta_{\rm imb}$, and $\gamma_{\rm imb}$, as well as an ℓ^2 norm, $||X||^2_{\rm imb} = (2\pi/n_{\rm g})^2 \sum_{i,j=1}^{n_{\rm g}} 1/2(|\mathbf{v}_{\rm imb}|^2 + gh^2_{\rm imb})$. The latter norm is simply the quadratic available energy of the imbalanced part of the flow.

To begin with, we compare the performance of WBSV and LPV. Figure 1 shows the foregoing measures against time for the weakly ageostrophic simulation ($HQ/f = 0.25$, $|Ro|_{\max} = 0.18$). Shown are the results for $\mathcal{O}(Ro^0)$ (dashed-dotted), $\mathcal{O}(Ro^1)$ (solid), and $\mathcal{O}(Ro^2)$ (dashed). One should note that for $||X||^2_{\rm imb}$, the $\mathcal{O}(Ro^1)$ and $\mathcal{O}(Ro^2)$ results for LPV as well as those at $\mathcal{O}(Ro^2)$ for WBSV are almost indistinguishable. For rms $\delta_{\rm imb}$, as expected, the two $\mathcal{O}(Ro^0)$ PV inversions agree. So do the four $\mathcal{O}(Ro^1)$ and $\mathcal{O}(Ro^2)$ PV inversions. The discrepancy between the LPV and WBSV results for rms values of $\gamma_{\rm imb}$ at $\mathcal{O}(Ro^0)$ is due to a very small-scale numerical error, which is explained in the appendix. Note that at $\mathcal{O}(Ro^0)$ we expect $\gamma_{\rm imb} = \gamma$. The same numerical error affects to a less extent the results for rms $\gamma_{\rm imb}$ at $\mathcal{O}(Ro^1)$, such that the actual values for the two procedures are closer than what Fig. 1 suggests. The effect of the small-scale numerical error on the other results in Fig. 1 and reported below was found to be insignificant. For rms $\gamma_{\rm imb}$, LPV at $\mathcal{O}(Ro^2)$ duplicates its $\mathcal{O}(Ro^1)$ counterpart. The immediate conclusion is that for this low Rossby-number flow, there is little to gain by going from $\mathcal{O}(Ro^1)$ to $\mathcal{O}(Ro^2)$ accurate balance conditions. In terms of the rms values of $q_{\ell{\rm imb}}$ and $||X||^2_{\rm imb}$, there is a clear distinction between the LPV and WBSV procedures of the same formal accuracy. In every case, the WBSV procedure exhibits less success than the LPV in capturing the existing balance, or equivalently in capturing the waves.

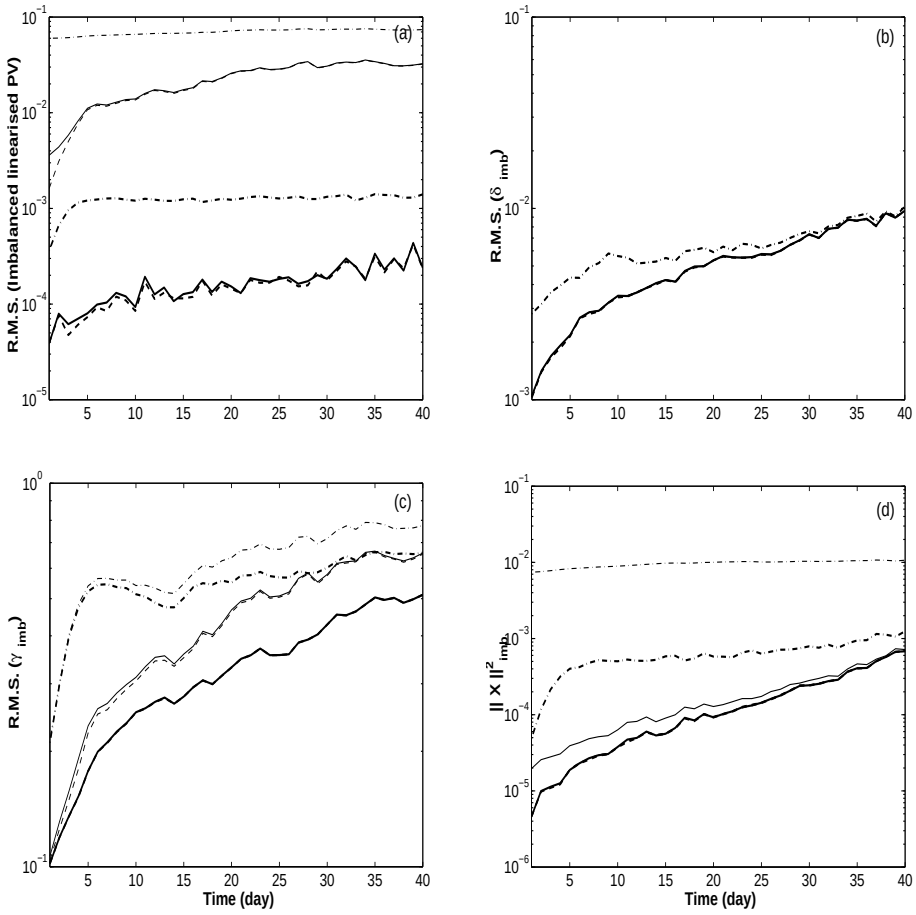

Figure 1. Different measures of imbalance for the weakly ageostrophic simulation. Plotted are root-mean-square fields of (a) $q_{\ell\mathrm{imb}}$, (b) δ_{imb}, (c) γ_{imb}, and (d) $||X||^2_{\mathrm{imb}}$. The heavy lines present the results for LPV, i.e. PV inversion by means of the Rossby-number expansion balance conditions that set $q_\ell = q_{\ell 0}$ (see section 3). The thin lines present the results for the WBSV procedure (see section 4). The dashed-dotted, solid, and dashed lines present the results formally accurate to $\mathcal{O}(Ro^0)$, $\mathcal{O}(Ro^1)$, and $\mathcal{O}(Ro^2)$, respectively.

Figure 2 shows the same measures of imbalance against time for the strongly ageostrophic simulation ($HQ/f = 1.0$, $|Ro|_{\max} = 0.88$) again for WBSV and LPV only. Except for the fact that the two $\mathcal{O}(Ro^0)$ PV inversions agree in their rms values of $\delta_{\rm imb}$ and $\gamma_{\rm imb}$, the other measures provide clear division between different orders of the asymptotic balance conditions on the one hand and the two PV inversion algorithms on the other. The division is particularly sharp for $q_{\ell{\rm imb}}$ and $||X||^2_{\rm imb}$, where $\mathcal{O}(Ro^2)$ strongly outperforms $\mathcal{O}(Ro^1)$. Like in the low-Rossby-number flow, the WBSV procedure leaves a noticeably larger contribution to imbalance, judging from the results for $||X||^2_{\rm imb}$ and the rms values of $q_{\ell{\rm imb}}$, $\gamma_{\rm imb}$. The picture emerging from the divergence field is not unequivocal. Whereas for the $\mathcal{O}(Ro^1)$ balance conditions the WBSV procedure does slightly better, the results for the $\mathcal{O}(Ro^2)$ balance conditions are again in favour of LPV. Overall, the results shown support the assertion that the WBSV procedure is not as accurate as the LPV.

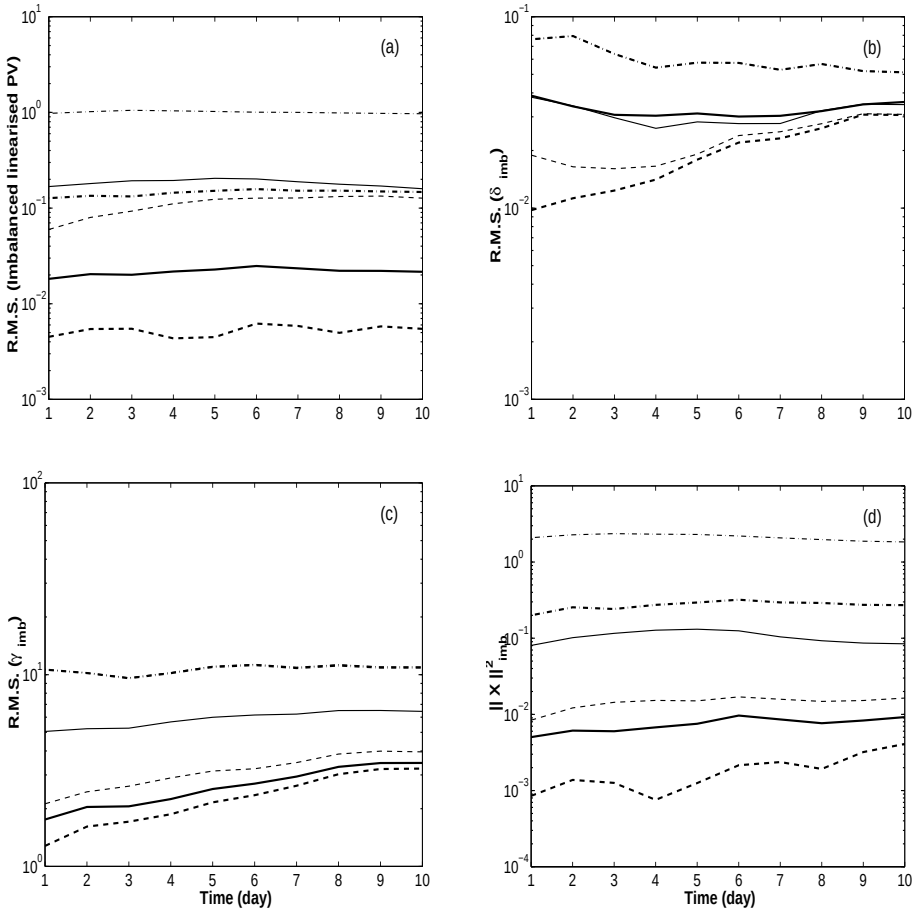

Figure 2. As in Fig. 1, but for the strongly ageostrophic simulation.

To give a composite picture of the performance of the PV inversions, we present in Fig. 3 the time-averaged values of $||X||^2_{\rm imb}$ against order of the balance conditions for (a,c) the weakly and (b,d) the strongly ageostrophic simulations by (a,b) the CASL and (c,d) the PS algorithms. Also shown are the results from PV inversion by means of δ–γ balances up to the highest convergent order. The results due to δ balance (not shown for

clarity) for orders greater than one follow closely those due to δ–γ balance. Regarding the other results absent in Fig. 3, the following considerations should be noted. First, MASS did not converge for the high-Rossby-number case. Second, the results for VORT and MASS for the low-Rossby-number case using the CASL algorithm are indistinguishable from the results for both LPV and δ–γ, and therefore were not included for the sake of clarity.

In essence, Fig. 3 leads us to the key findings of this study. Consistent with the conclusions reached from Figs 1 and 2, the WBSV procedure (thin solid lines) provides a quantitative measure for imbalance which is far from optimal, particularly so for moderate Rossby numbers. The optimal values can be considered the minimal values achieved by δ–γ balance, e.g. the second-order δ-γ balance in Fig. 3a. Note furthermore that the mWBSV procedure (circles) improves the performance of WBSV only slightly. Therefore, the poor performance of WBSV has more to do with its construction, i.e. the fact that the whole PV anomaly makes up the zeroth-order geostrophic solution, which results in inaccurate balance conditions like (21) at $\mathcal{O}(Ro^1)$. Among the Rossby-number expansion PV inversions, at $\mathcal{O}(Ro^1)$ the best results come from VORT, i.e. the procedure which is exact for steady axisymmetric vortices. But at $\mathcal{O}(Ro^2)$ it is LPV that is nearest to optimal. Notice the order of magnitude less imbalance for LPV compared with WBSV for the higher-Rossby number case. The last important point is that the results due to the PS algorithm are in general agreement with those due to the CASL one. The PS results alter none of the findings mentioned above despite the quantitative differences with the CASL results, especially at low Rossby numbers.

Such quantitative differences can be understood in the light of the results on numerical issues reported in MD00, a brief discussion of which is given here. The CASL algorithm results in the erroneous production of gravity waves as a result of mishandling the PV, which is presented down to the grid-scale. The effect is more manifest at the low Rossby, low Froude number regime of flow. The PS algorithm suppresses the gravity waves by removing the fine-scale features of PV as a result of diffusing vorticity. Therefore, the quantitative measures of imbalance presented in Fig. 3 should be considered as overestimates, for the CASL, and underestimates, for the PS, of imbalance. It is possible to improve the representation of gravity waves, without sacrificing the representation of PV, by designing CASL algorithms based on PV inversion hierarchies like the 'δ balance'. The best example of such CASL algorithms is the algorithm called CA_1 in MD00 whose prognostic variables are $(q\ ,\ \delta\ ,\ \delta_t)$. For the comparison of the relative performance of the PV inversion procedures, however, we expect the same overall insensitivity. For our purpose, it is not so important that the gravity waves generated are physical. What matters here is the extent to which various PV inversion procedures can capture the signature of balanced vortical flow.

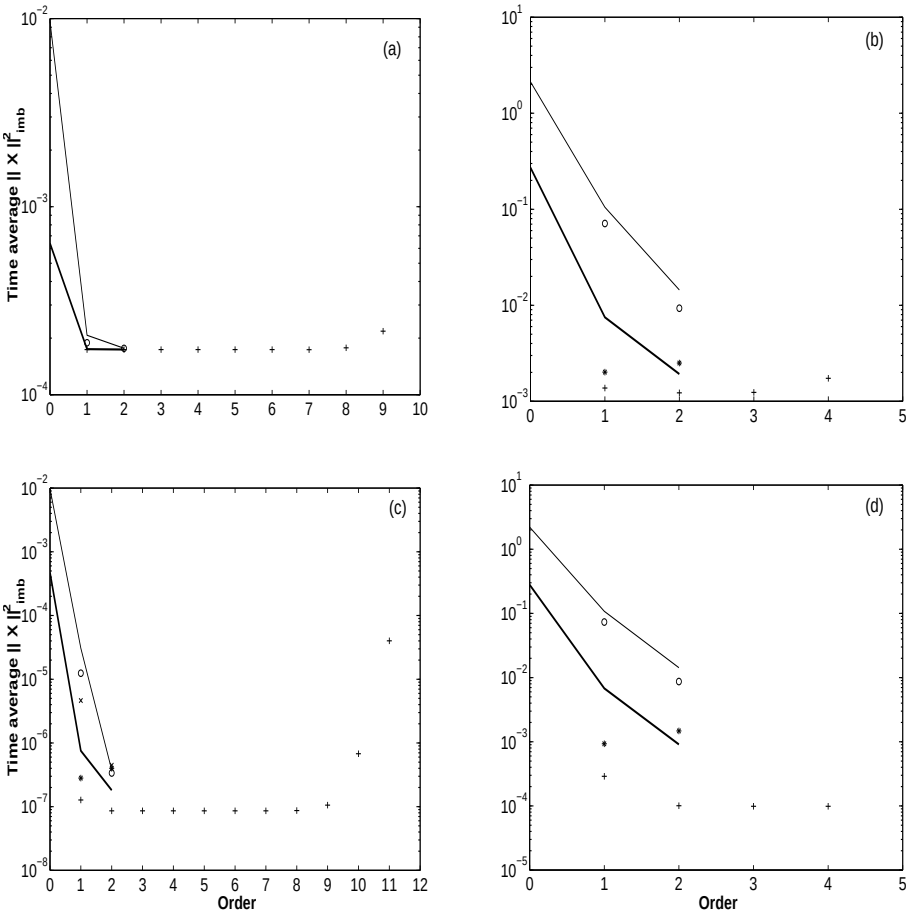

Figure 3. Time-averaged values of $||X||^2_{\rm imb}$ for (a,c) the weakly and (b,d) the strongly ageostrophic flows against the order of expansion. The results shown are for (a,b) the CASL algorithm and (c,d) the pseudo-spectral algorithm. The heavy line, $*$, and $\times$ denote the results for PV inversion by means of the Rossby-number expansion balance conditions that set $q_\ell = q_{\ell 0}$ (LPV), $\zeta = \zeta_0$ (VORT), and $h = h_0$ (MASS), respectively (see section 3). The thin line and ($\circ$) denote the results for the WBSV and mWBSV procedures, respectively (see section 4). Also shown, (+) are the results for PV inversion by means of δ–γ balance (see 5). For clarity, in (a) the symbols $*$ and $\times$ have not been plotted.

Finally, we provide qualitative information in support of the quantitative results presented. The objective is to show that more accurate PV inversions give us imbalanced fields that are more representative of freely-propagating gravity waves. In other words, the equivalence of less imbalance with higher accuracy assumed so far, e.g. in interpreting Fig. 3, is further established. Figure 4 shows plots of (a) actual height h and imbalanced height h_{imb} for the second-order (b) WBSV, (c) LPV, and (d) δ–γ balance. The plots are for the the highly ageostrophic CASL simulation at $t = 10$. Notice the different contour intervals used. The imbalanced height field due to WBSV is mainly contaminated with the signature of the balanced vortical flow. There are clear indications that by going from WBSV to LPV, and then to δ–γ balance, we recover an imbalanced height field ever more representative of freely-propagating gravity waves. In this regard, notice how h_{imb} due to LPV exhibits convergence to h_{imb} due to the second-order δ–γ balance, which is our optimal balance here (see Fig. 3a). It should also be noted that the plot of h_{imb} due to mWBSV (not shown) is only marginally different from that of WBSV.

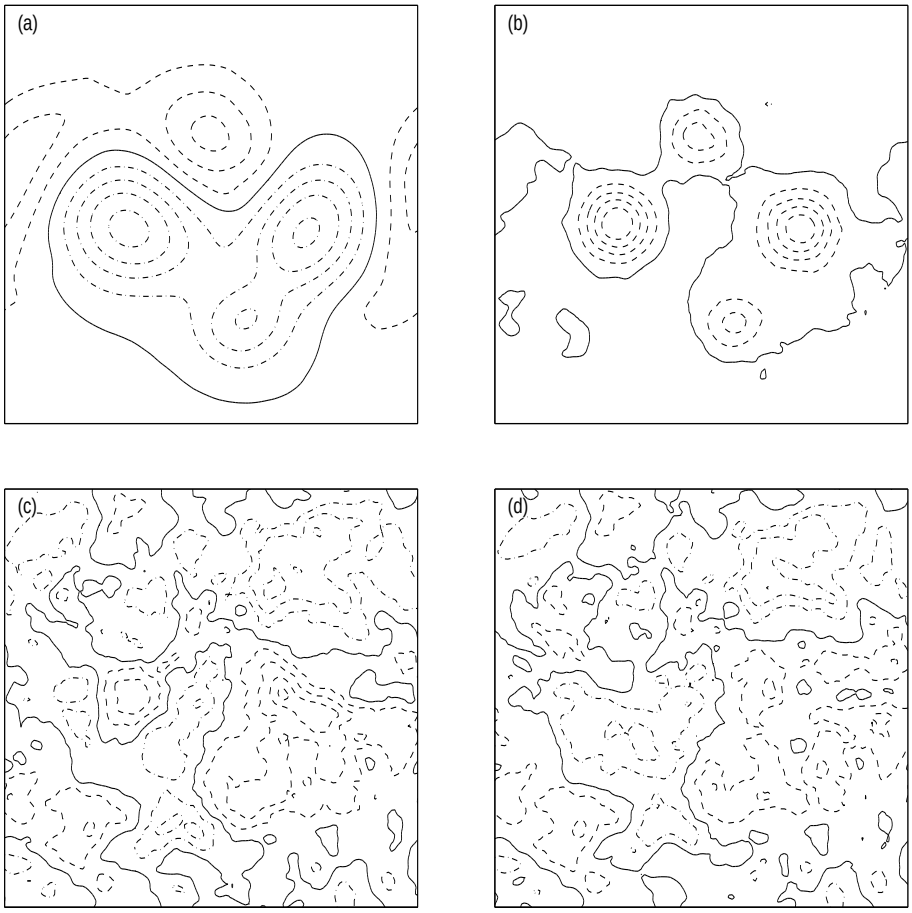

Figure 4. Plotted are (a) the actual perturbation height at $t = 10$ using a contour interval of 0.1 and the corresponding fields of imbalanced perturbation height given by the second-order PV inversions of (b) WBSV, (c) LPV, and (d) δ–γ balance. The contour intervals used in (b), (c), and (d) are respectively 0.005, 0.001, and 0.001. The plots are all for the CASL simulation with $HQ/f = 1.0$.

6. conclusion

The inherent indeterminacy of conventional Rossby-number expansions gives us, in principle, infinite freedom to construct balance conditions beyond leading-order geostrophy. For example, in the expansion we can set $q_\ell = q_{\ell 0}$ and obtain the modified expansion for the linearised PV. The resulting balance conditions can then be used to invert the PV anomaly. Another example is provided by the WBSV procedure, which is a modified expansion for the PV anomaly q' (not to be confused with q_ℓ). This procedure is unique among both asymptotic and non-asymptotic PV inversion procedures in the sense that at every order all of the inversion equations are linear. The linearity comes from (i) the non-expansion of the PV anomaly and thus the resulting balance conditions and (ii) the asymptotic truncation of the definition of PV anomaly (eqn. 16). The asymptotic truncation, however, results in an error in the PV. More precisely, the inverted fields do not satisfy the Rossby formula for PV anomaly exactly. It is possible to abandon the formal asymptotics at the highest order so that the exact Rossby formula for the PV anomaly is recovered. The resulting procedure, called mWBSV, helps us determine the relative significance of the balance conditions involved against the asymptotic truncation up to the highest order. It is worth emphasizing that the WBSV and mWBSV procedures of the same order share the same balance conditions. They are distinguished only by the error in PV. Another consequence of modifying WBSV to mWBSV is a weak nonlinearity of the highest-order equations.

Detailed numerical results for various measures of imbalance establish the fact that the WBSV procedure achieves linearity at the expense of accuracy. Nonlinearity is inescapable so far as accuracy is concerned. Further, the asymptotic truncation of the PV anomaly has a relatively small contribution to the error in the WBSV procedure. The main problem with this procedure seems to be a poor leading-order solution. Higher-order corrections fail to raise the accuracy of the solution to the level of other asymptotic procedures, as compared with the optimal one. Of course, the level of accuracy required strongly depends on the specific application of PV inversion and thus cannot be determined objectively. It is true that for low-Rossby-number flows the distinction between various measures of imbalance at $\mathcal{O}(Ro^1)$ and particularly at $\mathcal{O}(Ro^2)$ can be insignificant. But for moderate Rossby numbers, the inadequacy of the WBSV procedure even at $\mathcal{O}(Ro^2)$ can be a drawback. As well as the complexity of the expansion equations beyond $\mathcal{O}(Ro^2)$, for moderate Rossby numbers the divergent nature of the expansion may start to show up at $\mathcal{O}(Ro^3)$ as the results presented in MD01 indicate. In other words, by going from $\mathcal{O}(Ro^2)$ to $\mathcal{O}(Ro^3)$ the results may start to deteriorate. Therefore, it seems very unlikely that the WBSV procedure at $\mathcal{O}(Ro^3)$ can compete in accuracy with LPV at $\mathcal{O}(Ro^2)$, for instance.

Acknowledgement

The author wishes to thank O. Bühler, D. G. Dritschel, and M. E. McIntyre for helpful discussions and suggestions. The modification of the WBSV procedure satisfying the exact SW equation for PV anomaly was suggested by an anonymous referee to whom I am thankful.

Appendix

The numerical error that causes the discrepancy between the rms values of γ_{imb} at $\mathcal{O}(Ro^0)$ (Fig. 1) is explained here. The error is related to some subtle details of

the numerical PV inversion algorithms, the basic design of which is the same as that described in MD00 for the δ hierarchy. In the numerical algorithms for WBSV (and mWBSV) at various orders, after evaluating h_0 the zeroth-order vorticity field can be computed either (i) from the geostrophic formula, as carried out here, or (ii) from (18). Although the two computations are equivalent in the continuous case, the differences in the numerical errors for ζ_0 and h_0 make the two computations differ slightly at very small scales. It is the first computation that ensures $\gamma_0 = 0$. On the other hand, in the other numerical PV inversion algorithms including that for LPV, the vorticity field is always computed according to the Rossby formula for PV anomaly. The differences in the numerical errors for ζ_0 and h_0 make the LPV vorticity at zeroth-oder different from its geostrophic value. This discrepancy is marked only at very small scales. It occurs in the following way. The modified Helmholtz equations for height and divergence fields are solved in the spectral space. In order to remove the unresolvable waves smaller than the two-grid wave, a circular filter is applied to the inverse modified Helmholtz operator $\mathcal{H}^{-1}$ in spectral space. By removing the circular filter, for the low-Rossby-number case the $\gamma_{\rm imb}$ results due to the LPV and WBSV procedures would match at zeroth-order and be closer at the first-order. No noticeable effect of this kind, however, was found on the other numerical results reported in this paper.

References

Bokhove, O. 1997 Slaving principles, balanced dynamics, and the hydrostatic Boussinesq equations. *J. Atmos. Sci.*, **54,** 1662–1674

Dritschel, D. G., Polvani, L. M., and Mohebalhojeh, A. R. 1999 The contour-advective semi-Lagrangian algorithm for the shallow water equations. *Mon. Wea. Rev.*, **127,** 1551–1565

Ford, R., McIntyre, M. E., and Norton, W., A. 2000 Balance and slow quasi-manifold: some explicit results. *J. Atmos. Sci.*, **57,** 1236–1254

Hoskins, B., J., Draghici, I., and Davies, H., C. 1978 A new look at the ω-equation. *Q. J. R. Meteorol. Soc.*, **104,** 31–38

Hoskins, B. J., McIntyre, M. E., and Robertson, A. W. 1985 On the use and significance of isentropic potential-vorticity maps. *Q.J.R.Meteorol.Soc.*, **111,** 877–946

Leith, C. N. 1980 Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.*, **37,** 958–968

Lynch, P. 1989 The slow equations. *Q.J.R.Meteorol.Soc.*, **115,** 201–219

Machenhauer, B. 1977 On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitr. Phys. Atmos.*, **50,** 253–271

McIntyre, M. E., and Norton, W. A. 2000 Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, **57,** 1214–1235

Mohebalhojeh, A. R., and Dritschel, D. G. 2000 On the representation of gravity waves in numerical models of the shallow-water equations. *Q.J.R.Meteorol.Soc.*, **126,** 669–688

Mohebalhojeh, A. R., and Dritschel, D. G. 2001 Hierarchies of balance conditions for the f-plane shallow water equations. *J. Atmos. Sci.*, **58,** 2411–2426

Muraki, D. J., Snyder, C., and Rotunno, R. 1999 The next-order corrections to quasigeostrophic theory. *J. Atmos. Sci.*, **56,** 1547–1560

Pedlosky, J. 1987 *Geophysical Fluid Dynamics*, 2nd ed. Springer-Verlag, 710 pp.

Phillips, N. A. 1960 On the problem of initial data for the primitive equations. *Tellus*, **12,** 121–126

Rotunno, R., Muraki, D. J., and Snyder, C. 2000 Unstable baroclinic waves beyond quasigeostropic theory. *J. Atmos. Sci.*, **57,** 3285–3295

Vallis, G. K. 1996 Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Q. J. R. Meteorol. Soc.*, **122,** 291–322

Warn, T., O, Bokhove, O., Shepherd, T., G., and Vallis, G., K. 1995 Rossby number expansions, slaving principles, and balance dynamics. *Q. J. R. Meteorol. Soc.*, **121,** 723–739