J. Fluid Mech. (2007), *vol.* 570, *pp.* 365–383.
doi:10.1017/S0022112006002990 Printed in the United Kingdom

The persistence of balance in geophysical flows

By DAVID G. DRITSCHEL[1] AND ÁLVARO VIÚDEZ[2]

[1]School of Mathematics and Statistics, University of St Andrews, St Andrews, UK
dgd@mcs.st-and.ac.uk

[2]Institut de Ciènces del Mar, CSIC, Barcelona, Spain
aviudez@cmima.csic.es

(Received 13 May 2005 and in revised form 16 June 2006)

Rotating stably stratified geophysical flows can exhibit a near 'balanced' evolution controlled by the conservative advection of a single scalar quantity, the potential vorticity (PV). This occurs frequently in the Earth's atmosphere and oceans where motions tend to be weak compared with the background planetary rotation and where stratification greatly inhibits vertical motion. Under these circumstances, both high-frequency acoustic waves and lower-frequency inertia–gravity waves (IGWs) contribute little to the flow evolution compared with the even-lower-frequency advection of PV. Moreover, this 'slow' PV-controlled balanced evolution appears unable to excite these higher-frequency waves in any significant way – i.e. balance persists.

The present work pushes the limits of balance by systematically exploring the evolution of a range of highly nonlinear flows in which motions are comparable with the background rotation. These flows do not possess a frequency separation between PV advection and IGWs. Nonetheless, the flows exhibit a remarkable persistence of balance. Even when flows are not initialized to minimize the amount of IGWs initially present, and indeed even when flows are deliberately seeded with significant IGW amplitudes, the flow evolution – over many inertial periods (days) – remains strongly controlled by PV advection.

1. Introduction

Intermediate-to-large-scale fluid motions within the atmosphere and oceans are strongly influenced by both the planetary rotation and stable density stratification. Cyclones and other features tend to rotate slowly compared with the planetary rotation (the Rossby number is small) and have vertical shears which are weak compared with the buoyancy frequency (the Froude number is small). Exceptions occur at smaller scales, in convective regions, and in the upper atmosphere, but these tend to be localized. Otherwise, rotation and stratification combine to render fluid motions 'layerwise two-dimensional', with vertical motions often substantially weaker than horizontal motions (cf. Viúdez & Dritschel 2003, hereafter referred to as VD03, and references therein). Moreover, these motions are, often to astonishing accuracy (McIntyre & Norton 2000), governed by a single nearly conservative scalar field, the potential vorticity (PV). The generally higher-frequency inertia–gravity waves (IGWs) and acoustic waves play only a minor role. This PV-controlled motion is known as balance (Hoskins, McIntyre & Robertson 1985; Ford, McIntyre & Norton 2000 and references therein).

Balance is only a hypothetical state of motion – departures from balance, the IGWs or 'imbalance', appear to be inevitable, implying an inherent ambiguity in the concept

of balance (Ford *et al.* 2000). Yet these departures can be surprisingly weak, much weaker than a simple scale analysis would indicate, justifying balance even in nonlinear flow regimes, where PV and IGWs evolve on similar time scales (cf. Vanneste & Yavneh 2004; Vanneste 2004; Viúdez & Dritschel 2004). This has often been exploited, practically in meteorology (in data assimilation and forecast initialization), and theoretically in modelling fundamental fluid behaviour via reduced balanced models (cf. Vallis 1996).

The advantage of balanced models is that they filter the higher-frequency IGWs from the dynamics, reducing it essentially to PV advection. This filtering is normally accomplished by suppressing a pair of time derivatives in the original equations of motion. This transforms two 'prognostic' hyperbolic equations into a pair of 'diagnostic' elliptic equations. These equations, which are often nonlinear, together with the definition of PV, may then be inverted to find the original flow variables (velocity, density, etc.). There is considerable freedom in this procedure: the original equations can be time-differentiated multiple times and rewritten as prognostic equations for a pair of new (time-differentiated) variables which might appear more wavelike. Filtering these equations can be more effective in eliminating IGWs or, equivalently, in defining an accurate balanced evolution (Mohebalhojeh & Dritschel 2000, 2001).

The ambiguity in defining balance is inherent in the nonlinear equations governing the flow. No one balanced model is suitable for all regimes of flow. No one set of diagnostic relationships is suitable for defining balance. And without a definition of balance, it is impossible to quantify departures from balance, i.e. the existence of IGWs, in the full IGW-permitting equations of motion.

To overcome this obstacle, a new diagnostic procedure for defining balance was developed recently (Viúdez & Dritschel 2004). This procedure, called 'optimal PV balance', involves integrating the full equations of motion (recast to make PV conservation explicit) backwards and forwards in an iterative cycle designed to minimize the generation of IGWs. The balance so obtained depends only on the PV distribution at the diagnostic time and on the period of integration Δ_τ in the cycle. In practice, results prove insensitive to Δ_τ for Δ_τ values exceeding about three inertial periods $T_{\rm ip} = 2\pi/f$, where f is the Coriolis parameter (twice the background rotation). We may thus argue that the ambiguity has thereby been minimized, in a natural sense.

The goal here is to use this new procedure to quantify the importance of IGWs, in a three-dimensional non-hydrostatic isochoric flow, across a wide range of geophysically relevant flow parameters (Rossby number, Froude number, and scale ratio). To this end, we identify the characteristics of the PV distribution which trigger the emission of IGWs. Furthermore, we examine the accumulative impact of IGWs on the balanced flow, by varying the initial degree of imbalance for a fixed initial distribution of PV.

The paper is organized as follows. In the next section, we recall the governing fluid-dynamical equations, recast in a novel way to make PV conservation explicit. This is convenient both for the optimal PV (OPV) balance procedure and for the novel numerical method employed, as indicated in § 3. A simple illustration of OPV balance is provided in § 4 to clarify how it may be used to minimize IGW generation, thereby providing an accurate estimate of the underlying balance. In § 5, we turn to a class of complex flows involving the breakup of an unstable jet into cyclonic and anticyclonic eddies. These flows are investigated over a wide range of parameters (spanned by nearly 100 simulations), in order to better understand the role of IGWs and the persistence, or not, of balance. Flows are initialized either with minimal IGWs or with moderate-amplitude IGWs, and comparisons are made to assess the impact of IGWs on the underlying balance. The principal findings are reviewed in § 6.

2. Mathematical preliminaries

This study makes use of the simplest three-dimensional model of a rotating stratified fluid. We consider the isochoric (non-divergent) flow of an incompressible inviscid fluid. Thus, the density ρ is materially conserved. Moreover, density variations are considered small compared with a background linear reference profile $\bar{\rho}(z)$, to allow the Boussinesq approximation to be made. This reference profile defines the mean buoyancy frequency N. The fluid motion is viewed relative to a rotating reference frame, which rotates at the rate $\Omega = f/2$, where f is the Coriolis frequency. Both f and N are taken to be spatially uniform. This model is appropriate to the oceanic mesoscale away from boundaries and the free surface.

Under these assumptions, it is customary to write the governing equations in terms of the velocity $\boldsymbol{u} = (u, v, w)$ and the buoyancy $b = -g(\rho - \bar{\rho})/\rho_0$, where ρ_0 is a constant reference density (note that $N^2 = -g\bar{\rho}_z/\rho_0$). However, in this form $\nabla \cdot \boldsymbol{u} = 0$ is not explicitly enforced; an additional equation is needed to solve for the pressure. But perhaps more fundamentally, the conservation of potential vorticity Π for fluid particles is not explicit. The potential vorticity (PV), a scalar field, is proportional to the product of the absolute vorticity $\boldsymbol{\omega} + f\boldsymbol{k}$ (here the horizontal planetary vorticity is neglected) and the gradient of the density (or more generally, the thermodynamic entropy; see Hoskins *et al.* 1985). It may be defined in a dimensionless way by

$$\Pi = (\boldsymbol{k} + \boldsymbol{\omega}/f) \cdot (\boldsymbol{k} + \nabla b/N^2) \tag{2.1}$$

(cf. Dritschel & Viúdez 2003, hereafter DV03), and in an adiabatic inviscid flow it is conserved following fluid particles:

$$\dot{\Pi} \equiv \Pi_t + \boldsymbol{u} \cdot \nabla \Pi = 0 \tag{2.2}$$

(the subscripts t, x, y and z denote partial differentiation). Despite its importance in rotating stably stratified flows (cf. Ford *et al.* 2000), PV has not been used in the numerical simulation of these equations until recently (DV03, VD03). This, it would appear, is due to the nonlinear dependence of PV on the original flow variables $\boldsymbol{u}$ and b. Using PV requires one to pick two other independent variables and allow them to evolve and then to invert the definitions of PV and these two other variables to obtain $\boldsymbol{u}$ and b. Typically, this involves the inversion of nonlinear elliptic equations.

In DV03, the choice of these two other variables was motivated by anticipating that typical flows remain close to hydrostatic and geostrophic balance. Hydrostatic balance results from neglecting the vertical acceleration in the momentum equations, while geostrophic balance results from ignoring the horizontal acceleration (in the rotating frame of reference). When combined to eliminate the pressure, these two balances relate the vertical shear of the horizontal wind to the gradient of the buoyancy; they are known as the 'thermal wind' relations. It was argued that if the PV controls the balanced motions, the other variables should express the unbalanced motions, at least to leading order. To this end, the other variables were chosen to be the horizontal components of the dimensionless vector

$$\mathscr{A} \equiv \boldsymbol{\omega}/f + \nabla b/f^2. \tag{2.3}$$

The subscript h denoting 'horizontal part', the statement $\mathscr{A}_h = \boldsymbol{0}$ expresses hydrostatic–geostrophic balance to a high degree of accuracy (u_z and v_z dominate w_x and w_y in $\boldsymbol{\omega}_h$). This choice has many convenient mathematical properties too. As a vector expression, it is coordinate independent. Moreover, if we define a vector potential by $\mathscr{A} = \nabla^2\boldsymbol{\varphi}$, the original variables may be recovered from $\boldsymbol{u} = -f\nabla \times \boldsymbol{\varphi}$ and $b = -f^2\nabla \cdot \boldsymbol{\varphi}$. Note

that $\nabla \cdot \boldsymbol{u} = 0$ is explicit. (A related transformation was introduced by Muraki, Snyder & Rotunno (1999) for hydrostatic flows.)

The evolution equations for $\mathscr{A}_h$ are given in DV03 and are not repeated here. These equations involve $\boldsymbol{u}$ and b, which are found by inverting the definitions of $\mathscr{A}_h$ and Π each time. The horizontal part of the vector potential $\boldsymbol{\varphi}_h$ can be found simply by operating the inverse of Laplace's operator on $\mathscr{A}_h$. The vertical part ϕ, however, requires solving a double Monge–Ampère equation, a quadratically nonlinear equation which is normally elliptic but may in extreme cases be hyperbolic over parts of the flow (details are given in DV03). This is the price to pay for using PV as a prognostic variable, but in practice it is not costly. The explicit conservation of PV has major advantages, both for numerical accuracy and for defining the PV-controlled balanced flow, as seen below.

3. Outline of the numerical algorithm and optimal PV balance

3.1. *Numerical algorithm*

Full details of the numerical algorithm are given in the appendix of DV03, so only a brief review of the essential elements is provided here.

The algorithm integrates the equations for $\mathscr{A}_h$ and Π in an idealized triply periodic box of dimensions $L_X \times L_Y \times L_Z$. Without loss of generality, we take $L_Z = 2\pi$ and take $L_X = L_Y = 2\pi N/f$ (typically the frequency ratio N/f is about 10 to 100 in the ocean). Equal numbers of grid points are used in each direction, here 64 except where noted. This anisotropic geometry, a flat slab, is necessary to model accurately PV-controlled flows, which often exhibit horizontal to vertical scale ratios of $O(N/f)$ (see VD03; Reinaud *et al.* (2003), and refs.).

The algorithm exploits the periodicity by making use of a spectral representation of all fields to compute spatial derivatives and to invert operators etc. Products of fields are carried out on the grid after performing fast Fourier transforms. This approach is standard. Less standard is the contour representation used to ensure material conservation of PV to high accuracy. Here, the contour-advective semi-lagrangian (CASL) algorithm is employed (Dritschel & Ambaum 1997). The potential vorticity is represented by contours lying on material isopycnal (constant-density) surfaces. In the present study, 64 surfaces were used (the number of isopycnals is however independent of the number of vertical grid points). The contours represent jumps in Π, so that the PV is piecewise uniform. Each contour is represented by an ordered set of 'nodes' whose number may vary as the contour deforms and stretches. These nodes simply move with the local fluid velocity to ensure PV conservation. But one cannot follow contour stretching indefinitely, so exact conservation must be relaxed at some level. This is done in CASL by 'surgery', which limits the width of PV filaments to one-twentieth of the horizontal grid size. Additionally, a weak biharmonic hyperdiffusion is added to the $\mathscr{A}_h$ tendencies, to control the build-up of grid-scale noise. Except where noted, the damping (e-folding) rate on the highest wavenumber (here 32) is 1 per inertial period, as in previous studies. Overall, the total numerical dissipation is substantially weaker than that required in traditional grid-based approaches for numerical stability, as the comparison in DV03 illustrates. An accurate representation of the PV field, we argue, is a prerequisite in any proper assessment of IGW emission.

The inversion of the nonlinear Monge–Ampère equation for the vertical component ϕ of the vector potential $\boldsymbol{\varphi}$ requires gridded PV values. These are obtained by a fast interpolation algorithm described in DV03. The Monge–Ampère equation is solved iteratively, by inverting the linear operator part of the equation on the remaining terms,

found from the previous iterate for ϕ. This procedure converges exponentially fast, and normally only a few iterations are required.

Time integration is carried out using a third-order Adams–Bashforth method for the contours and a leap-frog method for $\mathscr{A}_h$. The time step is set to a tenth of the buoyancy period $2\pi/N$ for high accuracy.

3.2. *Optimal PV balance*

The main purpose of this paper is to quantify the degree to which the PV controls the evolution of rotating stably stratified flows. Do the inertia–gravity waves matter?

The full dynamics contains a mix of PV-controlled balanced motions and inertia–gravity waves (IGWs), or unbalanced motions. Separating the two has not proved to be easy, and indeed it may be impossible to do so precisely. For example, one might decide to invoke hydrostatic–geostrophic (or 'quasi-geostrophic') balance to define the PV-controlled flow, but then one would find that a part of the residual unbalanced motions is closely tied to the PV evolution and hence is actually part of the balanced motions. The difficulty is making this observation precise. The best we can do is to minimise the imbalance in any definition of balance.

This is the idea behind a new procedure for defining balance, called the 'optimal PV (OPV) balance' (Viúdez & Dritschel 2004, hereafter VD04). This procedure does not define balance by imposing specific balance relations (e.g. thermal-wind relations), but rather by integrating the full equations of motion backwards and forwards over a non-physical time τ in a way which reduces the imbalance after each integration cycle. During each integration, the PV anomaly $\varpi \equiv \Pi - 1$ on each fluid particle is multiplied by a ramp function $T(\tau)$ which varies smoothly from 0 to 1 from the beginning ($\tau = 0$) to the end ($\tau = \Delta_\tau$) of the ramp period. The fluid particles are allowed to move freely, reaching a new configuration at the end of each backwards integration (from $\tau = \Delta_\tau$ to 0). Any residual motion must be due to IGWs and is removed. The integration then restarts, now forwards, and the fluid particles again move freely, ending up close to the actual configuration (in the diagnosed flow) at the end of the integration. Any difference is then removed by resetting the PV distribution to the actual one, and the next integration cycle begins. This is repeated until the difference in the fields obtained at the end of successive cycles is less than some tolerance. The resulting fields define the balance – they have minimal IGWs.

This procedure has one adjustable parameter, namely the period of the integration Δ_τ. In practice, even for rapidly evolving flows (VD04) the balance is not sensitive to the choice of this parameter, for Δ_τ values greater than about three inertial periods ($T_{\rm ip} = 2\pi/f$). Additional results are provided in the following section, for a simple example.

4. A simple test problem

In this section, we consider the generation of IGWs during the growth and decay of a spheroidal vortex. This idealized flow is created by ramping up the PV anomaly ϖ on each fluid particle $\boldsymbol{X}$, in a flow initially at rest, over various periods T_I ranging from $T_I = 0.05T_{\rm ip}$ (abrupt initialization) to $T_I = 5.05T_{\rm ip}$. Specifically, we set $\varpi(\boldsymbol{X}, t) = W(t/T_I)\varpi_0(\boldsymbol{X})$, where $\varpi_0(\boldsymbol{X})$ is the specified material PV anomaly at $t = T_I$, see below, and $W(s) = \frac{1}{2}(1 - \cos \pi s)$ is the ramp function, identical to that used in previous works (DV03; VD03) and in OPV balance (VD04). The vortical flow thus created is then spun down again by reversing the ramp (i.e. using the ramp function until $t = 2T_I$). After $t = 2T_I$, the flow is left to evolve freely without any PV anomaly

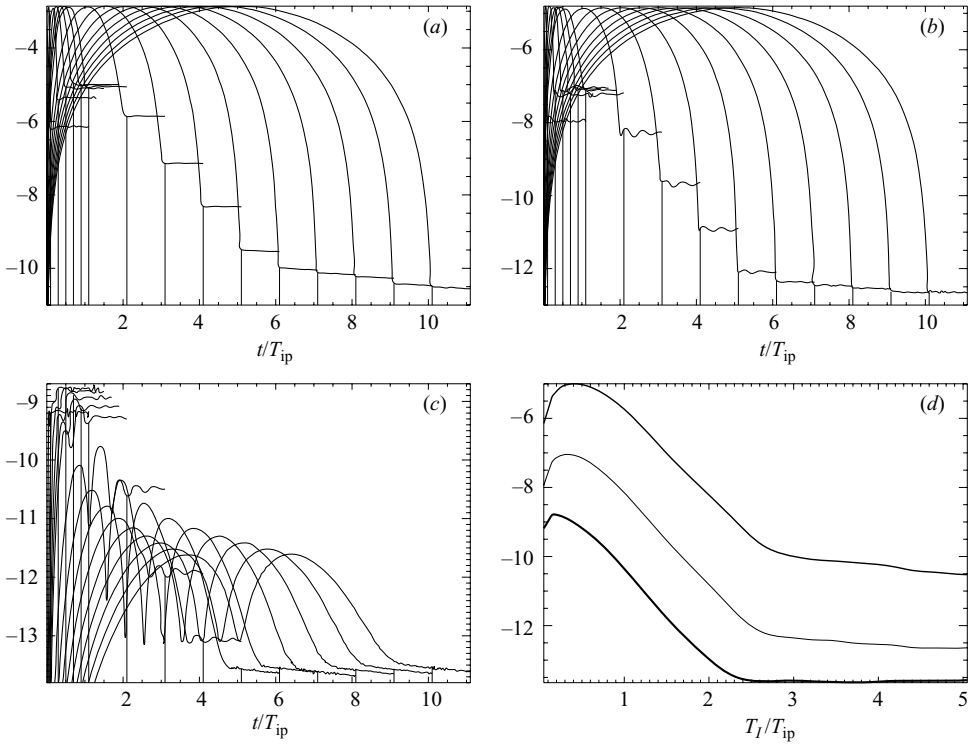

FIGURE 1. Time evolution of (a) mean-square kinetic energy $\langle(u^2+v^2)/2\rangle$, (b) mean-square isopycnal displacement $\langle\mathscr{D}^2\rangle$, and (c) mean-square vertical velocity $\langle w^2\rangle$, for the ramped-spherical-vortex simulations. All values have been log scaled. Curves are shown for selected ramp periods such that $T_I/T_{\rm ip}=0.05$, 0.15, 0.25, 0.35, 0.45, 0.55, 1.05, 1.55, 2.05, 2.55, 3.05, 3.55, 4.05, 4.55 and 5.05. There is no PV anomaly after $t=2T_I$, indicated by the thin vertical lines for each case. (d) The time averages of $\langle(u^2+v^2)/2\rangle$ (top), $\langle\mathscr{D}^2\rangle$ (middle) and $\langle w^2\rangle$ (bottom), over the final inertial period when only IGWs are present, vs. the ramp period T_I.

for a further inertial period. Any residual motions left at this stage must be IGWs. These waves must have been generated by the growth and decay of the vortex, that is, only by the initialization procedure, not by the spontaneous generation of IGWs by the balanced flow (there is no PV advection by the axisymmetric balanced flow).

A priori, we expect significant IGW generation for short ramp periods T_I. But what about long ramp periods? For the results presented next, we take

$$\varpi_0(\boldsymbol{X}) = \varpi_m(1 - X^2/a^2 - Y^2/a^2 - Z^2/c^2),$$

with $\varpi_m=-0.75$ (an anticyclonic anomaly), $a=N/f=100$, and mainly consider $c=1$. This corresponds to a spherical PV anomaly in vertically stretched material coordinates, $X'=X$, $Y'=Y$, and $Z'=NZ/f$. As the flow develops, isopycnals are displaced and PV contours expand or contract in physical space so the vortex distorts in shape yet remains roughly spherical. This distortion excites IGWs.

We examine this next by computing the mean-square kinetic energy $\langle(u^2+v^2)/2\rangle$, the mean-square isopycnal displacement $\langle\mathscr{D}^2\rangle$, where $\mathscr{D}=-b/N^2$ and b is the buoyancy, and the mean-square vertical velocity $\langle w^2\rangle$. Here $\langle\ \rangle$ means a spatial average over the complete three-dimensional domain. The evolution of these quantities is shown in figures 1(a)–1(c) for a selected set of ramp periods. Both $\langle(u^2+v^2)/2\rangle$ and $\langle\mathscr{D}^2\rangle$ exhibit a maximum very close to $t=T_I$ when the vortical flow is fully developed, and

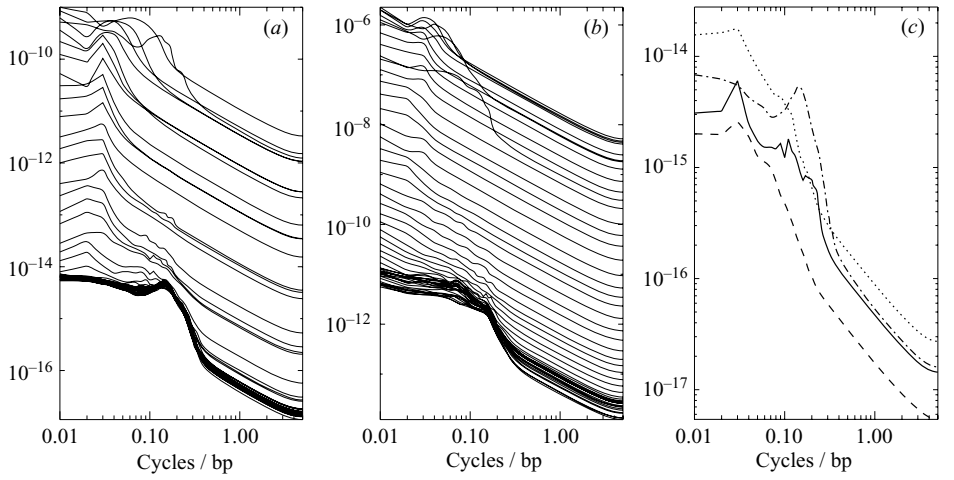

FIGURE 2. Frequency spectra of (*a*) w^2 (for the spherical vortex, with the profiles for all ramp periods T_I superposed), (*b*) as in (*a*) but for the kinetic energy $(u^2 + v^2)/2$, and (*c*) w^2 for $T_I = 5.05T_{\rm ip}$ but for four different vortex aspect ratios, $Nc/fa = 0.25$ (dashed), 0.5 (dotted), 1 (dash-dot), and 2 (solid). All vortices have the same total volume. In general, spectral amplitudes drop with increasing T_I. In the horizontal-axis label, 'bp' stands for 'buoyancy period'.

the value of this maximum varies little with T_I. After $t = 2T_I$, when there is no PV anomaly, these two quantities decrease by more than two orders of magnitude. That is, the residual flow consisting purely of IGWs is much smaller than the balanced PV-controlled flow which exists when the vortex is fully spun up. Moreover, the IGWs diminish rapidly in amplitude as T_I increases, up to about $T_I = 3T_{\rm ip}$, after which there is only a slow decrease. By then the residual flow is seven orders of magnitude smaller than the dominantly balanced flow at $t = T_I$. Interestingly, the largest excitation of IGWs does not occur for the shortest ramp period but rather for $T_I \approx 0.35T_{\rm ip}$, presumably because the time variation of the ramp is too fast to excite IGW frequencies lying between f and π/T_I.

The picture for $\langle w^2 \rangle$ is somewhat different. First of all, there are generally two main peaks, separated by a minimum at $t = T_I$, except for short ramp periods. These two peaks occur when the vertical displacement of isopycnals is changing most rapidly in response to the growth or decay of the vortex. The minimum occurs when the vortex momentarily stops growing at $t = T_I$. This evolution is again dominated by the balanced motion, as it appears to respond directly to the PV evolution on the ramp time scale. However, the IGWs are now much more apparent. They are at most two orders of magnitude smaller than the peaks in $\langle w^2 \rangle$ (some of this could be due to saturation of numerical error), but can be larger – by about a factor 3 – for short ramp periods. Also, $\langle w^2 \rangle$ is six to nine orders of magnitude smaller than $\langle (u^2 + v^2)/2 \rangle$ in all cases. As for $\langle (u^2 + v^2)/2 \rangle$ and $\langle \mathscr{D}^2 \rangle$, the residual flow in $\langle w^2 \rangle$ after $t = 2T_I$ diminishes rapidly up to about $T_I = 3T_{\rm ip}$ but then varies little for longer ramp periods.

This is summarized, for all three quantities, in figure 1(*d*), which plots their time-mean values over the final inertial period when only IGWs are present, as a function of the ramp period T_I. This verifies that there is a transition around $T_I = 3T_{\rm ip}$, beyond which the ramp period has little influence. Convergence occurs also in the IGW frequency spectra obtained during the last inertial period; see figures 2(*a*) and 2(*b*) for w^2 and $(u^2 + v^2)/2$. For very short T_I, the greatest excitation occurs around a

frequency π/T_I ($=10f$ or 0.1 cycles per buoyancy period when $T_I = 0.05T_{\rm ip}$). These IGWs are therefore a direct response to the ramped initialization. As T_I increases, they become much less important and, for $T_I \gtrsim 3T_{\rm ip}$, are overtaken by other very weak IGWs which do not depend on T_I. As shown in figure 2(*c*), these residual IGWs depend on the vortex aspect ratio, peaking around $Nc/fa = 0.5$. However, they are of very small amplitude in all cases (it is likely that they arise from the distortion of the initially circular vortex in the periodic domain). The converged spectrum for w^2, for the spherical vortex, exhibits two peaks, at f and around $15f$, before steeply falling at higher frequencies (note that the buoyancy frequency is hardly excited). The converged spectrum for $(u^2+v^2)/2$ exhibits a single broad peak at f.

This example, while highly idealized, is consistent with previous work indicating that IGW generation can be greatly suppressed by this ramped PV initialization procedure for moderately short ramp periods (VD03; DV03; VD04). This example also gives a lower bound on IGW generation, as more realistic flows are much more time dependent. Finally, it shows that the ramp period can be used to tune the initial amplitude of IGWs in a flow. This will be exploited in the results presented next.

5. Unstable jet simulations

5.1. *Flow specification and decomposition*

We next examine the flow conditions which favour the generation of IGWs and we quantify the importance of balance in the evolution of complex flows. The results were obtained from a set of 72 numerical simulations covering a wide range of parameter space spanned by the maximum PV anomaly $|\varpi|_{\max}$, the initial jet height–width aspect ratio H/L, and the initialization period T_I. The initial jet consists of two perturbed elliptic cylinders of height H, width L, and of opposite PV anomaly placed side by side with unperturbed axes parallel to the y axis. The cylinders are perturbed by displacing the PV contours by $\epsilon(\sin 2\hat{y} + \sin 3\hat{y})$ in x, where $\hat{y} = yf/N$ and $\epsilon = 0.1N/f$. The jet shape is taken to be nearly isotropic in scaled coordinates $x, y, Nz/f$ by choosing the scaled height–width aspect ratio NH/fL to be either $\sqrt{2}/2$, $\sqrt{2}$, or $2\sqrt{2}$. When $NH/fL = \sqrt{2}$ the unperturbed double-cylinder system just fits into a circular cylinder of radius L in scaled coordinates. In all cases, we take $L = 0.5N/f$ (the domain width is $2\pi N/f$) and a frequency ratio $f/N = 0.1$, comparable with typical oceanic values. The maximum PV anomaly $|\varpi|_{\max}$ is chosen to be either 0.25, 0.5, or 0.75, values which approximately correspond to the peak Rossby numbers observed during the evolution of these flows. The flow is spun up from rest by ramping up the PV anomaly ϖ on fluid particles between $t = 0$ and $t/T_{\rm ip} = T_I/T_{\rm ip} = 0.01$, 0.1, 0.25, 0.5, 1, 2, 4 and 5. Thereafter, the PV is fixed on fluid particles and the equations are integrated until $t = 30T_{\rm ip}$. The PV anomaly (ϖ) and the y component of the velocity (v) at the end of the longest initialization period are shown in the lower half of the domain and for $y = 0$ in figure 3, for the three jet aspect ratios mentioned above and for $|\varpi|_{\max} = 0.75$. The pattern of v is consistent with dominantly balanced motion, as demonstrated below.

These parameter variations allow for a wide exploration of flow nonlinearity or ageostrophy, measured by the Rossby and Froude numbers, $Ro = |\zeta|_{\max}/f$ and $Fr = |\boldsymbol{\omega}_h|_{\max}/N$, as well as the initial level of IGWs, which are controlled by T_I, as seen in the previous section. The objective is to determine the impact of IGWs in these flows – *when* do they matter or, in fact, do they matter at all?

To this end, we diagnose the balanced part of each flow $\boldsymbol{\varphi}$ every inertial period using the OPV balance procedure outlined above, with ramp period $\Delta_\tau = 5T_{\rm ip}$. The

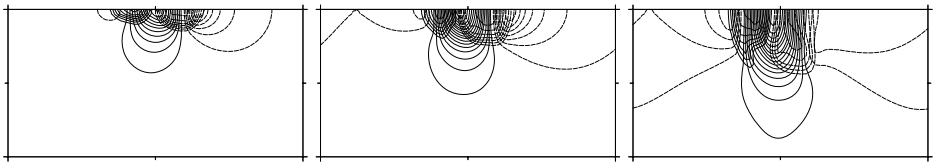

FIGURE 3. PV anomaly ϖ (bold contours) and y velocity component v (thin contours) at the end of the initialization period $t = T_I = 5T_{\rm ip}$ in the lower half of the domain in the plane $y = 0$, for three cases having different jet aspect ratios: (*a*) $NH/fL = \sqrt{2}/2$, (*b*) $NH/fL = \sqrt{2}$, and (*c*) $NH/fL = 2\sqrt{2}$. The three cases have the same maximum PV anomaly, $|\varpi|_{\max} = 0.75$. The contour intervals are $\Delta\varpi = 0.1$ and $\Delta v = 0.1$. The dashed contours correspond to negative field values. The zero contour is omitted. Note that the fields are symmetric about $z = 0$, and z has been stretched by $N/f = 10$.

difference $\boldsymbol{\varphi}_{\rm i}$ from the actual flow $\boldsymbol{\varphi}$ is defined to be the unbalanced flow: $\boldsymbol{\varphi}_{\rm i} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_{\rm b}$. (Recall that the original variables $\boldsymbol{u}$ and b are recovered from spatial derivatives of $\boldsymbol{\varphi}$.) $\boldsymbol{\varphi}_{\rm i}$ is associated with IGWs, though $\boldsymbol{\varphi}_{\rm b}$ is not entirely free of IGWs, as seen in the previous section, but is nearly so.

5.2. *Vortex dynamics*

We begin by illustrating the flow evolution in three selected cases differing only in the aspect ratio of the initial jet. The three cases share the same maximum PV anomaly $|\varpi|_{\max} = 0.75$ and were initialized with the same ramp period $T_I = 5T_{\rm ip}$. Figure 4 compares the PV field (PV contours) lying on the middle isopycnal surface ($z \approx 0$) at times $t = 10$, 20, and $30T_{\rm ip}$. In all cases, the initial jet breaks down into a street of cyclonic and anticyclonic vortices, which propagate throughout the domain. Note that the evolution advances at a faster pace for the larger-aspect-ratio cases. This is simply because tall vortices rotate faster – they have a larger Rossby number; see figure 5. The anticyclonic vortices are marginally more coherent and circular than the cyclonic ones, and they rotate faster as well (cf. figure 5), verifying that ageostrophic effects are significant. Finally, PV gradients sharpen dramatically, as seen by the rapid decrease in contour spacing in response to the stretching and twisting of material lines. This is a characteristic feature of inviscid fluid flows.

5.3. *Balance and imbalance*

We examine next the nature of the IGWs generated over the course of the flow evolution. Consider first two cases having very different levels of IGWs, one initialized with $T_I = 0$, i.e. with $\mathscr{A}_h = 0$ initially, and the other with $T_I = 5T_{\rm ip}$, but otherwise identical ($|\varpi|_{\max} = 0.75$ and $NH/fL = \sqrt{2}$). The full, balanced, and unbalanced vertical velocity fields (w, $w_{\rm b}$, and $w_{\rm i}$) at an early time are shown for contrast in figure 6. In the case with $T_I = 0$, the vertical velocity is dominated by IGWs (58 % of the total, see below), while in the other case, the vertical velocity is essentially balanced (the imbalance is only 3 %). One might therefore expect a significant impact of IGWs in the first case. However, the flow is hardly being driven by vertical motions, which – in all the simulations – are never much in excess of 1 % of the horizontal motions. Moreover, the latter are dominantly balanced, as shown in figure 7 for $|\boldsymbol{u}_h|$ using the same format as figure 6.

In time, IGWs are generated and build up even in the carefully initialized flow with $T_I = 5T_{\rm ip}$. This is shown in figure 8, for the percentage imbalance in vertical velocity, in the nine cases, encompassed by $|\varpi|_{\max} = 0.25$, 0.5, and 0.75 (the dashed, dotted,

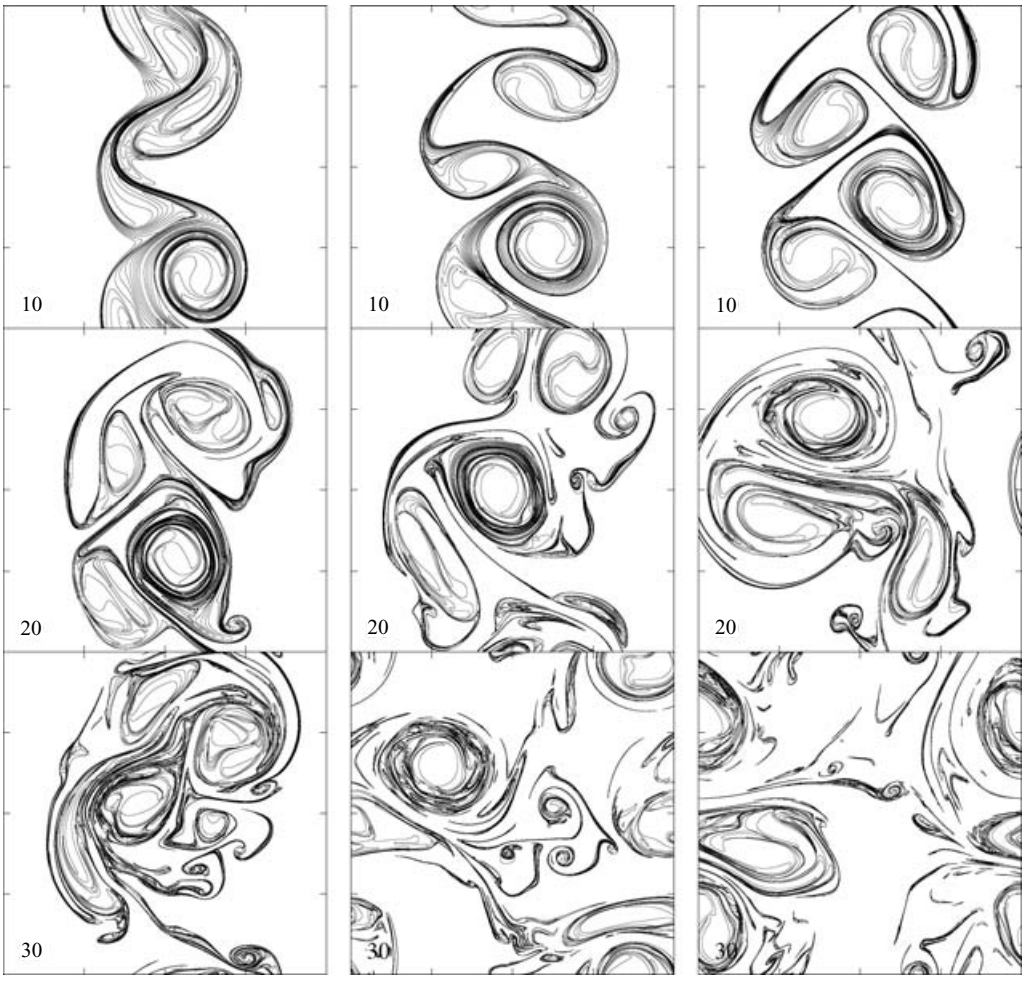

FIGURE 4. Snapshots of the PV field (at $10T_{\rm ip}$, $20T_{\rm ip}$ and $30T_{\rm ip}$) in the middle isopycnal surface ($z \approx 0$) for the three cases illustrated in figure 3: (*a*) $NH/fL = \sqrt{2}/2$, (*b*) $NH/fL = \sqrt{2}$, and (*c*) $NH/fL = 2\sqrt{2}$. The contour interval is $\Delta\varpi = 0.075$.

and solid curves) and $NH/fL = \sqrt{2}/2$, $\sqrt{2}$, and, $2\sqrt{2}$ (the thin, medium, and thick lines). For $|\varpi|_{\max} = 0.75$, a sharp rise in the percentage imbalance occurs between $7T_{\rm ip}$ and $8T_{\rm ip}$ – this is identified as spontaneous IGW emission (and described in detail in Viúdez & Dritschel 2006). For lower values of $|\varpi|_{\max}$ the rise occurs later and is less pronounced, resulting in less imbalance at late times. Remarkably, there is little dependence on the flow aspect ratio, NH/fL.

5.4. *Impact of inertia–gravity waves*

While the IGWs in the above case for $T_I = 0$ are small, one may wonder if, over a long period, their effects might become significant. To check this, we next compare the above case with another starting from the OPV balanced flow at $t = 0$, integrated to $t = 30T_{\rm ip}$. In every respect the simulations are identical, apart from the initial IGWs, which are absent in the initially balanced flow. Subsequently, imbalance grows in the initially balanced simulation but remains well below the level found in the original

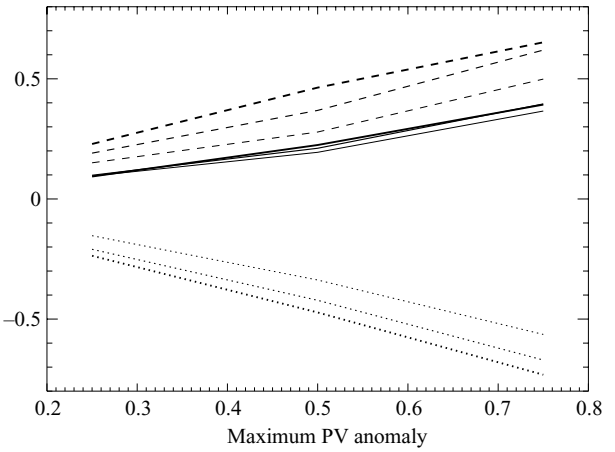

FIGURE 5. Time-mean maximum (dashed) and minimum (dotted) Rossby numbers and maximum (solid) Froude numbers vs. $|\varpi|_{\max}$. Bold, medium, and thin lines are used for $NH/fL = 2\sqrt{2}$, $\sqrt{2}$ and $\sqrt{2}/2$ respectively. The time-mean is computed between $t = 5T_{\mathrm{ip}}$ and $t = 30T_{\mathrm{ip}}$. These values are independent of T_I to within the plotted line width. Note that the Froude number is nearly independent of NH/fL.

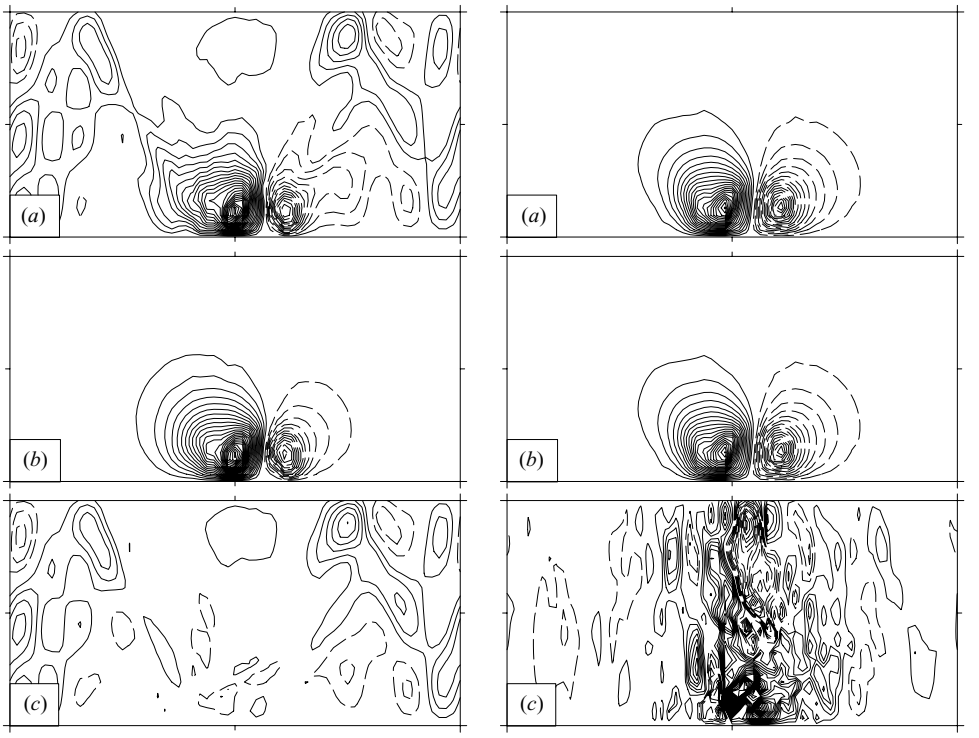

FIGURE 6. Vertical cross-sections of (*a*) w, (*b*) w_{b}, and (*c*) w_{i} at $y = 0$ (and for $z \geqslant 0$ only) in two jet simulations: with $T_I = 0$ (left) at $t = 2T_{\mathrm{ip}}$, and $T_I = 5T_{\mathrm{ip}}$ (right) at $t = 5T_{\mathrm{ip}}$. Both simulations have $|\varpi|_{\max} = 0.75$ and $NH/fL = \sqrt{2}$. The contour interval is 2×10^{-4}, except for w_{i} (right), for which it is 100 times smaller. The fields are antisymmetric across $z = 0$.

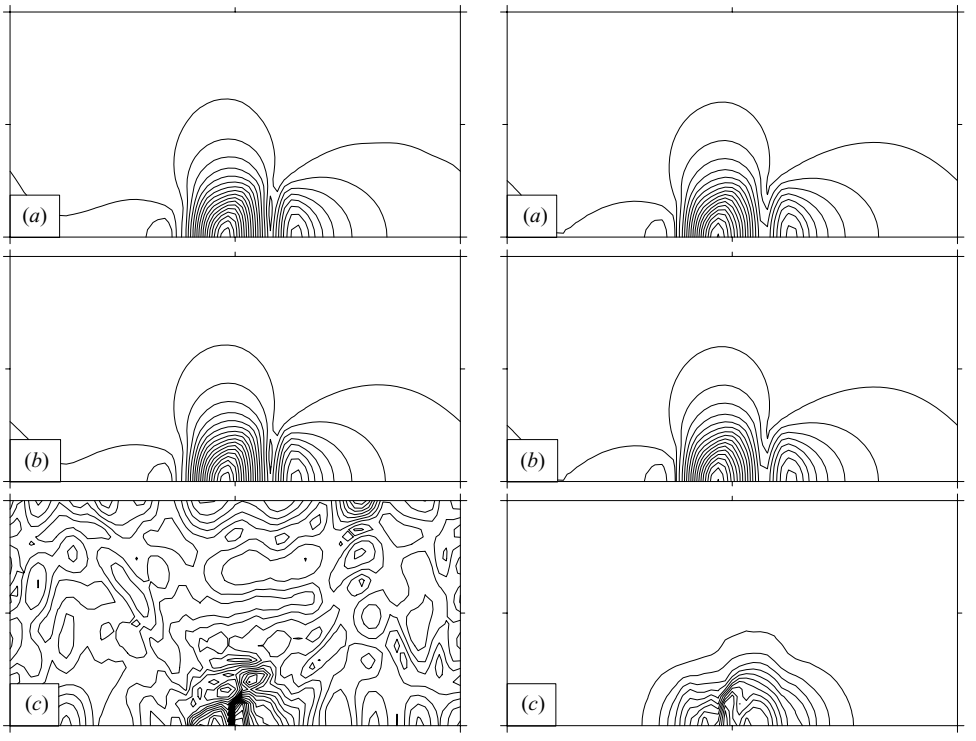

FIGURE 7. As in figure 6, except that here cross-sections of (*a*) $|\boldsymbol{u}_h|$, (*b*) $|\boldsymbol{u}_{h\mathrm{b}}|$, and (*c*) $|\boldsymbol{u}_{h\mathrm{i}}|$ are shown. The contour interval is 0.1 for $|\boldsymbol{u}_h|$ and $|\boldsymbol{u}_{h\mathrm{b}}|$, while it is 0.001 for $|\boldsymbol{u}_{h\mathrm{i}}|$.

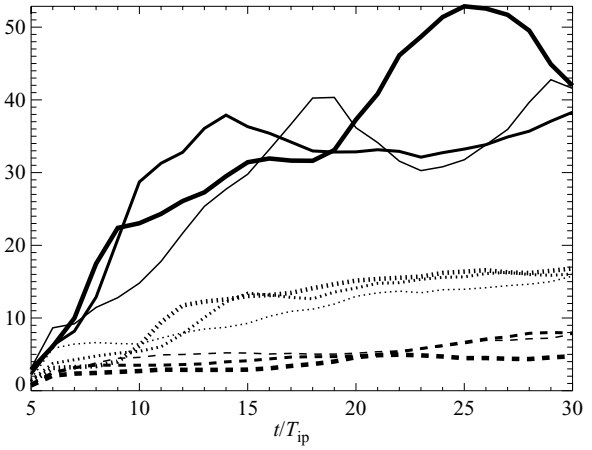

FIGURE 8. The percentage imbalance in w, as measured by the ratio of mean absolute values $\langle|w_{\mathrm{i}}|\rangle/\langle|w|\rangle$ vs. time, for nine simulations all initialized using $T_I = 5T_{\mathrm{ip}}$ but differing in the maximum PV anomaly $|\varpi|_{\max}$ and the cylinder aspect ratio NH/fL. The dashed, dotted, and solid curves correspond to $|\varpi|_{\max} = 0.25$, 0.5, and 0.75, while the thin, medium, and thick lines correspond to $NH/fL = \sqrt{2}/2$, $\sqrt{2}$, and $2\sqrt{2}$.

simulation; see figure 9. This imbalance results from spontaneous emission of IGWs from the balanced flow occurring around $t = 7T_{\mathrm{ip}}$ (cf. figure 8) and amounts to about

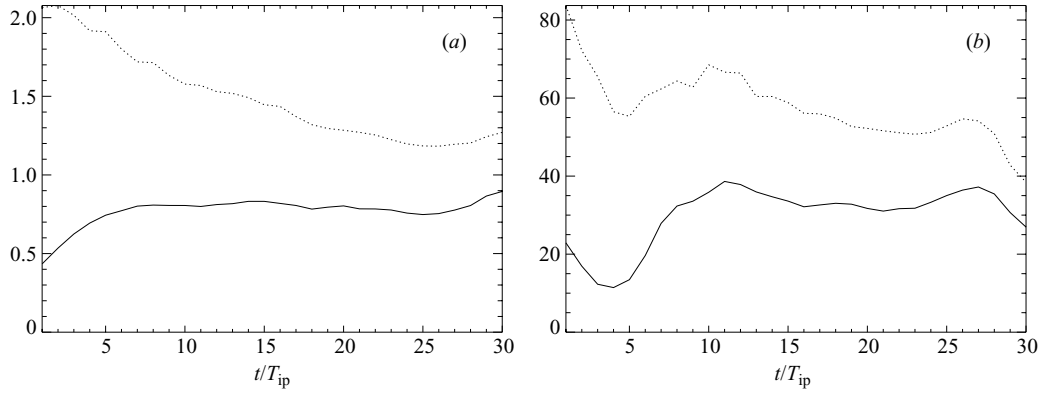

FIGURE 9. The percentage of imbalance (*a*) in $\boldsymbol{u}_h$, and (*b*) in w, vs. time, for two simulations: an initially unbalanced one with $T_I = 0$ (dashed lines), and an initially balanced one (solid lines), as described in the text. Note that t starts from $1T_{\rm ip}$ since $w = 0$ at $t = 0$ in the unbalanced simulation.

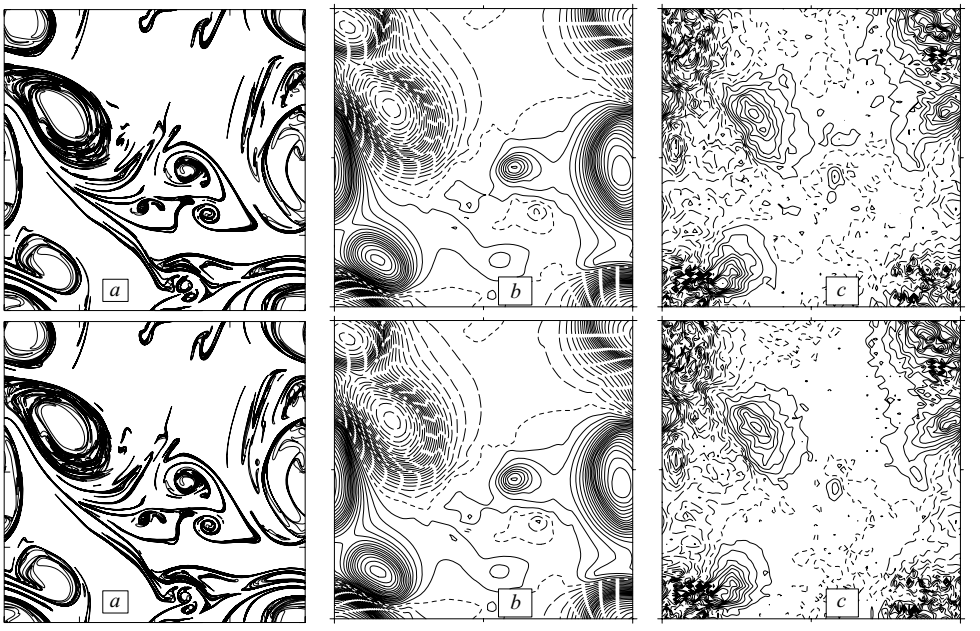

FIGURE 10. Horizontal cross-sections of (*a*) ϖ, (*b*) $\mathscr{D}$, and (*c*) w at $t = 30T_{\rm ip}$ in the initially unbalanced (top row) and balanced (bottom row) simulations illustrated in figure 9. The PV is shown in the middle isopycnal surface ($z \approx 0$), while $\mathscr{D}$ and w are shown at $z = -0.3927$, the level of maximum w. The contour intervals are $\Delta\varpi = 0.075$, $\Delta\mathscr{D} = 5 \times 10^{-3}$, and $\Delta w = 2.5 \times 10^{-4}$. After $t = 7T_{\rm ip}$, strong IGW activity is observed in w in both simulations.

32 % of the mean magnitude of the vertical velocity and 0.8 % of the horizontal velocity at late times. So what about the long-term impact? The fields of PV ϖ, displacement $\mathscr{D}$, and vertical velocity w may be compared at the final time $t = 30T_{\rm ip}$ in figure 10. Hardly any differences at all can be seen in the fields ϖ and $\mathscr{D}$. Even the w fields are closely similar, despite the 11 % difference in unbalanced vertical velocity at late times (cf. figure 9). This level of difference is less than the numerical error, as

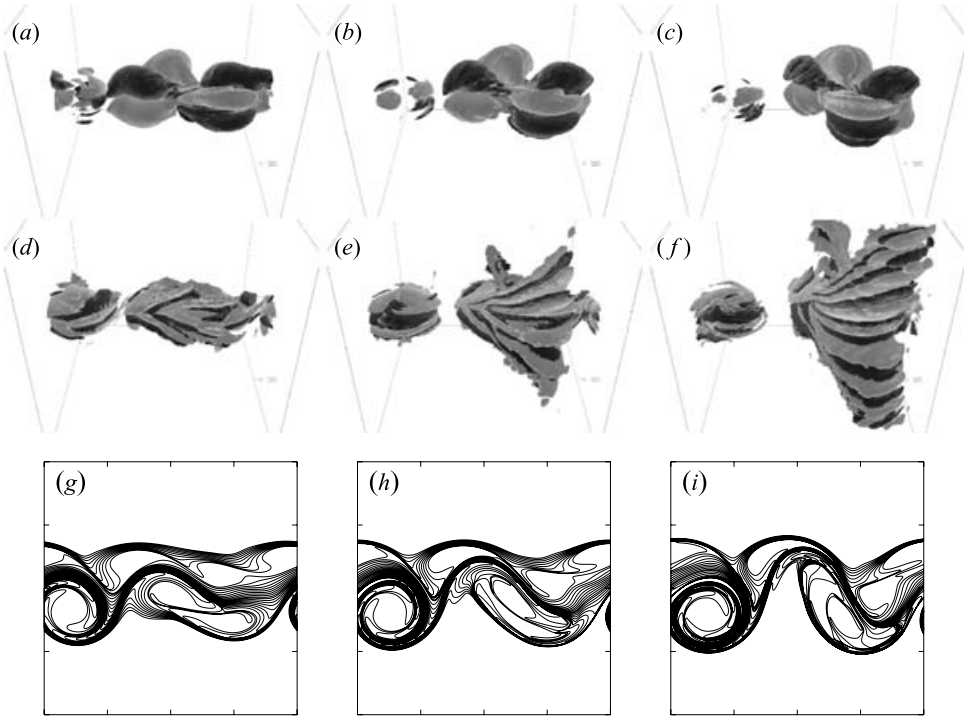

FIGURE 11. Three-dimensional perspective images of the vertical velocity field w at $t/T_{\rm ip} = 6$, 7, and 8 (top row, left to right), $w_{\rm i}$ at the same times (middle row), and PV contours at $z = 0$ (bottom row). The initial jet axis runs from left to right through the centre of the domain (the PV contours have been rotated by 90° for clarity). The isosurfaces shown are $w = \pm 10^{-3}$ and $w_{\rm i} = \pm 10^{-4}$, with negative values darkly shaded and positive values lightly shaded. These images were obtained from a special double-resolution simulation with $|\varpi|_{\max} = 0.9$, $NH/fL = \sqrt{2}/2$, and $T_I = 5T_{\rm ip}$.

deduced by doubling the resolution (convergence results may be found in Viúdez & Dritschel 2006). While IGWs are better resolved at higher resolution, their effects on the balanced flow appear to be no more significant.

5.5. *Three-dimensional structure of IGW emission*

Up to a maximum PV anomaly of $|\varpi|_{\max} = 0.75$, the IGWs are so weak that they are difficult to visualize, particularly at 64^3 resolution. A special simulation was therefore conducted at double resolution (in all directions) and for a larger PV anomaly, $|\varpi|_{\max} = 0.9$, for a shallow jet with $NH/fL = \sqrt{2}/2$. In this case, the time-mean Froude number is 0.7, and the maximum Froude number exceeds 0.8 (cf. figure 5), values which *a priori* might be associated with much greater IGW activity. This case closely parallels those for $|\varpi|_{\max} = 0.75$ described above (cf. figures 6–8). Here, a steep rise in the percentage imbalance in w occurs a little earlier, between $6T_{\rm ip}$ and $7T_{\rm ip}$ (not shown), but does not result in substantially more imbalance, in percentage terms, during the course of the simulation. While $\langle |w_{\rm i}| \rangle$ is roughly ten times larger here than in the corresponding case with $|\varpi|_{\max} = 0.75$, $\langle |w| \rangle$ is roughly twenty times larger. That is, the ageostrophic balanced vertical velocity also grows steeply with $|\varpi|_{\max}$ (or with Froude number), keeping the imbalance-to-total ratio in check. (Resolution and numerical damping differences also probably contribute to the observed differences; this is currently under investigation.) The vertical velocity field between $6T_{\rm ip}$ and $8T_{\rm ip}$ is shown in figure 11, in a three-dimensional perspective view, for both w (top row) and

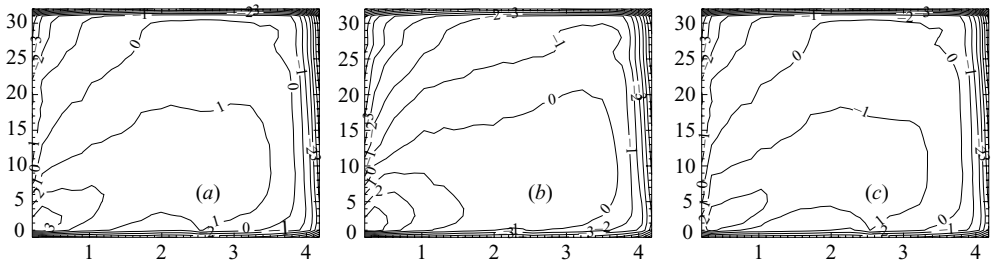

FIGURE 12. Two-dimensional (k_h, k_z) spatial spectra for (a) w, (b) w_b, and (c) w_i for a simulation with no initialization, $T_I = 0$, $|\varpi|_{\max} = 0.75$, and $NH/fL = \sqrt{2}$. The spectra are time-averaged from $t = 11T_\mathrm{ip}$ to $30T_\mathrm{ip}$.

w_i (middle row). Note that the magnitude of w_i shown is only a tenth of the magnitude of w. The PV contours at $z = 0$ are shown in the bottom row at the same times. These images reveal that the source of IGWs is located near the jet axis, most prominently in the three-way interaction occurring for $y > 0$ between two merging anticyclonic vortices and one cyclonic vortex, which subsequently form a mismatched vortex pair. In time, the waves fan out and propagate upwards and downwards, although some waves appear trapped in the vicinity of the larger, more circular, anticyclonic vortex for $y < 0$. Nevertheless, these represent weak motions, even weaker in the horizontal velocity for which the time-mean percentage imbalance is just 1.3 %. The upshot is that the IGWs have little impact on the evolution of these flows, even for $O(1)$ Rossby and Froude numbers. Much larger IGWs are needed. Large IGWs cannot be generated by ramping up the PV anomaly in this formulation of the equations. This is an important result and is discussed further in the conclusions.

5.6. *Spectra*

Spatial and temporal spectra are next examined for the two cases, discussed in § 5.3, for which $T_I = 0$ and $T_I = 5T_\mathrm{ip}$. For the case with no initialization, the time-averaged spatial spectra of w, w_b, and w_i, obtained by averaging over the horizontal wavenumber shells k_h for each vertical wavenumber k_z, show a clear separation of balanced and unbalanced motions; see figures 12(a)–(c). (Similar but less distinctive results are found for $T_I = 5T_\mathrm{ip}$.) Balanced vertical motions dominate at low wavenumbers, in both k_h and k_z, while imbalanced motions dominate elsewhere. More significantly, the spectra are nearly isotropic in the scaled wavenumbers $(N/f)k_h$ and k_z. This is characteristic of balanced dynamics at small Rossby and Froude numbers, yet here even IGWs exhibit approximate isotropy. That is, they exhibit vertical-to-horizontal scale ratios comparable with those exhibited by the balanced vortical structures. The structure of the PV field appears to dictate the structure of the IGW field.

The temporal frequency spectra for the two cases $T_I = 0$ and $T_I = 5T_\mathrm{ip}$ are compared next. These spectra are obtained by spatially averaging the single-point spectra at a subset of $8 \times 8 \times 16$ grid points distributed uniformly throughout the domain. Spectra are computed for the vertical velocity and, additionally, for the horizontal ageostrophic vorticity $\mathscr{A}_h$, which contains horizontal motions due to IGWs. In particular, $\mathscr{A}_h$ can measure near-inertial oscillations (having no vertical motion), whose frequencies lie in the vicinity of f. Figures 13(a) and 13(b), for $T_I = 0$ and $T_I = 5T_\mathrm{ip}$ respectively, contrast the clockwise and anticlockwise spectra corresponding to $\mathscr{A}_h$. The difference, more apparent in (a), is associated with near-inertial oscillations (NIOs): for such motions, horizontal wavevectors rotate exclusively clockwise in time. The spread in

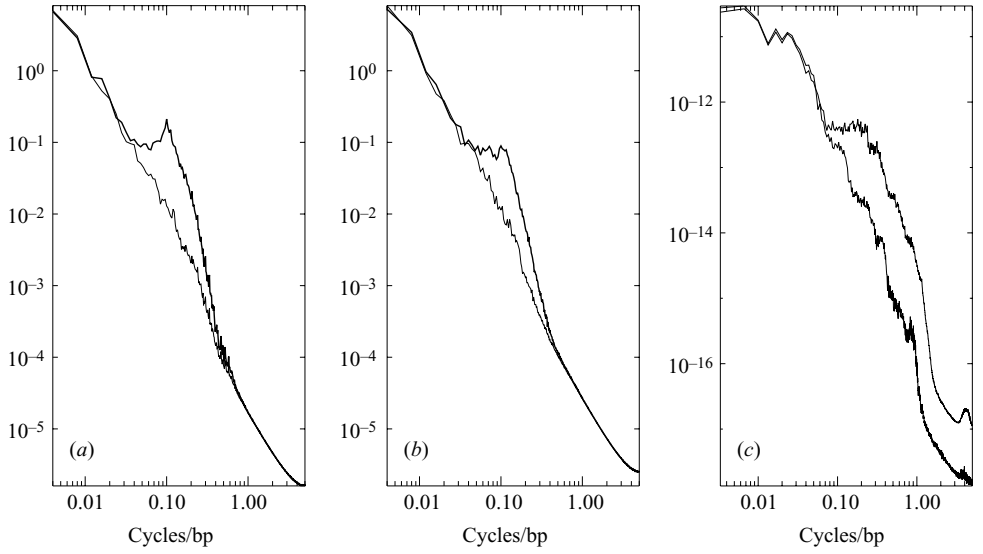

FIGURE 13. Frequency spectra for (*a*) $\mathscr{A}_h$ in the case $T_I = 0$, with the clockwise spectrum rendered bold and the anticlockwise spectrum rendered thin. (*b*) $\mathscr{A}_h$ in the case $T_I = 5T_{\rm ip}$, and (*c*) w for both $T_I = 0$ (bold line) and $T_I = 5T_{\rm ip}$ (thiner line).

frequencies about f is caused by the PV anomalies, which reduce or increase the effective background rotation. The remaining IGWs, which depend on stratification, rotate in either direction and show up clearly in the anticlockwise spectrum. That the NIOs in (*b*) are weaker shows that they can be partially suppressed by careful initialization.

Careful initialization also reduces the amplitudes of the remaining IGWs, as seen in figure 13(*c*) for the vertical velocity spectrum. IGWs can be seen to extend from around f to N, with much weaker amplitudes near N than near f. Note that there is no peak around f because NIOs have no vertical velocity (albeit only in the absence of PV anomalies).

5.7. *Summary of parameter sweep*

The results of all 72 simulations are summarized next, in figures 14–16, in which the time-mean percentage of imbalance in $\boldsymbol{u}_h$ and w is plotted as a function of the initialization period T_I for the three jet aspect ratios. In all cases, the curves level off for $T_I \gtrsim 2T_{\rm ip}$, indicating that a nearly balanced flow state is achieved for moderately short initialization periods. Longer initialization does not reduce the imbalance further because of IGW emission throughout the flow evolution. This emission is clearly strongest for the largest PV anomalies, corresponding to the largest Rossby numbers, as expected. However, it depends only weakly on the Froude number. Note that the percentage of imbalance approximately doubles as one goes from small to moderate Rossby numbers. In the limit of small Rossby numbers, quasi-geostrophic scaling (using geostrophic and hydrostatic balance) implies that $\boldsymbol{u}_h = O(Ro)$ and $w = O(Ro^2 f/N)$ (cf. McKiver & Dritschel 2006). These apply to the balanced flow. Assuming that this scaling is valid here, the results shown in figures 14–16 for $T_I \gtrsim 2T_{\rm ip}$ indicate that the unbalanced flow components are one order of Ro smaller, for $Ro \ll 1$.

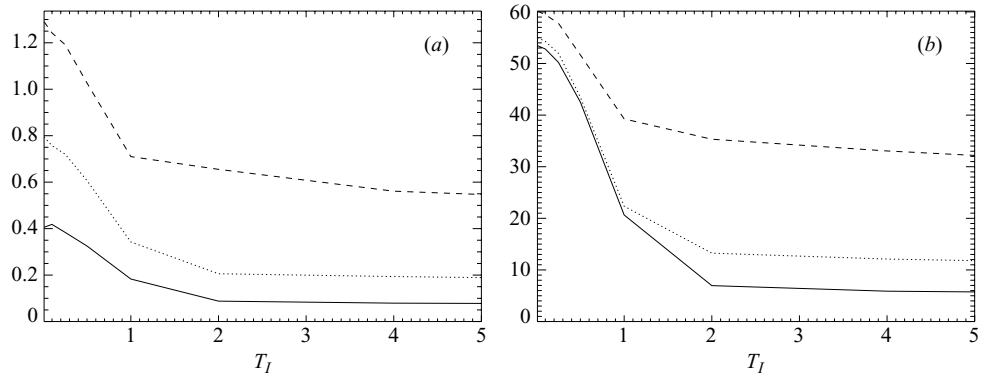

FIGURE 14. The time-mean percentage of imbalance (*a*) in $\boldsymbol{u}_h$, and (*b*) in w, vs. the initialization period T_I, for a jet aspect ratio of $NH/fL = \sqrt{2}/2$. The dashed, dotted, and solid lines correspond to $|\varpi|_{\max} = 0.75$, 0.5, and 0.25, respectively. The time mean is computed between $10T_{\rm ip}$ and $30T_{\rm ip}$.

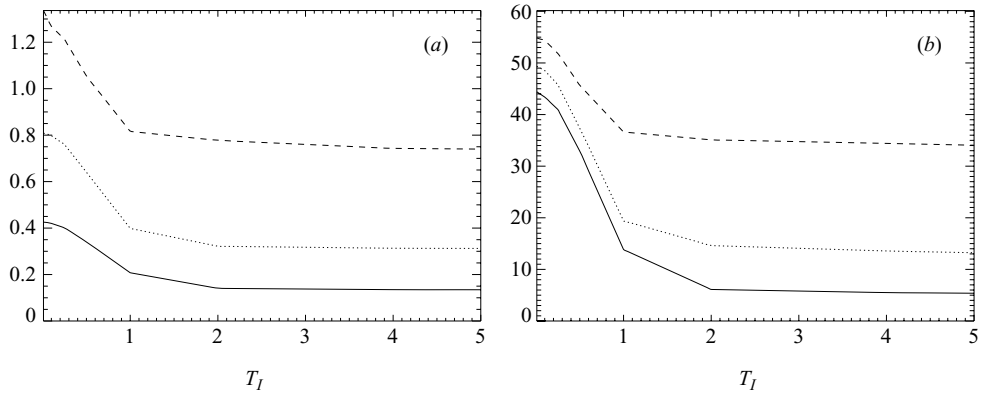

FIGURE 15. As in figure 14, but for a jet aspect ratio of $NH/fL = \sqrt{2}$.

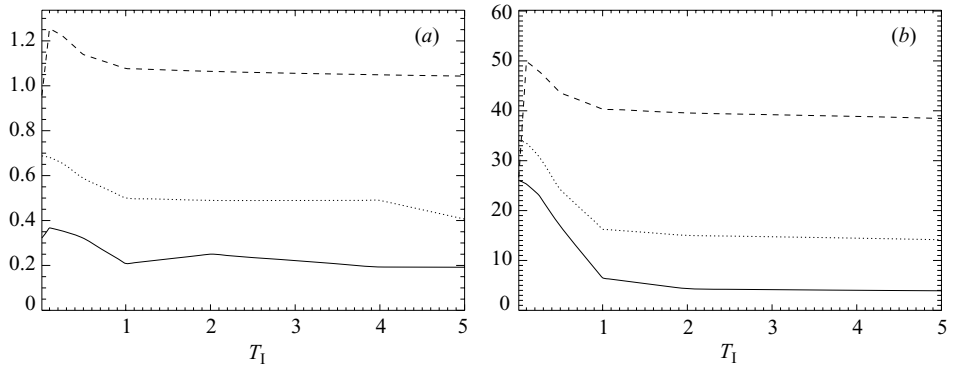

FIGURE 16. As in figure 14, but for a jet aspect ratio of $NH/fL = 2\sqrt{2}$.

6. Concluding remarks

Potential vorticity (PV) has long been recognised as the key quantity driving large-scale fluid motions in the atmosphere and oceans (Hoskins *et al.* 1985; Ford

et al. 2000). Fundamentally, this notion rests upon the constraints imposed by rotation and stratification, which act together to inhibit the overturning of material surfaces and the intensification of vorticity. Certain balances are established, approximately, between the PV and other dynamical fields which would otherwise be free to evolve with far greater independence. Hydrostatic and geostrophic balance are examples but are only the simplest. The question is: how much of the flow does the PV actually control?

The present work has demonstrated that a wide class of flows within realistic geophysical parameters exhibit a very high degree of balance, to the extent that the residual motions, the inertia–gravity waves, play almost no role. This is consistent with the recent small-Rossby-number asymptotic analysis of Zeitlin, Reznik & Ben Jelloul (2003), who demonstrated a decoupling of balanced and imbalanced motions for times $t \gg T_{\rm ip}/Ro$. Our results suggest that this decoupling remains a good approximation up to $Ro = O(1)$, at least for flows which are nearly balanced initially. Similar conclusions were reached in a recent work on fully turbulent flows (McKiver & Dritschel 2006). We conjecture that the persistence of balance is limited only by inertial and static stability.

Interestingly, this class of flows was created using a novel approach which features PV explicitly (DV03). This forces one to reformulate the governing equations in terms of a new set of variables, from which the original set (velocity, buoyancy, and pressure) are recovered diagnostically, via the inversion of elliptic operators. Here initial conditions are created for the other dynamical fields (specifically, the ageostrophic horizontal vorticity), by artificially ramping up the PV anomaly on fluid particles to its desired level while integrating the full equations of motion. Previous studies found that long ramp periods, of several inertial periods or more, establish a nearly balanced initial state (cf. VD04). It was anticipated that short ramp periods would excite significant inertia–gravity waves simply because the flow has to adjust rapidly to the growing PV anomaly. Inertia–gravity waves are excited, it is true, but their amplitudes are too weak to affect the dominant flow evolution over long periods of time, i.e. at least tens of inertial periods.

This observed persistence, we suggest, is fundamentally related to the explicit use of PV. A large part of the flow field is found by inverting an elliptic operator on the PV, and only a small part is left under the control of the other independent variables. Even that small part is mostly carried along by PV, almost adiabatically, by some implicit underlying 'balance' relations. These findings are fully consistent with previous studies of balance in shallow-water flows, where an exceptional degree of balance (i.e. a virtual absence of gravity waves) was found when explicitly using PV, that is, its exact unapproximated form, in the definition of balance (McIntyre & Norton 2000; Mohebalhojeh & Dritschel 2000, 2001; McKiver & Dritschel 2006).

The present model though highly-idealised, has nevertheless allowed us to assess the significance of inertia–gravity waves in the simplest context, in which those waves can arise only in response to the advection of PV. It is evident that strong inertia–gravity waves must be established in some other manner (i.e. via topography, adiabatic processes, convection, etc. in real flows). Another possibility, not yet considered in this context, is IGW emission from a sharp change in stratification, like the 'tropopause' in the atmosphere. Observations (Zhang 2004) and numerical simulations (Plougonven, Teitelbaum & Zeitlin 2003) both suggest that IGWs appear at the level of or just above the tropopause. Indeed, atmospheric jets tend to straddle the tropopause, and there is growing evidence that such jets are effective sources for IGWs, which have a major impact on the upper-level circulation (see Fritts & Alexander 2003 and

references therein). Is the tropopause the main element responsible for IGW emission, or is atmospheric compressibility (the exponential decrease of density with height) important? The present model, which assumes that the density varies only weakly, can help to disassociate these two elements and as such may help to clarify the mechanisms behind IGW emission in rotating stratified flows.

Support for this research has come from the UK Engineering and Physical Sciences Research Council (grant number XEP294), and the Spanish Ministerio de Ciencia y Tecnología (grant number REN2002-01343). We wish to thank Dr Riwal Plougonven and three anonymous referees for their helpful suggestions.

REFERENCES

Dritschel, D. G. & Ambaum, M. H. P. 1997 A contour-advective semi-Lagrangian algorithm for the simulation of fine-scale conservative fields. *Q. J. R. Met. Soc.* **123**, 1097–1130.

Dritschel, D. G. & Viúdez, A. 2003 A balanced approach to modelling rotating stably-stratified geophysical flows. *J. Fluid Mech.* **488**, 123–150.

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

Fritts, D. C. & Alexander, M. J. 2003 Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.* **41**(1), art. no. 1003.

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

McIntyre, M. E. & Norton, W. A. 2000 Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.* **57**, 1214–1235, Corrigendum **58**, 949.

McKiver, W. J. & Dritschel, D. G. 2006 Balance in non-hydrostatic rotating stratified turbulence. *J. Fluid Mech.* (under review).

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

Mohebalhojeh, A. R. & Dritschel, D. G. 2001 Hierarchies of balance conditions for the f-plane shallow water equations. *J. Atmos. Sci.* **58**(16), 2411–2426.

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

Plougonven, R., Teitelbaum, H. & Zeitlin, V. 2003 Inertia-gravity wave generation by the tropospheric mid-latitude jet as given by the FASTEX radiosoundings. *J. Geophys. Res.* **108(D21)**, art. no. 4686.

Reinaud, J., Dritschel, D. G. & Koudella, C. R. 2003 The shape of vortices in quasi-geostrophic turbulence. *J. Fluid Mech.* **474**, 175–191.

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

Vanneste, J. 2004 Inertia-gravity wave generation by balanced motion: revisiting the Lorenz-Krishnamurthy model. *J. Atmos. Sci.* **61**, 224–234.

Vanneste, J. & Yavneh, I. 2004 Exponentially small inertia-gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.* **61**, 211–223.

Viúdez, A. & Dritschel, D. G. 2003 Vertical velocity in mesoscale geophysical flows. *J. Fluid Mech.* **483**, 199–223.

Viúdez, A. & Dritschel, D. G. 2004 Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.* **521**, 343–352.

Viúdez, A. & Dritschel, D. G. 2006 Spontaneous generation of inertia-gravity wave packets by balanced geophysical flows. *J. Fluid Mech.* **553**, 107–117.

Zeitlin, V., Reznik, G. M. & Ben Jelloul, M. 2003 Nonlinear theory of geostrophic adjustment. Part 2. Two-layer and continuously stratified primitive equations. *J. Fluid Mech.* **491**, 207–228.

Zhang, F. 2004 Generation of mesoscale gravity waves in upper-tropospheric jet-front systems. *J. Atmos. Sci.* **61**(4), 440–457.