

Balance dynamics in rotating stratified turbulence

Hossein A. Kafiabad^{1,†} and Peter Bartello^{1,2}

¹Department of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke ouest, Montreal, Quebec H3A 0B9, Canada

²Department of Mathematics and Statistics, McGill University, 805 Sherbrooke ouest, Montreal, Quebec H3A 0B9, Canada

(Received 8 July 2015; revised 22 December 2015; accepted 22 February 2016;
first published online 22 April 2016)

If classical quasigeostrophic (QG) flow breaks down at smaller scales, it gives rise to questions of whether higher-order nonlinear balance can be maintained, to what scale and for how long. These are naturally followed by asking how this is affected by stratification and rotation. To address these questions, we perform non-hydrostatic Boussinesq simulations where the initial data is balanced using the Baer–Tribbia nonlinear normal mode initialization scheme (NNMI), which is accurate to second order in the Rossby number, as the next-order improvement to first-order QG theory. The NNMI procedure yields an ageostrophic contribution to the energy spectrum that has a very steep slope. However, as time passes, a shallow range emerges in the ageostrophic spectrum when the Rossby number is large enough for a given Reynolds number. It is argued that this shallow range is the unbalanced part of the motion that develops spontaneously in time and eventually dominates the energy at small scales. If the initial flow is not nonlinearly balanced, the shallow range emerges at even lower Rossby number and it appears at larger scales. Through numerous simulations at different rotation and stratification, this study gives a clear picture of how energy is cascaded in different initially balanced regimes of rotating stratified flow. We find that at low Rossby number the flow mainly consists of a geostrophic part and a balanced ageostrophic part with a steep spectrum. As the Rossby number increases, the unbalanced part of the ageostrophic energy increases at a rate faster than the balanced part. Hence, the total energy spectrum displays a shallow range above a transition wavenumber. This wavenumber evolves to smaller values as rotation weakens.

Key words: atmospheric flows, rotating turbulence, stratified turbulence

1. Introduction

In the study of rotating stratified flow two avenues of research have been explored. The first focuses on ‘balance dynamics’. In the limit of strong rotation and stratification, the linearized governing equations describe a distinct separation between fast and slow time scales. Since atmospheric and oceanic flows vary slowly, balance

[†] Email address for correspondence: hossein.aminikafiabad@mail.mcgill.ca

models try to reduce the dynamics to only the slow subset. In order to make the total nonlinear solution slowly varying, it is necessary to introduce judiciously some variance in the high-frequency linear modes. This is what we refer to as nonlinear balance and is described in detail below. We emphasize here that, although there are high-frequency linear solutions, there is no *a priori* reason why the total nonlinear solution has to display wave-like motion. In the second avenue, attempts have been made to extend turbulence theory by studying the statistics of rotating stratified flow, mainly the energy and potential-ensrophy transfer and their cascades in different limits. In this section, we briefly review both groups. Then, we discuss the focus of this study, which is merging these two avenues by looking at balance dynamics from a turbulence perspective. In particular, we investigate how an initially balanced flow evolves into a more general form of turbulence in well-resolved numerical simulations. Throughout this paper, we consider the regimes of low Rossby number (Ro) and low Froude number (Fr) defined as

$$Ro = \frac{U}{fL}, \quad Fr = \frac{U}{NH}, \quad (1.1a,b)$$

where f is the Coriolis parameter, N the Brunt–Väisälä frequency, L the horizontal length scale, H the vertical length scale and U the horizontal velocity scale.

As a result of computational advances, different aspects of high-Reynolds-number rotating stratified turbulence have been simulated recently. Bartello (2010) used quasigeostrophic forcing in the non-hydrostatic Boussinesq equations with different Rossby numbers to present large-scale quasigeostrophic flow with a small-scale transition to a more general form of turbulence. Marino *et al.* (2013) changed the relative linear frequency of gravity to inertial waves, N/f , and looked at its effect on the inverse cascade of kinetic energy. Deusebio, Vallgren & Lindborg (2013) studied the route to dissipation in strongly rotating stratified regimes by looking at the energy cascades at different rotations. Whitehead & Wingate (2014) numerically analysed the effect of fast ageostrophic modes on slow modes in three different asymptotic limits of stratification and rotation. Pouquet & Marino (2013) showed the existence of a dual cascade in one simulation. Starting from an unbalanced initial condition and randomly forcing a midrange wavenumber, their simulation showed an inverse cascade of energy at large scales, with the slope of -3 invariant to Ro . However, at wavenumbers larger than the forcing, a forward cascade of energy was established as Ro increased and the slope of the energy spectrum changed from -3 to $-5/3$.

In all these studies a wide variety of initial conditions or forcings were used, but none were nonlinearly balanced. However, a number of studies used geostrophic initial conditions or geostrophic forcing, which can be referred to as linearly balanced. Starting from an unbalanced initial condition, Bartello (1995) showed the ageostrophic energy is cascaded to the dissipation range via a catalytic interaction with geostrophic modes at low Fr and Ro . This can be seen as a form of nonlinear geostrophic adjustment that takes place by transferring ageostrophic energy to smaller scales and dissipating it. This transfer of ageostrophic energy was seen in the simulations of Deusebio *et al.* (2013) as well. It provides a mechanism to explain the observed reality of the large-scale atmosphere and ocean, which is a slowly varying balance (or something rather close to it). These scales can be thought of as synoptic scales in the atmosphere and mesoscales in the ocean, where there is little energy with frequencies much higher than planetary motion. Therefore, the fundamental question is whether balanced or unbalanced initial conditions and/or forcing change the turbulence in a

fundamental way. We address this question and show that, at low Rossby number, initially balanced flow shows considerably different evolution, at least over several eddy turnover times. However, Ro can be increased to the point that initially balanced and unbalanced flows act rather alike. Although this is not surprising, we explore the limits of these regimes.

In studies of balance dynamics the approach is different. Unlike most studies of rotating stratified turbulence, many of these start with a nonlinear balanced state and investigate if it is maintained. Of course, this depends on one's ability to define such a state and to project initial data onto it. If so, and it is indeed found that balance is maintained by the subsequent evolution, then it has been speculated there exists a slow manifold (Leith 1980; Lorenz 1980), or at least something close to it. It is strictly defined as an invariant manifold in phase space on which the dynamics are devoid of any high-frequency variability (Leith 1980; Warn *et al.* 1995; Ford, McIntyre & Norton 2000), implying the maintenance or breakdown of balance can be equated to the degree of invariance of the slow manifold. At the other extreme, in studies where fast motion is initially present, the issue is whether the state evolves towards the slow manifold.

Numerous studies have discussed the existence of the slow manifold (Lorenz 1986; Warn & Menard 1986; Lorenz & Krishnamurthy 1987; Lorenz 1992; Warn 1997; Ford *et al.* 2000). They show that the solution does not necessarily stay on it, but remains close to it in some sense. Equivalently, some high-frequency motion is observed, but its amplitude remains sufficiently small. Following this, Warn & Menard (1986) suggested the concept of a fuzzy manifold. They demonstrated that as Rossby number increases, less information on the fast modes can be deduced from the slowly varying modes alone, implying they act as independent degrees of freedom. Hence, it is not a manifold in the strict sense, but rather a 'thin' region of the full dimension of phase space.

A more recent line of research exploring the non-invariance of the slow manifold is spontaneous wave generation. Ford *et al.* (2000) showed that unsteady rotating stratified vortical flows 'must emit inertia-gravity waves'. Following this, Vanneste & Yavneh (2004) started from a particular balanced solution and showed the amplitude of inertia-gravity waves generated from an initially balanced solution grew exponentially with time as the negative of the inverse Rossby number. By using multi-scale time expansion and assuming small Ro , Zeitlin, Reznik & Jelloul (2003) showed that time splitting is valid for $t \leq (fRo)^{-1}$ (f being the Coriolis parameter). Beyond this they modified the quasigeostrophic (QG) equations to the frontal geostrophic set. Unlike in QG, they reported that in the frontal geostrophic regime time splitting is incomplete. Therefore, the vortical component and inertial oscillations evolve with similar time scales.

As an example of spontaneous wave generation, vortex dipoles have also been studied. Snyder *et al.* (2007) showed that quasigeostrophic vortex dipoles generate inertia-gravity waves that are close to stationary relative to them. They persisted after a long integration time. Therefore, it was concluded that these were inherent features of the dipoles. Viúdez (2007, 2008) investigated the characteristics of these waves further in his numerical simulations. For a comprehensive review on balance and spontaneous generation, see Vanneste (2013).

Many studies of balance dynamics and the slow manifold are restricted to low-order dynamical systems. Therefore, their results are difficult to generalize to the infinite-dimensional system governed by the Boussinesq equations. For instance, Kreiss & Lorenz (1994) showed the spatially discrete version of their

system generated an exactly invariant slow manifold, whereas its spatially continuous counterpart could not. To overcome this shortcoming, at least numerically, we study the breakdown of balance from a turbulence perspective. Unlike low-order dynamical systems, the turbulence platform of this study considers a broad distribution of wavenumbers and their interactions. In our view, a very limited number of modes can only present a limited number of triadic Rossby and Froude numbers, whereas in reality these parameters can be considered as a function of scale or wavenumber, k , ranging over many orders of magnitude in geophysical and astrophysical flows.

Even though most studies of rotating stratified flow examined either turbulence or balance aspects, there have been some exploring both. A series of papers by Dritschel and co-workers are highlights of this group. Dritschel & Viúdez (2003) first proposed a numerical approach based on integration of a balance and two imbalance variables. They chose potential vorticity as their balance and ageostrophic horizontal vorticity components as imbalance variables based on geostrophic and hydrostatic balances. In their subsequent papers, they introduced optimal potential vorticity to render their balance more precise than geostrophic and hydrostatic balance (Viúdez & Dritschel 2004). Their definition of balance was constructed to minimize inertia-gravity waves. Using nearly balanced initial conditions in this sense, McKiver & Dritschel (2008) studied properties of rotating stratified turbulence over a range of Rossby numbers. They also extracted the balanced part of the flow using optimal potential vorticity balance to access the degree of imbalance generated in time. In another recent study, Dritschel & McKiver (2015) thoroughly investigated the generation of imbalance in initially balanced flows with different frequency ratios N/f . In this series of studies, the authors found balance to be very robust in the regime of low Rossby numbers and large frequency ratios, yielding steep spectra characteristic of quasigeostrophic flow.

Another recent study that looked at balance dynamics in a turbulence context was carried out by Nadiga (2014). In this paper, the author constructed an initial condition that was in linear balance. He then investigated its subsequent evolution in two parallel simulations of the non-hydrostatic Boussinesq and quasigeostrophic equations. Comparing the two simulations using diagnostics of balance, he examined how the unbalanced part of the flow developed.

Motivated by previous work we investigate the effect of the following parameters in limiting the accuracy of balance dynamics (or the degree of approximate invariance of the slow manifold):

- (i) the order of balance;
- (ii) time;
- (iii) the length scale; and
- (iv) the strength of rotation and/or stratification.

In so doing, we first generate a set of geostrophic data using a QG model and then use a high-order initialization scheme, namely that of Baer & Tribbia (1977), to produce nonlinearly balanced ageostrophic modes. The Baer–Tribbia scheme yields initial data that are slowly varying at second order in Ro , at least at $t=0$. These sets of ageostrophic and geostrophic modes provide us with a representation of balance dynamics suitable as initial conditions in a more general model. Starting with these initial conditions we run a set of well-resolved non-hydrostatic Boussinesq simulations to study how and when balance breaks down. The compromise that is made concerns the flow boundaries. Our simulations are carried out in a triply periodic configuration

that is the only possibility in numerical studies of homogeneous turbulence. We explain the limitations and advantages of this choice.

Our platform provides several advantages that suit the detailed study of influential parameters in the breakdown of balance. First, we can see how different orders of balance affect its breakdown. The Baer–Tribbia initialization scheme that we employ can be carried out to an order higher than the QG approximation. In a similar study Bartello (2010) used fully QG initial conditions, which are linearly balanced, and studied their evolution in a non-hydrostatic Boussinesq model with forcing only in the geostrophic modes. Linear balance cannot form a complete representation of the slow manifold, as it sets all ageostrophic modes to zero, thereby rapidly generating high-frequency oscillations. Higher-order nonlinear balance schemes, on the other hand, yield sets of ageostrophic modes producing more slow evolution.

The other advantage of a turbulence-theory-based analysis is to look at the breakdown of balance through the cascade and transfer of energy. Below, we rely on the energy spectrum of normal modes. Their evolution in time draws a sufficiently clear picture of the interplay between linear fast and slow modes in a nonlinear context whose temporal variability can take on any time scale in between.

In light of analysing the breakdown of balance, we also make reference to the shallow range of the Gage–Nastrom spectrum. Gage & Nastrom (1986) used the Global Atmospheric Sampling Program (GARP) data to calculate the atmospheric kinetic and potential energy spectra. These consisted of two distinct ranges; a steep range of k^{-3} , followed by a more shallow range of $k^{-5/3}$ at smaller scales, with a relatively sharp transition between the two. The steep part of the spectrum is explained by quasigeostrophic turbulence resulting from the injection of eddy energy at the deformation scale (Charney 1971). However, QG theory does not predict the existence of the shallow range without invoking other sources or boundaries. To explain this shallow range, two separate, but not mutually exclusive, paths have been followed:

- (i) The effects of external boundaries such as topography (Vanneste 2013) or internal boundaries such as the tropopause (Tulloch & Smith 2006) that break down the balance and create a shallow range.
- (ii) The generation of unbalanced motion and more general forms of turbulence that project on other degrees of freedom (Bartello 1995, 2010; Vallgren, Deusebio & Lindborg 2011; Nadiga 2014).

Being aware of the important role of boundaries in the breakdown of balance dynamics, we choose to focus on the second issue. In addition to being more tractable statistically, it can conceivably describe the energy cascade in the ocean/atmosphere interior far from boundaries. Moreover, it shows that, even without boundaries, there is a mechanism that generates unbalanced ageostrophic flow in homogeneous turbulence, even when starting from balanced large-scale dynamics at low Ro and Fr .

Aside from the idealized simulations cited above, the shallow part of the Gage–Nastrom spectrum has been studied with Numerical Weather Prediction (NWP) and Global Climate Models (GCM) recently (Skamarock 2004; Hamilton, Takahashi & Ohfuchi 2008; Evans *et al.* 2013). The models used in these studies can simulate the atmosphere with more realistic geometry, inhomogeneity and physics. Of this group, Waite & Snyder (2009) is one of the most relevant to the current paper, since it employed a relatively high vertical resolution with a simple geometry and started from a linearly balanced initial condition describing a baroclinically unstable zonal jet. It was then perturbed with the ‘fastest-growing gravest normal mode’.

This unbalanced perturbation was added to make the unstable jet transition to turbulence, whose characteristics were then studied. In their simulations the Rossby number increased with time, which led to the shallowing of the energy spectrum at certain vertical levels. Using the Advanced Research core of the Weather Research and Forecast (WRF) model, they examined the energy spectrum at different levels in the troposphere and lower stratosphere.

It should be noted that the dynamics of NWP/GCM models are generally not nearly as well resolved as in turbulence simulations in simplified geometry, since much of the computational effort is devoted to realistic features such as surface fluxes, convection, radiation, topography and vertical inhomogeneity. Clearly, one of the weak points of such simulations, when compared with turbulence simulations, is the relatively coarse vertical grid. In addition, many global models still employ the hydrostatic approximation. The degree to which it reproduces the true dynamics at scales where balance breaks down is still an open question.

To have reliable statistics of stratified turbulence it was shown that the buoyancy scale, U/N , which is around 1 km near the tropopause, should be resolved (Waite & Bartello 2004, Lindborg 2006, Waite 2011). It has even been argued that it is necessary to resolve down to the small-scale regime of 3D isotropic turbulence, i.e. the Ozmidov scale, which is of the order of tens of metres (Brethouwer *et al.* 2007; Bartello & Tobias 2013). Knowing that these resolution requirements may not be as severe with the addition of weak rotation, the authors' view is that it is best to explore them using idealized simulations of only the dynamics in the first instance, without adding all the other complications of meteorology and oceanography.

These scales are not resolved by GCM/NWP models with today's computers. For example, Hamilton *et al.* (2008) used 24 vertical levels from the ground to approximately 1 hPa, which led to approximately 1.5 km grid spacing in the upper troposphere. Evans *et al.* (2013) employed 26 vertical levels with a model top at 2.2 hPa, resulting in an even coarser grid. Given that vertical momentum and thermal diffusion restrict the fully nonlinear scales to larger than 1.5 km, it is clear these studies come short of resolving the vertical outer scale, U/N , of stratified turbulence. However, with their subgrid dissipation schemes these simulations still offer realistic large-scale dynamics. For this reason, the authors feel their mesoscale transitions must at this point be interpreted with caution.

In addition to providing higher vertical resolution, our idealized configuration enables us to study a wider range of parameters compared to NWP/GCM models. These models are difficult to tune in order to explore the effect of resolution, rotation, stratification and dissipation. Since we model only the dynamics in as simple a geometry as possible, we were able to explore objectively their sensitivity to Rossby, Froude and Reynolds numbers (at much higher effective Reynolds numbers) in order to advance our understanding of these dynamics. For this reason we feel that studies such as this are complementary to studies using more realistic models.

The organization of this paper is as follows. Section 2 lays out the governing equations of motion and briefly describes the normal mode decomposition used in this paper. The initialization scheme that balances the ageostrophic modes given the geostrophic modes is explained in §3. In §4, we propose a three-step procedure to study the breakdown of balance dynamics. In §5, we show how the initially balanced and unbalanced flow differ in various Rossby number regimes. Our results demonstrate that the balanced part of the ageostrophic energy has a steep spectrum. However, by increasing the Rossby number, a shallow unbalanced range emerges. In fact, it seems reasonable to speculate that such a shallow tail in the ageostrophic spectrum will emerge at any Ro after a sufficient time if the Reynolds number is sufficiently large. These ideas are developed in our concluding remarks.

2. Mathematical formalism and normal modes

2.1. Governing equations

The three-dimensional equations governing the motion of rotating stratified flow under the Boussinesq approximation are

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla p + b \hat{\mathbf{z}} + \mathbf{D}_u, \tag{2.1a}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{2.1b}$$

$$\frac{\partial b}{\partial t} + \mathbf{u} \cdot \nabla b = -N^2 w + D_b, \tag{2.1c}$$

where $\mathbf{u} = (\mathbf{u}_h, w) = (u, v, w)$ is the velocity field; b is the buoyancy perturbation (which can be defined based on potential temperature as $-g\theta'/\theta_0$, or based on density as $-g\rho'/\rho_0$); p is the pressure perturbation divided by a constant reference density, ρ_0 . The operator D_q represents the dissipation of quantity q . We also assume the Coriolis parameter, f , and Brunt–Väisälä frequency, N , to be constants.

2.2. Normal mode decomposition

To close the problem mathematically one needs a set of boundary and initial conditions. We will discuss the choice of the initial conditions in the following sections. As for boundaries, we assume periodicity. This configuration maintains statistical homogeneity and provides for efficient direct numerical simulation (DNS) using pseudospectral methods.

Following Leith (1980) and Bartello (1995), we use the eigenvectors of the linearized equation as our orthonormal basis. To derive these we first take the Fourier transform of (2.1), then linearize the system around a state of rest. For each wavenumber $\mathbf{k} = (k_x, k_y, k_z)$, one finds the following eigenfrequencies (in dimensional form)

$$\lambda_{\mathbf{k}}^{(0)} = 0, \quad \lambda_{\mathbf{k}}^{\pm} = \pm \sigma_{\mathbf{k}} = \frac{(f^2 k_z^2 + N^2 k_h^2)^{1/2}}{k}, \tag{2.2a,b}$$

where $k_h = (k_x^2 + k_y^2)^{1/2}$ and $k = (k_h^2 + k_z^2)^{1/2}$. We denote the associated orthonormal eigenvectors by $\xi_{\mathbf{k}}^{(0)}$ and $\xi_{\mathbf{k}}^{(\pm)}$. We can group all physical variables (e.g. velocity components and buoyancy) in a dynamical state vector $\mathbf{X}_{\mathbf{k}}$. Then, the normal modes can be derived by projecting $\mathbf{X}_{\mathbf{k}}$ onto the set of eigenvectors

$$G_{\mathbf{k}} = \mathbf{X}_{\mathbf{k}} \cdot \overline{\xi_{\mathbf{k}}^{(0)}}, \quad A_{\mathbf{k}} = \mathbf{X}_{\mathbf{k}} \cdot \overline{\xi_{\mathbf{k}}^{(\pm)}}, \tag{2.3a,b}$$

where over-bar denotes the complex conjugate. We refer to the slow modes $G_{\mathbf{k}}$ as geostrophic (also known as rotational or vortical modes), and the fast modes $A_{\mathbf{k}}$ as ageostrophic (also known as gravitational or internal wave modes).

By applying the projections in (2.3) to the Fourier-space form of the linear terms of (2.1) and then non-dimensionalizing it, the evolution equations of the normal mode amplitudes follow as

$$\frac{\partial G_{\mathbf{k}}}{\partial t} = Ro \Phi_{\mathbf{k}}(G, A) + D_G, \tag{2.4a}$$

$$\frac{\partial A_{\mathbf{k}}}{\partial t} + i\sigma_{\mathbf{k}} A_{\mathbf{k}} = Ro \Psi_{\mathbf{k}}(G, A) + D_A. \tag{2.4b}$$

In (2.4), we also grouped all quadratic nonlinearities in terms of $\Phi_k(G, A)$ and $\Psi_k(G, A)$ for notational economy. They can be expanded as convolution sums. For instance

$$\Phi_k(G, A) = \sum_{k=p+q} [\phi_{GG}G_pG_q + \phi_{AG}A_pG_q + \phi_{AA}A_pA_q], \tag{2.5}$$

where the coefficients, ϕ , are functions of k, p, q and constant parameters defined in (1.1). A similar description holds for $\Psi_k(G, A)$.

Multiplying (2.4) by $\overline{G_k}$ and $\overline{A_k}$, the evolution of geostrophic and ageostrophic modal energy follows

$$\frac{\partial G_k \overline{G_k}}{\partial t} = Ro \Phi_k \overline{G_k} + D_G \overline{G_k} + \text{c.c.}, \tag{2.6a}$$

$$\frac{\partial A_k \overline{A_k}}{\partial t} = Ro \Psi_k \overline{A_k} + D_A \overline{A_k} + \text{c.c.}, \tag{2.6b}$$

where c.c. denotes the complex conjugate.

The nonlinear transfers in (2.6) can be summed as below to derive the horizontal energy fluxes

$$\Pi_G(k_h) = - \sum_{|k'-k' \cdot \hat{z}| < k_h} Ro \Phi_{k'} \cdot \overline{G_{k'}} + \text{c.c.}, \tag{2.7a}$$

$$\Pi_A(k_h) = - \sum_{|k'-k' \cdot \hat{z}| < k_h} Ro \Psi_{k'} \cdot \overline{A_{k'}} + \text{c.c.} \tag{2.7b}$$

3. Balance dynamics and nonlinear normal mode initialization

3.1. Time scales

From (2.2) there are two time scales in the flow when $Ro \rightarrow 0$. The time scale of G_k is L/U , which is determined by the nonlinear terms since $\lambda_k^{(0)} = 0$. The time scale of A_k , on the other hand, is derived as the inverse of the linear ageostrophic mode frequencies, $1/\sigma_k$. Therefore the ratio of the two is

$$\frac{t^*}{T} = \epsilon = \frac{1/\sigma_k}{L/U} \leq Ro, \tag{3.1}$$

where t^* and T denote the time scales of A_k and G_k , respectively.

This separation enables us to split the time derivative into two parts

$$\frac{\partial(\cdot)}{\partial t} = \frac{\partial(\cdot)}{\partial t^*} + \epsilon \frac{\partial(\cdot)}{\partial T}. \tag{3.2}$$

This two-time-scale method is a foundation of a nonlinear normal mode initialization (NNMI) scheme proposed by Baer & Tribbia (1977). More details on it can be found in classic texts on perturbation techniques, such as Cole (1968).

Rewriting (2.4) in terms of fast and slow time derivatives leads to

$$\epsilon \frac{\partial G_k}{\partial T} = \epsilon \Phi_k(G, A) + D_G, \tag{3.3a}$$

$$\frac{\partial A_k}{\partial t^*} + \epsilon \frac{\partial A_k}{\partial T} + i\sigma_k A_k = \epsilon \Psi_k(G, A) + D_A. \tag{3.3b}$$

The fast time derivatives of the geostrophic modes are zero by construction, as they are slowly varying.

A more advanced time-scale separation method was used by Reznik, Zeitlin & Ben Jelloul (2001) and Zeitlin *et al.* (2003) to develop the theory of nonlinear geostrophic adjustment. In these studies, instead of two time scales, the method of multiple time scales was used. More specifically, in addition to a fast time scale t^* , they used several slow time scales such that each one is $O(Ro)$ slower than the next. As will be seen in the following sections, their analytical results are compatible with the numerical simulations of the current study.

3.2. The Baer–Tribbia scheme

The essence of the Baer–Tribbia scheme is to ensure evolution on the slow time scale by forcing $\partial A_k / \partial t^*|_{t=0} = 0$ in (3.3b). This is done by expanding (3.3) in terms of ϵ , and then applying perturbation techniques up to the desired order. It is useful to introduce a new notation for variable expansion in powers of both T and ϵ as

$$X = (X^{0,0} + X^{0,1}T + X^{0,2}T^2) + (X^{1,0} + X^{1,1}T + X^{1,2}T^2)\epsilon + O(\epsilon^2, T^3), \tag{3.4}$$

where X can be any variable in (3.3). In other words, $X^{i,j}$ refers to terms proportional to ϵ^i and T^j .

After setting $\partial A_k / \partial t^*|_{t=0} = 0$, expanding both sides of (3.3) and equating each order in ϵ , one obtains

$$A_k^{0,0} = O(T), \tag{3.5a}$$

$$i\sigma_k A_k^{1,0} + i\sigma_k A_k^{1,1}T = \Psi_k^{0,0} + \Psi_k^{0,1}T + O(T^2), \tag{3.5b}$$

$$A_k^{1,1} + i\sigma_k A_k^{2,0} = \Psi_k^{1,0} + O(T). \tag{3.5c}$$

The time derivative of (3.5b) yields one more equation

$$i\sigma_k A_k^{1,1} = \Psi_k^{0,1} + O(T). \tag{3.6}$$

By setting $T = 0$ in (3.5) and (3.6), the initial ageostrophic modes are derived up to $O(\epsilon^2)$

$$A_k^{0,0} = 0, \tag{3.7a}$$

$$A_k^{1,0} = \frac{\Psi_k^{0,0}}{i\sigma_k}, \tag{3.7b}$$

$$A_k^{2,0} = \frac{\Psi_k^{1,0}}{i\sigma_k} - \frac{A_k^{1,1}}{i\sigma_k} = \frac{\Psi_k^{1,0}}{i\sigma_k} + \frac{\Psi_k^{0,1}}{\sigma_k^2}. \tag{3.7c}$$

Since our focus is on higher-order balance, it is necessary to keep terms up to $O(\epsilon^2)$ as a minimum requirement since QG theory is only one order lower. Equation (3.7a) shows that the ageostrophic modes are $O(Ro)$. Therefore, the scaling of ageostrophic energy goes as $O(Ro^2)$. With some algebra, the nonlinear terms in (3.7) can be

expressed in terms of the geostrophic modes as

$$\left. \begin{aligned} \Psi_k^{0,0} &= \sum_{k=p+q} \psi_{GG} G_p G_q, \\ \Psi_k^{1,0} &= \sum_{k=p+q} \frac{\psi_{GA}}{i\sigma_q} G_p \left(\sum_{q=m+n} \psi_{GG} G_m G_n \right), \\ \Psi_k^{0,1} &= 2 \sum_{k=p+q} \psi_{GG} G_p \left(\sum_{q=m+n} \phi_{GG} G_m G_n \right). \end{aligned} \right\} \quad (3.8)$$

Note that we only expanded the fast modes in terms of ϵ and kept geostrophic modes unchanged, as justified by Warn *et al.* (1995). The direct calculation of convolution sums in (3.8), especially nested sums, is too costly. To circumvent this problem we propose a novel algorithm that uses the pseudospectral method to calculate them. It is briefly described in appendix A. The dissipation terms can be kept in the procedure above. We performed initialization with and without viscosity and diffusion. The results were very similar, other than in the very small scales. This is to be expected, since balance is most apparent at larger scales where viscosity and diffusion are rather unimportant.

Knowing that the geostrophic spectrum admits a slope of -3 , one might try to analytically derive the slope of the balanced ageostrophic energy spectrum at least to first order. Thus, $A_k^{1,0} \cdot \overline{A_k^{1,0}}$ (over-bar denoting the complex conjugate) should be calculated from (3.7b) and (3.8)

$$\begin{aligned} A_k^{1,0} \cdot \overline{A_k^{1,0}} &= \frac{1}{\sigma_k^2} \Psi_k^{0,0} \cdot \overline{\Psi_k^{0,0}} = \frac{1}{\sigma_k^2} \sum_{k=p+q} \psi_{GG}(\mathbf{k}, \mathbf{p}, \mathbf{q}) G_p G_q \sum_{k=m+n} \overline{\psi_{GG}(\mathbf{k}, \mathbf{m}, \mathbf{n}) G_m G_n} \\ &= \sum_{k=p+q} \sum_{k=m+n} \frac{1}{\sigma_k^2} \psi_{GG}(\mathbf{k}, \mathbf{p}, \mathbf{q}) \overline{\psi_{GG}(\mathbf{k}, \mathbf{m}, \mathbf{n})} G_p G_q \overline{G_m G_n}. \end{aligned} \quad (3.9)$$

Knowing only that the spectrum formed by geostrophic modes scales as k^{-3} does not lead to a solution for the inertial range of the ageostrophic energy spectrum unless one calculates or approximates the coefficient $\psi_{GG}(\mathbf{k}, \mathbf{p}, \mathbf{q}) \overline{\psi_{GG}(\mathbf{k}, \mathbf{m}, \mathbf{n})} / \sigma_k^2$, which is a function of wavenumber, and then performs the summation. We therefore explore the slope of ageostrophic energy spectrum numerically and discuss it in §4.2.

4. Balancing procedure and numerical configuration

In this section, we develop a procedure to investigate the effect of initial Baer-Tribbia balance as well as the role of rotation and stratification in the transition from balance dynamics to more general smaller-scale turbulence. It is carried out in three steps: first, generation of the geostrophic data; second, finding balanced ageostrophic modes using NNMI; and finally, exploring the robustness of balanced initial conditions in the more general dynamical context of the non-hydrostatic Boussinesq simulations. Before describing each in more detail, we lay out the numerical configuration of our simulations.

Our Boussinesq model employs the de-aliased pseudospectral method with second-order time stepping. The state variables that are integrated in time are 3D vorticity and

buoyancy. We set the size of our domain to be 2π in the horizontal and $2\pi \times (f/N)$ in the vertical. In this way the aspect ratio, $\alpha = H/L$, is equal to f/N , which makes the domain a cube of $(2\pi)^3$ using Charney (1971) scaling. Throughout this paper we match stratification to rotation by fixing $N/f = L/H$. Therefore, the change in Ro is equivalent to a change in Fr , following classical QG scaling. For the sake of brevity we usually refer only to changes in rotation (or Ro), but the reader should keep in mind this implies changes to stratification as well.

In the real atmosphere the ratio of N to f is usually reported as 100, and in the ocean as 30–50. We also know that at least $O(10^2)$ grid points are required to resolve something more than just viscous and diffusional coupling of vertical layers. This vertical grid, together with high $L/H (= N/f)$, would require a large horizontal resolution, as explained in Bartello (2010). To maintain the vertical resolution high enough while keeping horizontal resolution within our computational resources, we reduced the ratio N/f to 8. This compromise distances us from exact atmospheric and oceanic scaling somewhat, but we have varied N/f and are certain that the features described below are generic, while recognizing the future need to explore this question with larger computers.

For the physical domain of $[0, 2\pi]^2 \times [0, 2\pi f/N]$, two different grids were employed:

- (i) Grid (1) $\Delta x = \Delta y = (N/f)\Delta z$. This grid is more loyal to QG structures since it is isotropic in Charney stretched coordinates.
- (ii) Grid (2) $\Delta x = \Delta y = \Delta z$. This grid is isotropic in unstretched physical space. Hence, it is felt to be a better choice to capture the smaller scales where the dynamics, and hence the aspect ratio, are still unknown.

If N_h and N_z are the number of horizontal and vertical grid points, respectively, grid (2) requires $N_z = (f/N)N_h$ vertical points, whereas the Charney grid requires $N_z = N_h$, which is far larger. Grid (2) therefore allows for a wider range of horizontal scales, for given computational resources, in addition to providing unbiased numerics for dynamics in the small-scale transition range.

We performed most of our simulations using grid (2) as we believe it is a better choice in analysing the breakdown of balance. Nevertheless, we performed several simulations using grid (1) and compared them. Table 1 shows a list of all simulation parameters used in this study.

The other parameters that need to be set in the Boussinesq model are the viscosity and diffusion. In addition to the Laplacian operator of the Boussinesq set, an iterated Laplacian (hyperviscosity) was often used. To be consistent with our cylindrical truncation of Fourier modes, we employed a cylindrical dissipation operator expressed as

$$D_u = D_b = \nu_h(-1)^{n+1}\nabla_h^{2n} + \nu_z(-1)^{n+1}\frac{\partial^{2n}}{\partial z^{2n}}, \quad (4.1)$$

which provides the same dissipation for buoyancy and velocity. In all our simulations on grid (2) we keep $\nu_h = \nu_z = \nu$. The Newtonian Laplacian operator is recovered by setting $n = 1$. Hyperviscosity, here set to $n = 4$, restricts dissipation to a narrower range of scales. For further details on the comparison between the two in rotating stratified turbulence, see Bartello, Métais & Lesieur (1996). Although the majority of our simulations used hyperviscosity, we also performed a number of sensitivity DNS runs in figure 14 below. Although there are quantitative differences, the observed trends are the same. All the times reported were normalized with the r.m.s. geostrophic vertical vorticity, τ , which is measured at the end of the preliminary QG runs and can be found in table 1.

Run	N_h	N_z	N	f	Ro_u	Ro_ω	Fr	τ	D_u	ν_h	ν_z
QGv18	1536	192	—	—	—	—	—	4.35	∇^8	5×10^{-18}	5×10^{-18}
QGv19	1536	192	—	—	—	—	—	5.28	∇^8	5×10^{-19}	5×10^{-19}
QGv20	1536	192	—	—	—	—	—	6.45	∇^8	5×10^{-20}	5×10^{-20}
QGg1	1024	1024	—	—	—	—	—	12.73	∇^8	5×10^{-18}	1.6×10^{-23}
QGg2	1024	128	—	—	—	—	—	6.23	∇^8	1×10^{-18}	1×10^{-18}
B768r005	768	96	861.97	107.75	0.005	0.044	0.015	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r01	768	96	430.98	53.87	0.01	0.089	0.030	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r03	768	96	143.66	17.96	0.03	0.269	0.091	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r05	768	96	86.20	10.77	0.05	0.453	0.160	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r07	768	96	61.57	7.70	0.07	0.646	0.244	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r09	768	96	47.89	5.99	0.09	0.852	0.351	4.8	∇^8	1×10^{-17}	1×10^{-17}
B768r11	768	96	39.18	4.90	0.11	1.078	0.491	4.8	∇^8	1×10^{-17}	1×10^{-17}
B1536r01lv	1536	192	492.05	61.50	0.01	0.105	0.036	6.4	∇^8	5×10^{-20}	5×10^{-20}
B1536r02lv	1536	192	246.02	30.75	0.02	0.211	0.075	6.4	∇^8	5×10^{-20}	5×10^{-20}
B1536r03lv	1536	192	164.02	20.50	0.03	0.320	0.118	6.4	∇^8	5×10^{-20}	5×10^{-20}
B1536r07lv	1536	192	70.29	8.79	0.07	0.817	0.403	6.4	∇^8	5×10^{-20}	5×10^{-20}
B1536r01	1536	192	407.52	50.94	0.01	0.104	0.036	5.3	∇^8	5×10^{-19}	5×10^{-19}
B1536r03	1536	192	135.84	16.98	0.03	0.314	0.112	5.3	∇^8	5×10^{-19}	5×10^{-19}
B1536r05	1536	192	81.50	10.18	0.05	0.553	0.204	5.3	∇^8	5×10^{-19}	5×10^{-19}
B1536r07	1536	192	58.22	7.28	0.07	0.771	0.330	5.3	∇^8	5×10^{-19}	5×10^{-19}
B1536r09	1536	192	45.28	5.66	0.09	1.041	0.511	5.3	∇^8	5×10^{-19}	5×10^{-19}
B2048r01	2048	256	456.91	57.11	0.01	0.112	0.04	6.4	∇^8	5×10^{-21}	5×10^{-21}
B2048r03	2048	256	152.30	19.04	0.03	0.343	0.131	6.4	∇^8	5×10^{-21}	5×10^{-21}
B2048r05	2048	256	91.38	11.42	0.05	0.596	0.262	6.4	∇^8	5×10^{-21}	5×10^{-21}
B2048r07	2048	256	65.27	8.16	0.07	0.896	0.474	6.4	∇^8	5×10^{-21}	5×10^{-21}
B2048r09	2048	256	50.77	6.35	0.09	1.278	0.809	6.4	∇^8	5×10^{-21}	5×10^{-21}
B2048r01n	2048	256	276.26	34.53	0.01	0.084	0.024	3.0	∇^2	5×10^{-5}	5×10^{-5}
B2048r05n	2048	256	55.25	6.91	0.05	0.425	0.124	3.0	∇^2	5×10^{-5}	5×10^{-5}
B2048r09n	2048	256	30.70	3.84	0.09	0.781	0.248	3.0	∇^2	5×10^{-5}	5×10^{-5}
B2048r11n	2048	256	25.11	3.14	0.11	0.970	0.327	3.0	∇^2	5×10^{-5}	5×10^{-5}

TABLE 1. The runs starting with QG are based on the quasigeostrophic equations and the rest are based on the non-hydrostatic Boussinesq equations. The values of Ro_u , Ro_ω , Fr and τ in the Boussinesq runs are calculated at $t = 0$.

4.1. Generation of geostrophic data

In this step we run a decaying QG model to produce the geostrophic modes that are later used in our balancing scheme. Our QG model is obtained from the Boussinesq model by setting the ageostrophic modes to zero at each time step. Although this is not efficient, it suffices for the generation of our initial data. As shown in figure 1, our initial energy distribution peaks at $k_i = 20$ to provide a wavenumber range for the considerable upscale transfer of energy. Then the model is run until $t = 5\tau$, based on the initial r.m.s. vorticity. As time evolves, the spectrum fills out quickly. Subsequent changes to its shape occur much more slowly.

According to Charney (1971), QG flow is mathematically analogous to 2D isotropic turbulence, where the logarithmic slope of the energy spectrum in the potential-*enstrophy* cascade range is predicted to be -3 . Our slope, not surprisingly, is steeper than this, as in previous studies at comparable resolutions, arguably due to emerging coherent structures (cf. McWilliams, Weiss & Yavneh 1994). To avoid any

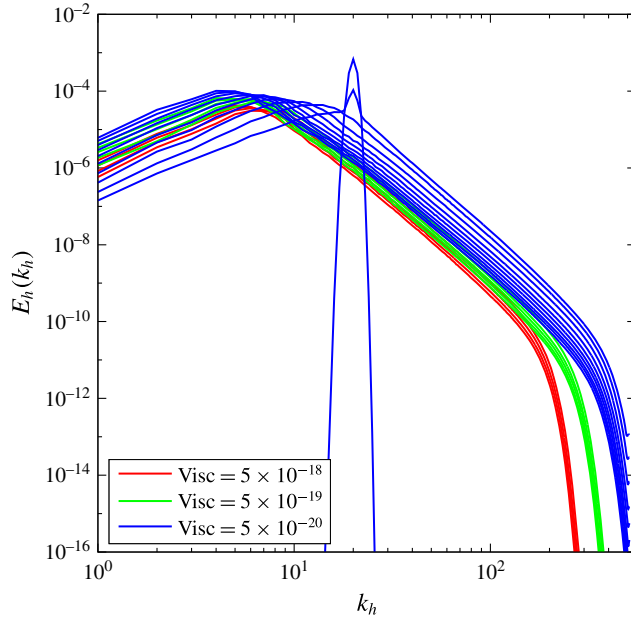


FIGURE 1. (Colour online) The spectra of total energy for QG simulations with different viscosities: runs listed as QGv18, QGv19 and QGv20 in table 1. For QGv20, the spectra are plotted from $t=0$ to 5τ . For the other two simulations only $t \in [3\tau, 5\tau]$ is plotted.

rapid small-scale adjustment when inputting these QG modes into the non-hydrostatic Boussinesq model, we decided to perform separate preliminary QG simulations for all values of (hyper)viscosity used below.

As explained in § 4, two types of grid were employed; grid (1) is isotropic in the Charney stretched coordinate and grid (2) is isotropic in unstretched real coordinates. Figure 2 shows the horizontal and vertical spectra of the total energy using grid (1) with 1024^3 collocation points (Run QGg1 in table 1) and grid (2) at $1024^2 \times 128$ (Run QGg2). The latter had the same hyperviscosity in all directions, whereas in the former the vertical viscosity was reduced by a factor $(f/N)^8$ to keep νk_{max}^8 the same in each direction, since we used a ∇^8 operator.

Figure 2 shows the horizontal and vertical energy spectra at $t=4\tau$ using $1024^2 \times 128$ points on grid (2) and 1024^3 points on grid (1). The slope of the horizontal spectrum is similar at large horizontal wavenumbers. Not surprisingly the spectrum corresponding to grid (2) is lower, since the vertical dissipation is higher. The vertical spectra of grid (1) extends to much higher vertical wavenumbers. Nonetheless, they are reasonably parallel at low k_z . Acknowledging that grid (1) resolves QG structures isotropic in Charney stretched coordinates better, figure 2 shows that the results of grid (2) are reliable as well for the QG flow, while at the same time not biasing the small-scale dynamics emerging from the shallow spectra discussed below.

4.2. Generation of balanced ageostrophic modes using NNMI

Once geostrophic data are generated, they can be inserted into the balance scheme to derive the ageostrophic modes that ensure initial evolution on the slow time scale. Our high-order scheme and its numerical implementation are described in § 3 and

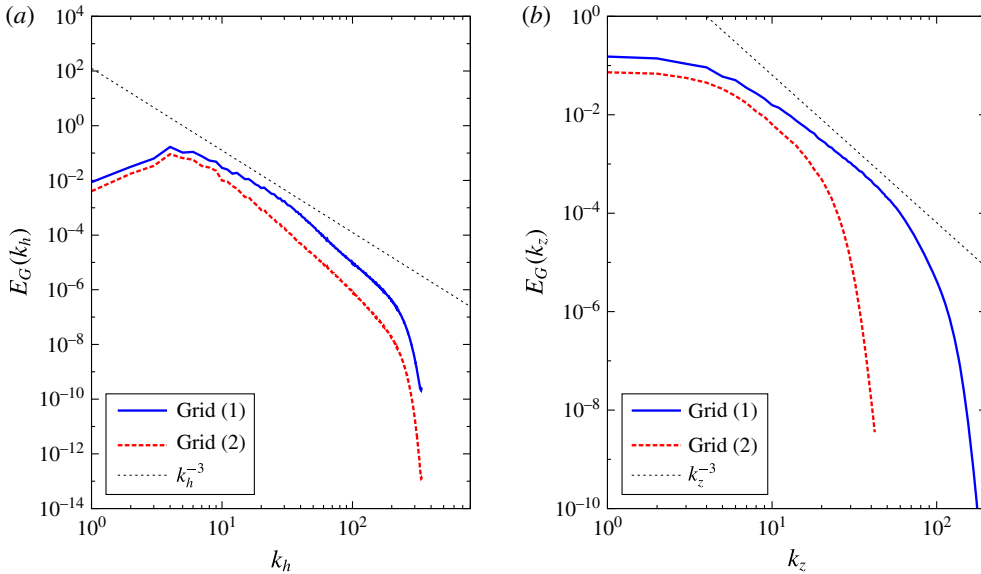


FIGURE 2. (Colour online) Horizontal (a) and vertical (b) spectra of total energy for QG simulations for the two different grids.

appendix A, respectively. The initialized data depend on Ro (and consequently Fr), since the frequency, σ_k , in (2.2) depends on f and N . There are different ways to calculate Ro in fully developed turbulence. Following its definition in (1.1), it can either be derived based on mean-square vorticity or velocity. The vorticity-based Rossby number is formulated as

$$Ro_\omega = \frac{\langle (\nabla \times \mathbf{u}(x, y, z) \cdot \hat{\mathbf{z}})^2 \rangle^{1/2}}{f}, \tag{4.2}$$

whereas the velocity-based Rossby number is given by

$$Ro_u = \frac{\langle \mathbf{u} \cdot \mathbf{u} \rangle^{1/2}}{fL}. \tag{4.3}$$

In both definitions, $\langle \rangle$ represents a spatial average. In (4.3), L is the characteristic length scale, plausibly calculated by inverting the wavenumber at which the energy spectrum peaks.

The vorticity-based Rossby number is more affected by the dynamics at small scales than its velocity-based counterpart. Since we want to classify different rotating regimes based on their large-scale QG-like motion, we favour the velocity-based definition and drop the subscript. As shown below, some of our simulations show a transition to a more shallow spectrum at larger wavenumbers. While this affects the root mean-square vorticity, and hence Ro_ω , its influence on Ro_u is much weaker. In addition, the slope of this shallow range shows some dependence on model dissipation, which further justifies our choice of velocity-based Rossby number. Note that, even in the QG model, the vorticity-based Rossby number is approximately one order of magnitude larger than Ro_u .

Ro_u	Ro_ω	Ro_{max}
0.01	0.10	0.81
0.03	0.31	2.016
0.09	1.04	16.49

TABLE 2. Different Rossby numbers at $t=0$ on a grid of $1536^2 \times 192$.

A third definition of Rossby number is based on the extremal value of vertical vorticity in real space, $Ro_{max} = |\nabla \times \mathbf{u}(x, y, z) \cdot \hat{\mathbf{z}}|_{max}/f$. This definition has been used frequently by Dritschel and co-workers (see for instance Dritschel & Viúdez (2003), Viúdez & Dritschel (2004) and Dritschel & McKiver (2015)). This maximum Ro is far greater than Ro_ω since the vertical vorticity at large Reynolds numbers displays a close-to-exponential distribution. We derived Ro_{max} , Ro_ω and Ro_u after initialization and compared them in table 2. This considerable difference should be taken into account when values are compared to Ro_{max} and Ro_ω employed in other studies. It may be argued that values of Ro_{max} in excess of unity are unrealistically large. However, Hakim (2000) and Hoskins & Hodges (2002) have shown this not to be the case. In fact, the former study shows relative vorticity values in excess of f in 40% of the 500 hPa weather charts considered.

To investigate the effect of dissipation and rotation separately, we keep Ro constant for different viscosities. In doing so a slight change of the Coriolis parameter (and consequently a change of N) is necessary, as the peak of the energy spectra, L , and the total energy, U^2 , in the initial data resulting from the QG run are slightly different for different viscosities.

To describe the initialization we investigate the ageostrophic modal spectrum that the Baer–Tribbia scheme produces. We begin by taking the slow modes from an output of a QG simulation. While keeping their complex Fourier phases constant, we scale their amplitudes to yield geostrophic modal spectra with various slopes. After implementing the initialization scheme on each set of these data, we produce the spectrum of the balanced ageostrophic modes. It is, of course, also interesting to see how the balanced output changes at different Rossby numbers. Figure 3 presents the slope of the Baer–Tribbia ageostrophic energy (E_A) spectrum versus the slope of the input geostrophic spectrum (E_G).

The real quasigeostrophic data, derived from a decaying QG simulation, have a slope of -4.2 , as seen in figure 1, whose value is marked with a large circle. It is interesting that for the true QG data, all Rossby numbers have almost the same ageostrophic spectral slope, which is -6.8 . As we shall see, in all our simulations at different Ro , viscosities and resolution, the slope of the balanced ageostrophic spectrum is between -6.5 and -7 , that is, much steeper than the slope of the geostrophic spectrum. We see similar results at other geostrophic slopes as well since the entire curve falls below the $y = x$ line, but as we vary the geostrophic spectral slope further from the circled value, the ageostrophic slope becomes a function of the Rossby number. Instead of using our QG simulation complex phases, we followed the same procedure with completely random phases and obtained qualitatively similar results. The conclusion is simply that balanced turbulence occurs in conjunction with rather steep energy spectra, as the total energy spectrum follows that of the less steep geostrophic modal spectrum at small scales. As such, balanced rotating stratified turbulence cannot explain the shallow $-5/3$ range in subdeformation scale atmospheric and oceanic data.

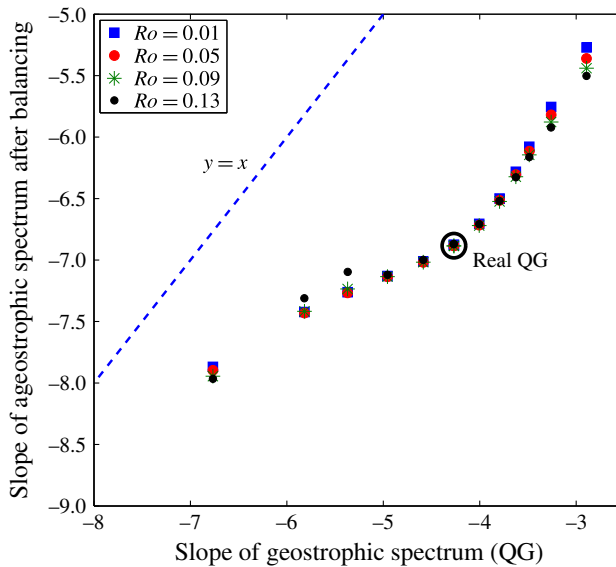


FIGURE 3. (Colour online) The logarithmic slope of the ageostrophic energy spectrum after the Baer–Tribbia balancing scheme as a function of the geostrophic slope. The resolution was $1536^2 \times 192$, $\nu = 5 \times 10^{-20}$.

4.3. Boussinesq simulations

We ran the non-hydrostatic Boussinesq model using the geostrophic modes from the QG model and the balanced ageostrophic data derived in the NNMI step as an initial condition in a more general dynamical setting. Our goal is to establish the robustness of this sort of balance within a non-hydrostatic Boussinesq framework. It is, of course, well known that this will depend on the Rossby number. We therefore calculated separate ageostrophic initial conditions for a variety of Ro . The following section presents the results of these simulations.

5. Results

Our aim is first to examine the effect of balance by comparing initially balanced simulations with their unbalanced counterparts. The unbalanced initial data were generated by taking the output ageostrophic modes of the Baer–Tribbia data and scrambling their Fourier phases. In this way the comparison is fair, since the balanced and unbalanced ageostrophic modes share the same amplitude. Hence, all simulations of §§ 5.1–5.3 start with the same E_G , and for each Rossby number, initially balanced and unbalanced simulations have the same E_A and indeed the same spectrum. In fact, the energetics of the two simulations are identical. After generating a set of balanced and unbalanced data, we examine:

- (i) the time series of their spatially averaged quantities such as energy;
- (ii) the evolution of their energy spectra in time; and
- (iii) the frequency spectrum of a flow variable at a grid point in physical space.

5.1. The time series of initially balanced and unbalanced flow

In figure 4, the total energy, which is the sum of the geostrophic and ageostrophic energy, is depicted. At low Rossby number the energy stays relatively more constant in

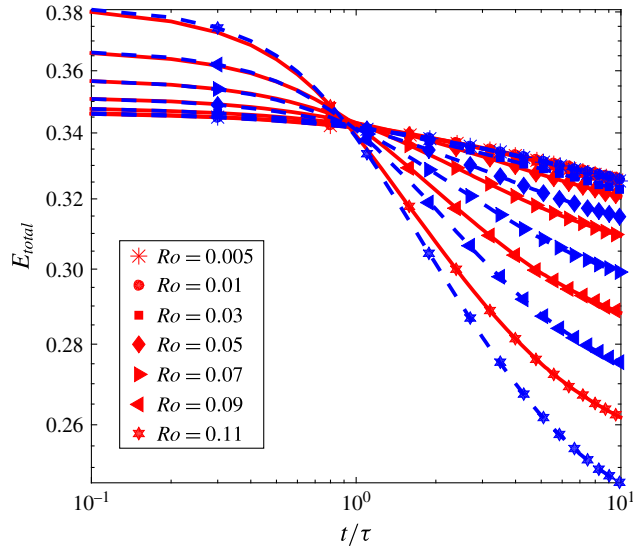


FIGURE 4. (Colour online) The time series of the total energy. The solid lines are the initially balanced flow, dashed lines are the initially phase-scrambled flow. The plots correspond to runs starting with B768 listed in table 1.

time, signalling that the dominant direction of transfer is towards larger scales. As Ro increases, more forward transfer takes place and more energy is dissipated. Comparing the initially balanced and unbalanced flows shows that, at strong rotations, both have similar levels of energy. However, as Ro increases, more energy is dissipated when initial conditions are unbalanced, implying that most of the forward cascade is due to the ageostrophic modes.

Figure 5 shows the total geostrophic and ageostrophic energy of the flow as a function of time. Noting the scale of the vertical axis of both panels, one can conclude that the decay of ageostrophic energy is much larger than that of geostrophic energy. This is consistent with an inverse cascade for the geostrophic modes and a forward cascade for ageostrophic modes, as described for initially unbalanced turbulence by Bartello (1995).

Comparing the two Rossby numbers in figure 5(a) one finds that E_G of the higher Rossby number decays faster than that of the lower Rossby number. Since the dominant transfer of geostrophic energy by geostrophic modes alone is towards larger scales, the faster geostrophic decay at higher Ro is presumably due to stronger interaction with ageostrophic modes. This is in accordance with the ageostrophic time series of these Rossby numbers in figure 5(b). The higher Rossby number has higher E_A , and it decays faster.

Both initially phase-scrambled and balanced flows have the same level of geostrophic and ageostrophic energy at $t = 0$. This is not visible on the logarithmic horizontal axis of figure 5. At $Ro = 0.01$, the geostrophic energies of the initially balanced and unbalanced flow stay very close to each other during the entire integration time. The ageostrophic energy of the initially unbalanced flow, however, is larger at this Ro . This shows that the ageostrophic modes of the unbalanced flow take more energy from geostrophic modes and dissipate it at small scales. In spite of this, the total geostrophic energy is not much affected, since E_A is several

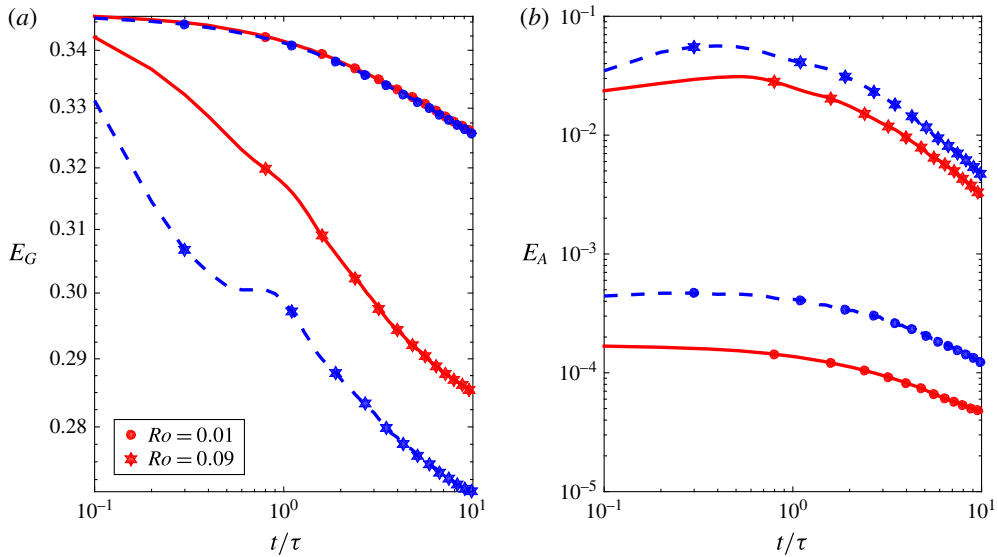


FIGURE 5. (Colour online) (a) Geostrophic energy versus time. (b) Ageostrophic energy versus time. The solid lines are the time series of the initially balanced flow and the dashed lines are the initially phase-scrambled flow. The simulations correspond to B768r01 and B768r09.

orders smaller than E_G at $Ro = 0.01$. Unlike the low Rossby number, at the higher $Ro = 0.09$, geostrophic energies exhibit a clear difference. Compared to $Ro = 0.01$, E_A is several orders larger at $Ro = 0.09$, implying the interaction between geostrophic and ageostrophic modes is much stronger.

Figure 6 illustrates how balance dynamics breaks down with time in different rotating regimes. In this figure, the total ageostrophic energy is depicted versus the total geostrophic energy on logarithmic axes. Each solid curve with identical marker shapes connects different Ro at the time given in the legend. The dashed lines connect the same Rossby numbers at different times. At $t = 0$, the curve is a vertical line, as we started all our simulations with the same amount of geostrophic energy. However, for different Ro the initial ageostrophic energy produced by the Baer–Tribbia scheme is different. More specifically, (3.7a) shows that $A_k \sim O(Ro)$, since $A_k^{0,0} = 0$. Therefore, the ageostrophic energy produced by the Baer–Tribbia scheme scales as Ro^2 .

As time increases, the dynamics fall into two distinct regimes. First, one which is balanced and remains as vertical lines. Since rotation is strong in this regime, the dominant transfer is an inverse cascade. Hence, the geostrophic energy stays similar for different Ro , but the balanced ageostrophic energy changes. The second regime occurs at weaker rotation and behaves as horizontal lines. The marked feature of this regime is that different Rossby numbers seem to approach the same level of ageostrophic energy, a characteristic we could not infer from other figures. The geostrophic energies at different Rossby numbers are, as expected, different as Ro is increased, since more of it is transferred to ageostrophic modes, where it is then dissipated via a forward cascade. The other notable feature of this figure is that the transition zone between these two dynamics gets sharper with time.

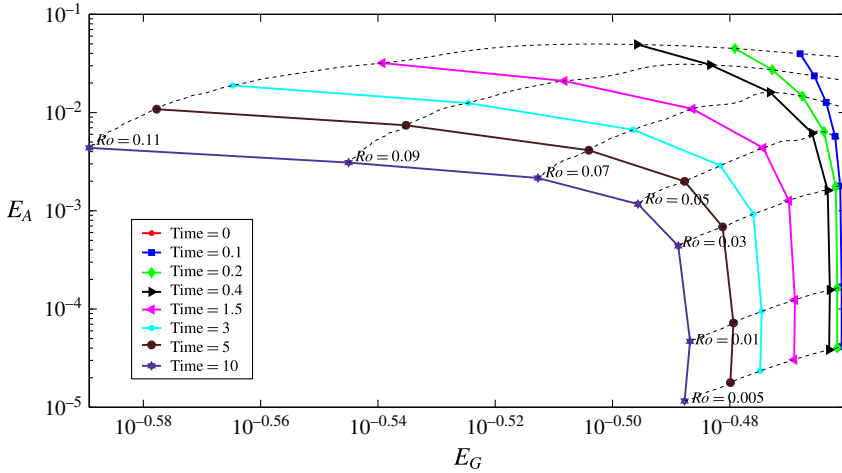


FIGURE 6. (Colour online) The total ageostrophic energy versus total geostrophic energy at different times for different Ro . Solid lines connect different Ro at the same time, while dashed lines do the reverse. The plots correspond to runs starting with B768 listed in table 1.

5.2. *The energy cascade of initially balanced and unbalanced flows*

The energy cascades of fast and slow modes can be the simplest indicator of balance. In unforced simulations we speculate that the ageostrophic modal spectrum is a power law whose amplitude decays with time (Bartello 1995). Figure 7 shows the evolution of initially balanced and unbalanced states at $t = 10\tau$. As explained previously, the initially unbalanced data is generated by scrambling the phases of the ageostrophic modes, implying their initial spectra coincide.

The geostrophic spectrum of both balanced and unbalanced flow evolves similarly in time. However, the ageostrophic part of the energy spectrum is considerably different for initially balanced and unbalanced data. For $Ro = 0.03$ (figure 7a), this difference both in terms of value and slope is drastic, whereas for $Ro = 0.09$ (figure 7b) it is much less. In the lower Rossby number case, the slope of the ageostrophic balanced spectrum is much steeper than its unbalanced counterpart, showing that the initialization scheme prevented the excitation of wave modes. In the higher Ro simulation, the growth of balanced ageostrophic modes is more similar to unbalanced modes. This is expected, as the Baer–Tribbia scheme uses perturbation techniques based on expansions in terms of Ro . As it increases, convergence is less rapid, presumably degrading the quality of the balance.

5.3. *The frequency spectrum of initially balanced and unbalanced flows*

The role of the initial balancing is perhaps more directly observed in the frequencies of the flow variables. To derive these in the current statistically homogeneous geometry we simply take the Fourier transform with respect to time of a flow variable at a point in our physical-space collocation grid. The choice of variable does not affect the result as long as both geostrophic and ageostrophic modes contribute to it. Here, we arbitrarily chose the x component of velocity. Figure 8 shows its time series and frequency spectra for both initially balanced and unbalanced conditions in the simulation denoted B768r03. High-frequency fluctuations make the evolution of

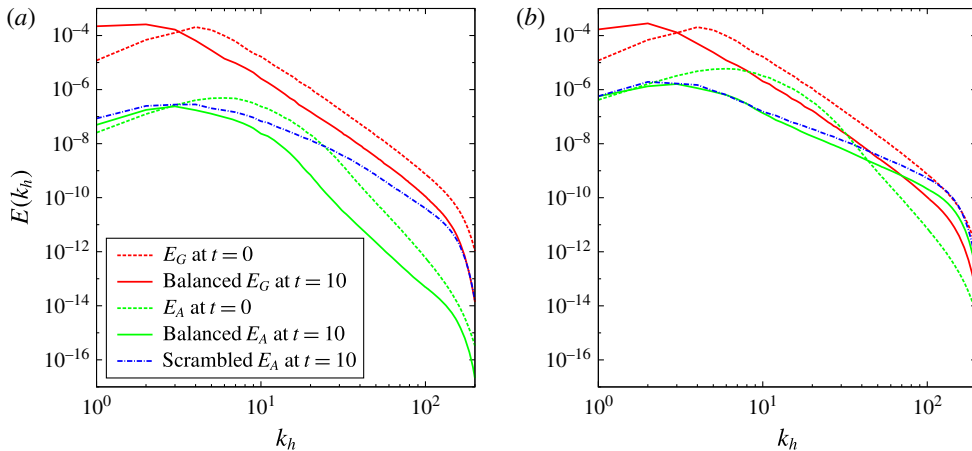


FIGURE 7. (Colour online) The upper dashed curve corresponds to the geostrophic spectrum at $t=0$, and the lower dashed curve corresponds to the ageostrophic spectrum at $t=0$. The solid curves represent the spectrum of the initially balanced data at $t=10\tau$, where the geostrophic spectrum is above the ageostrophic curve. The dash-dotted line is the ageostrophic spectrum of initially phase-scrambled data. The corresponding simulations are B768r03 and B768r09. (a) $Ro=0.03$. (b) $Ro=0.09$.

unbalanced initial data clearly different from their balanced counterpart. However, the slow part of the flow evolves similarly at early times. Later, even the slowly varying part of the motion displays differences between initially balanced and unbalanced simulations.

To take the Fourier transform we chose a time interval that satisfies two conditions. First, the beginning is after $t=\tau$, discarding early non-stationary effects. Second, the function at its two ends had approximately the same value. The interval is marked by two thick vertical lines in figure 8(a). Its Fourier transform is portrayed on figure 8(b).

Equation (2.2) shows that the linear frequencies of the ageostrophic modes lie between the Coriolis parameter, f , and the Brunt–Väisälä frequency, N . These parameters are shown with vertical lines in figure 8(b). Unsurprisingly, in this range the amplitude of balanced flow is less than that of the unbalanced flow.

To understand the effect of initial balancing in different rotating regimes better, the frequency spectra at four different values of Ro are shown in figure 9. Similar to the those in figure 8, the most significant feature is the difference between the amplitudes of balanced and unbalanced flows in the band of inertia-gravity frequencies. Here we see a jump in amplitude of the unbalanced system at low Rossby numbers. The jump diminishes as the Rossby number increases. Note that the wave band moves to the left as Ro increases, since we decrease both N and f . However, the logarithmic width of the band does not change, as we keep N/f constant. Not surprisingly, we see that reducing the frequency disparity between fast and slow modes reduces the difference in amplitude between initially balanced and unbalanced simulations, which makes the interaction between geostrophic and ageostrophic modes stronger.

‘Subinertial’ frequencies (less than f) correspond to the slow part of the dynamics. At low Rossby numbers it is almost the same for balanced and unbalanced flow, indicating they are dominated by the geostrophic modes, which are only weakly coupled to the ageostrophic modes. As Ro is increased, even the slowly varying part

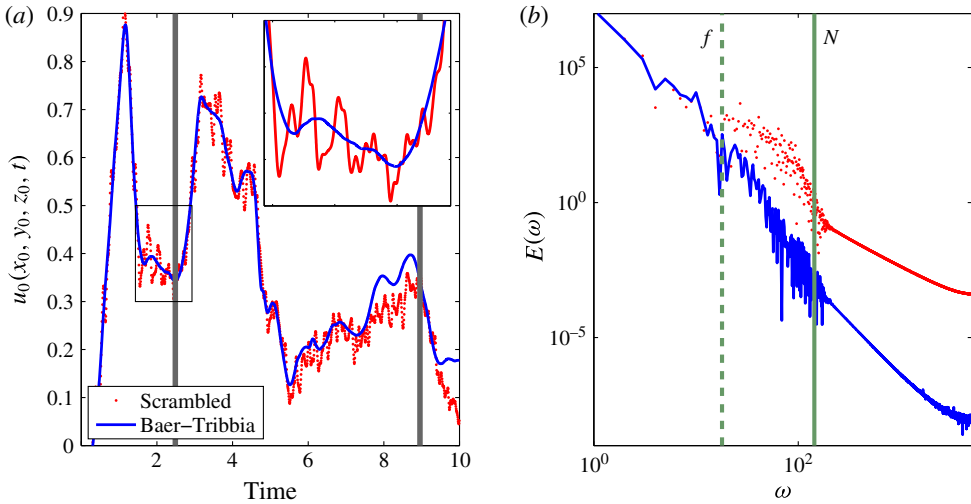


FIGURE 8. (Colour online) (a) The time series of $u_0(x_0, y_0, z_0, t) = u_0(t)$. The thicker smooth line shows the evolution of the balanced condition in time. The thinner fluctuating line shows the evolution of initially phase-scrambled data. (b) The Fourier transform of $u_0(t)$ with respect to time, for initially unbalanced and balanced conditions. The dashed vertical line represents the Coriolis parameter, f , and the solid vertical line shows the Brunt–Väisälä frequency, N .

of the flow shows differences, a fact that was also seen in the time series of figure 8. This, together with decreasing differences in amplitude of the wave band, makes fast and slow dynamics less distinct in the high- Ro regime. Although this behaviour was expected and predicted by previous studies employing two-time-scale approaches, here we can quantify it.

5.4. The role of wave frequency in the breakdown of balance

After studying the difference between the dynamics of initially balanced and unbalanced flow, henceforth we focus on initially balanced flow and try to analyse its breakdown as a function of Ro and Fr , while also exploring sensitivity to the dissipation.

To investigate the role of rotation, in figure 10 we vary Ro . In each panel the geostrophic and ageostrophic parts of the energy are portrayed together. Since the large-scale dynamics are close to QG at low Rossby numbers, E_G spectra are well above E_A . However, this difference is reduced as rotation becomes weaker. Even in the initial condition (dashed lines) the difference between the two spectra tapers off with increasing Rossby number, which is consistent with the scaling of ageostrophic energy ($E_A \sim Ro^2$).

For all the plots in figure 10, the slope of E_A is steeper than E_G at $t=0$, as expected from figure 3. As time increases, the slope of E_G stays relatively insensitive to Ro at this resolution. By contrast, the ageostrophic slope changes substantially with Ro number at later times.

When rotation is as strong as $Ro=0.01$ in figure 10(a) the slope of the ageostrophic spectrum remains steeper than the geostrophic slope. This is consistent with the maintenance of balance, as unbalanced wave modes stay weaker than the higher-order

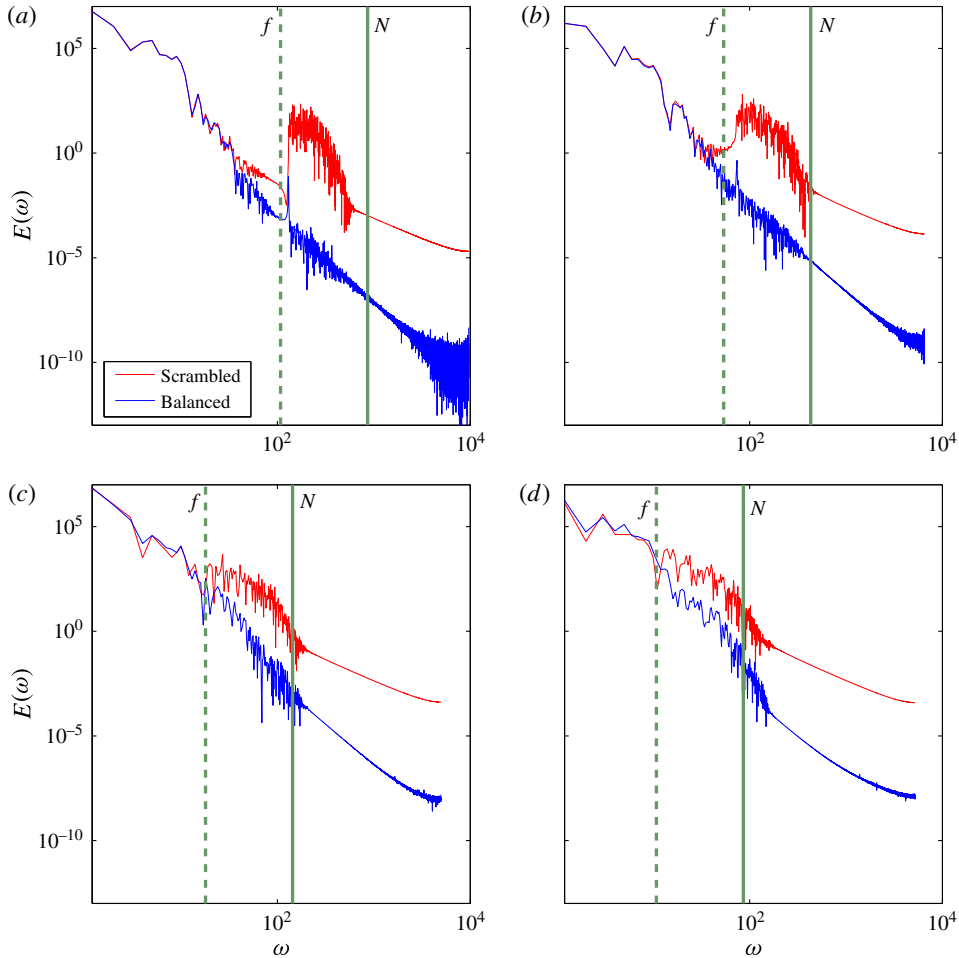


FIGURE 9. (Colour online) The time spectrum of $u(x_0, y_0, z_0, t)$ for four different Rossby numbers in the runs of B768r005, B768r01, B768r03 and B768r05 (a) $Ro = 0.005$. (b) $Ro = 0.01$. (c) $Ro = 0.03$. (d) $Ro = 0.05$. The dashed vertical line represents the Coriolis parameter, f , and the solid vertical line shows Brunt–Väisälä frequency, N . The time spectrum of phase-scrambled initial condition is above the balanced initial condition in Rossby numbers.

corrections to geostrophic flow. As Ro increases, the ageostrophic slope quickly becomes more shallow. This can lead to a crossing of the two spectra when rotation is weak enough. This crossover may suggest strong interaction between geostrophic and ageostrophic modes. For instance, when Ro is increased to 0.07 in figure 10(c), E_A becomes larger than E_G above a certain wavenumber. However, these results are derived using grid (2) defined in § 4; we performed similar simulations on grid (1) and found that imbalance grew even more quickly. This confirms that the generation of balance is a real feature of the flow, and is not imposed by our choice of numerical grid.

The geostrophic energy at a higher resolution (or equivalently higher Re) and at an earlier time is presented in figure 11. Unlike figure 10(c), the geostrophic spectrum develops a shallow tail. The geostrophic shallow range, however, always

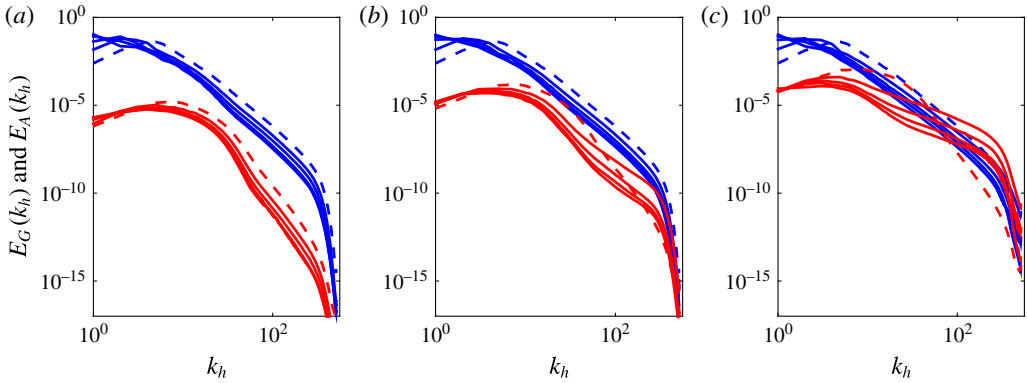


FIGURE 10. (Colour online) The geostrophic (top curves) and ageostrophics (lower curves) contributions to the energy spectrum for different Rossby numbers and at different times. (a) $Ro = 0.01$. (b) $Ro = 0.03$. (c) $Ro = 0.07$. The dashed lines correspond to $t = 0$, i.e. after Baer–Tribbia initialization and before Boussinesq simulation. The solid curves represent later times, namely $t = 5\tau$, $t = 10\tau$, $t = 15\tau$ and $t = 20\tau$. The corresponding runs are B1536r01lv, B1536r03lv and B1536r07lv.

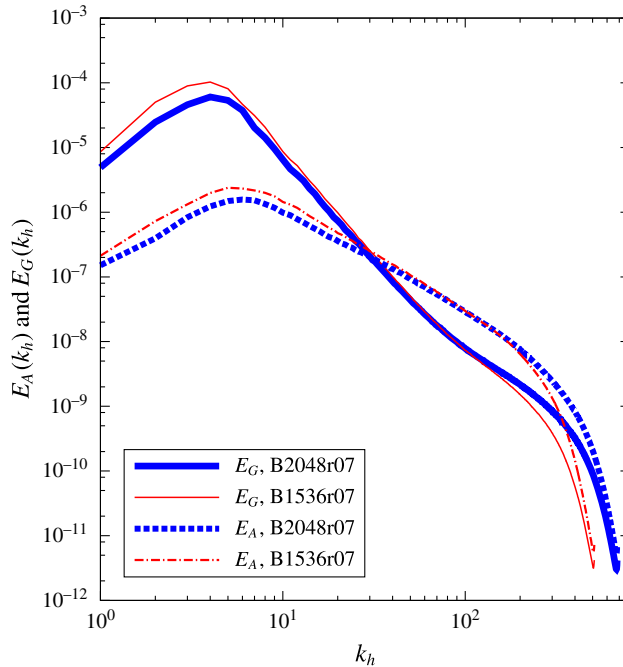


FIGURE 11. (Colour online) E_G and E_A spectra at $t = \tau$ for B1536r07 and B2048r07, whose initial Rossby number was 0.07.

occurs at wavenumbers larger than the onset of the ageostrophic shallow spectrum. If the resolved wavenumber range is sufficient, we expect both E_G and E_A to exhibit parallel slopes of $-5/3$ at very small scales, as in Bartello (2010). The E_A spectrum in this range is twice as large as E_G in amplitude since there are two ageostrophic

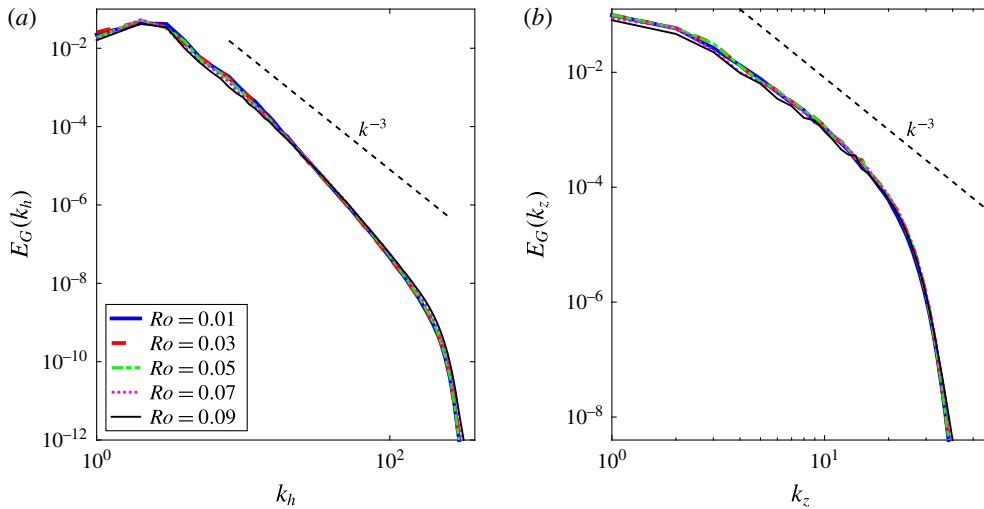


FIGURE 12. (Colour online) (a) Horizontally integrated E_G as a function of vertical wavenumber, for different Ro . (b) Vertically integrated E_G as a function of horizontal wavenumber, for different Ro . All the spectra are derived at $t = 10\tau$. Dashed lines in both plots have the slope of -3 . The plots correspond to B1536r01, B1536r03, B1536r05, B1536r07 and B1536r09.

modes corresponding to $\pm\sigma_k$ (cf. (2.2) and (2.3)). At these small scales the turbulence becomes more 3D and closer to isotropic, with our linear decomposition losing its meaning. Similar results were reported by Deusebio *et al.* (2013) in unbalanced turbulence. Even in the analysis of atmospheric data, Callies, Ferrari & Bühler (2014) observed that the tail of the geostrophic energy spectrum flattens out at smaller scales. They expressed that this flattening ‘is likely an artifact, because at these scales the geostrophic component makes up a small fraction of the observed signal’. It should also be noted that in our simulations the shallow tails decay in amplitude and are eventually suppressed as time advances, the turbulence decays, and Ro decreases.

Figure 12 portrays the geostrophic part of the energy spectrum for different rotations. In figure 12(a) horizontal spectra are shown, while figure 12(b) displays vertical spectra. It reveals that the geostrophic energy at different Ro displays very similar spectra over this range. However, it was shown in figure 11 that, at larger Re and at early times, the geostrophic spectrum can develop a shallow tail. The slope of the vertical geostrophic spectra is very close to -3 , as predicted by the potential-entrophy cascade phenomenology, whereas the slope of the horizontal spectra is somewhat steeper. As discussed in §4.1, the steeper spectra are often attributed to coherent vorticity structures.

Unlike $E_G(k_h)$, $E_A(k_h)$ is highly affected by the strength of rotation and stratification. Figure 13 shows this dependence. Figure 13(a) presents E_A spectra for different Rossby numbers. They are scaled by the instantaneous $Ro(t)^2$, as justified by (3.7a). The utility of this scaling appears in the ageostrophic energy spectra of large scales in figure 13, where those at different Rossby numbers collapse to the same curve. Clearly, the balanced curve is the steep envelope, with individual simulations departing from it at wavenumbers that decrease as Ro increases.

Similar to the observations made of figure 10, in figure 13 the spectral slopes at this fixed time become shallower as Ro increases. This difference can better be

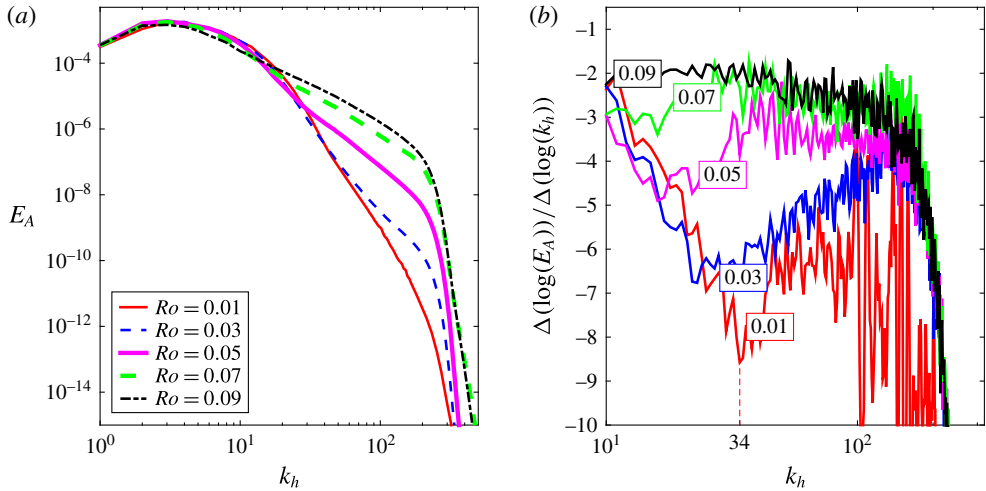


FIGURE 13. (Colour online) (a) E_A spectrum at different Ro . (b) The slope of E_A corresponding to (a). The same simulations as those of figure 12 were used.

seen in figure 13(b), where the slopes of the ageostrophic spectra are depicted. The slopes are derived by calculating $\Delta(\log(E_A))/\Delta(\log(k_h))$ (Δ denoting forward Euler differencing). The flat range in figure 10(b) corresponds to the power-law range of spectra in figure 10(a), and the local minima portray the kinks in the ageostrophic spectra.

At the lowest Rossby number, $Ro = 0.01$, there is a clear power-law range with a slope between $[-7, -6]$, similar to the results of figure 3. At low Rossby numbers, there is a well-pronounced local minimum that indicates the kink in the spectrum. Another distinguishing feature of the $Ro = 0.01$ simulation is the fluctuation in slopes, which are greater than simulations at higher Rossby numbers. To understand their cause we should recall that, in figure 13(a), E_A is scaled with Ro^2 . The unscaled spectrum of $Ro = 0.01$ is below those of the higher Rossby numbers and its slope is steeper. Therefore, the ageostrophic energy is very small at large wavenumbers, where its value is close to machine round-off error. This is corroborated by the fact that the slope gets noisier as we go to smaller scales. As Ro is increased the kink moves to larger scales. This is best explained by the onset of spontaneous imbalance at larger scales at weaker rotations. When rotation is weak enough, the kink disappears and the slope asymptotes to a constant that is close to -2 . Other simulations with lower hyperviscosity coefficients have more shallow slopes, approaching $-5/3$, but whose dissipation was not judged to be sufficiently well resolved to be presented here.

In the limits of $Ro \rightarrow 0$ and $Ro \rightarrow \infty$, geostrophic and 3D turbulence are recovered, respectively. In these regimes, there are inertial ranges with universal slopes. The appearance of a power-law range in the spectra of figure 10 gives rise to the question of whether there is a similar universal inertial range. To demonstrate this it would be necessary to show that its slope is not affected by small-scale dissipation. In fact, for the resolutions and Ro considered in this study, there is a dependence on the viscosity and diffusion coefficients, even outside of the dissipation range, and especially for midrange Rossby numbers. Despite this fact, the general characteristics of figure 13, such as the flattening of spectra at increased Ro , were observed at four other hyperviscosity coefficients.

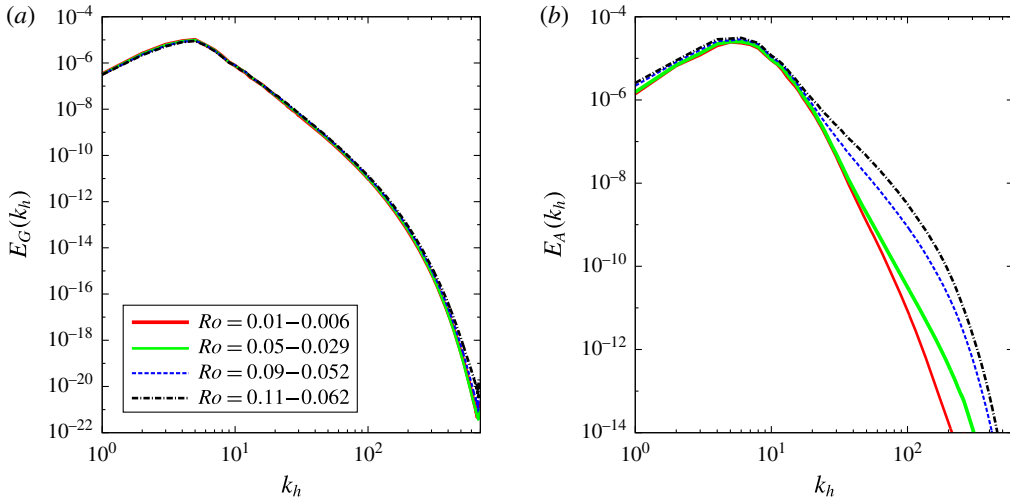


FIGURE 14. (Colour online) The geostrophic (a) and ageostrophic (b) part of energy spectrum for simulations with Laplacian viscosity at $t = 6\tau$. The ageostrophic energy spectra are scaled with Ro^2 as in figure 13. The corresponding runs are B2028r01n, B2028r05n, B2028r09n and B2028r11n.

To verify these results using hyperviscosity we performed several simulations with Laplacian viscosity and diffusion at the higher resolution of $2048^2 \times 256$ (figure 14). Similar to the geostrophic spectra of figure 12, in figure 14 there is not much variation in the geostrophic spectra at the different Rossby numbers. Similarities can also be seen in the ageostrophic spectra of figures 13 and 14 as well. As before, we see in figure 14 that the spectrum of ageostrophic energy becomes more shallow as Rossby number is increased. The slope of the ageostrophic modes in the lowest Rossby number simulation in figure 14 is between -6 and -7 , similar to the rapidly rotating cases using hyperviscosity.

For a better understanding of the dynamics, examining the energy fluxes and transfers in tandem with the energy spectra is very helpful. Figure 15 depicts the net transfer from the geostrophic to ageostrophic modes through nonlinear interactions shown in (2.6). At low Rossby number this transfer is almost zero, showing that the exchange of energy between geostrophic and ageostrophic modes is very small. As rotation weakens, the transfer between geostrophic and ageostrophic modes becomes larger. This exchange is higher at early times and it gets smaller with time. At the higher Ro , the transfer is dominantly from geostrophic to ageostrophic modes, which is consistent with the energy plots of figure 5. In figure 5, the level of E_A increases at the beginning. Given there is no forcing, this increase must come from the geostrophic modal energy.

The total energy flux, which is the sum of $\Pi_G(k_h)$ and $\Pi_A(k_h)$ defined in (2.7), is plotted in figure 16. At $Ro = 0.01$ it follows the QG scenario. There is a negative trough in the large scales, implying an inverse cascade of energy. As the Rossby number increases, a positive-flux range emerges at scales below the peak in the spectrum that signals a forward cascade of energy in these flows. It is stronger at early times and becomes weaker in time as Ro decreases.

To provide more insight on the role of rotation and stratification on the breakdown of balance, the local Rossby number is portrayed in 3D in figure 17. It is derived

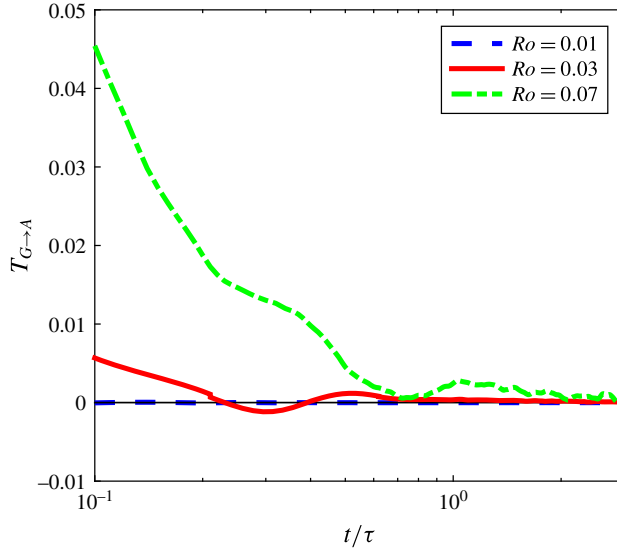


FIGURE 15. (Colour online) The total energy transferred from the geostrophic to ageostrophic modes via nonlinear interactions ($T_{G \rightarrow A}$) versus time for the runs: B2048r01, B2048r03 and B2048r07.

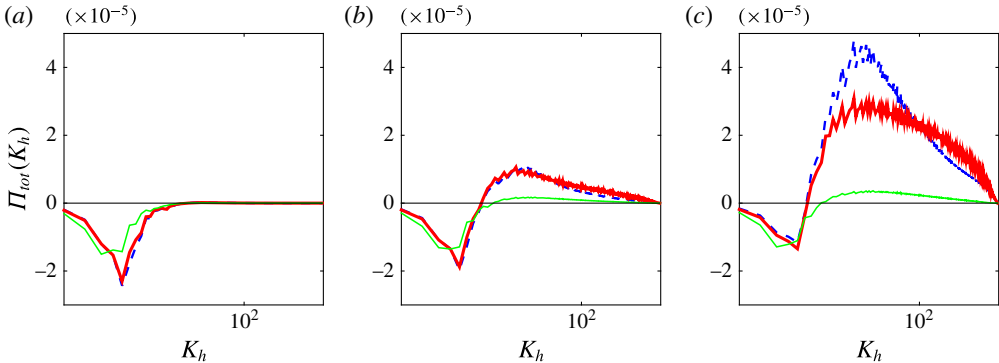


FIGURE 16. (Colour online) The energy flux spectra for B2048r01, B2048r05 and B2048r07 at different Ro . (a) $Ro = 0.01$, (b) $Ro = 0.05$, (c) $Ro = 0.07$. The dashed lines (blue online) are the flux spectra averaged over $[0, 0.2\tau]$, the thick solid lines (red online) averaged over $[0.2\tau, 0.4\tau]$ and the thin solid lines (green online) averaged over $[1.6\tau, 1.8\tau]$.

by scaling the vertical vorticity by the Coriolis parameter, i.e. $Ro(x, y, z, t) = (\nabla \times \mathbf{u}(x, y, z, t) \cdot \hat{\mathbf{z}})/f$. To better show the vertical structure, we scale the vertical axis by N/f . The transparency curve in the upper left corner of the figure is designed such that the points with large absolute values are opaque and visible, whereas lower absolute magnitudes of vorticity are almost transparent. This technique emphasizes the intense structures of the flow.

The coherent structures of QG flow have been extensively studied in the past, for example in McWilliams *et al.* (1994) and McWilliams, Weiss & Yavneh (1999). The results are similar to the vortex structures visible in figure 17(a,c) (low Ro).

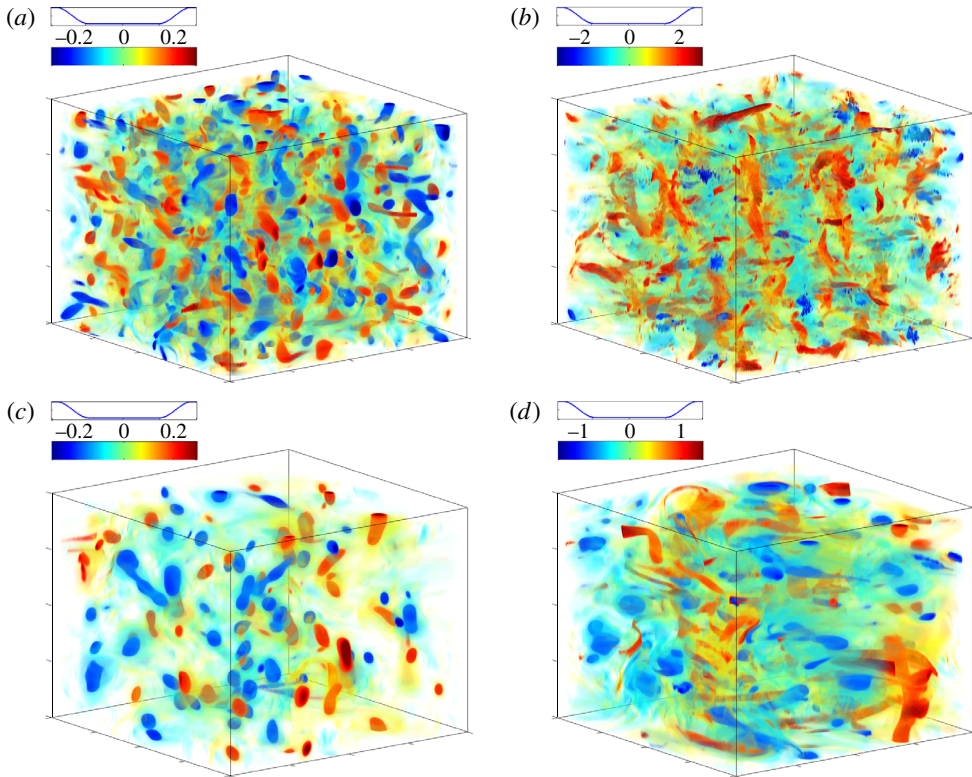


FIGURE 17. (Colour online) Local Rossby number, $Ro(x, y, z) = (\nabla \times \mathbf{u}(x, y, z) \cdot \hat{\mathbf{z}})/f$, in 3D space at $t = 5\tau$ (*a,b*) and $t = 20\tau$ (*c,d*). The initial Ro is set to 0.01 in (*a,c*) and 0.09 in (*b,d*). Values are displayed by assigning a colour and a degree of transparency to them. There are two horizontal bars on the top left corner showing the transparency (upper curve) and the hue (lower colourbar) of grid values as a function of $Ro(x, y, z)$.

This is quite expected as the strongly rotating flow remains dynamically close to QG, especially at larger scales, where coherent vortex structures appear. As explained in a number of previous studies, there are two important mechanisms in the formation and evolution of vortices in geostrophic flows: the merger of two like-sign vortices (Melander, Zabusky & McWilliams 1988; Polvani, Zabusky & Flierl 1989) and vertical alignment (McWilliams 1989; Polvani 1991). Due to the weaker rotation, these structures are observed less at $Ro = 0.09$. Figure 17 also manifest the difference in slope of the energy spectrum at low and high Rossby numbers. At $Ro = 0.01$, where the energy spectrum is steeper, the vortex structures have distinct boundaries, whereas at $Ro = 0.09$ the vortex cores are smaller and more diffuse. There is also stronger mixing at the higher Ro due to the shallower spectrum. The coherent structures make the flow more anisotropic at low Rossby numbers. Hence, we see that the slope of the energy spectrum is steeper than predicted by isotropic theory. However, at the higher Ro of figure 17, the flow looks more isotropic. Hence, the energy spectra of weakly rotating flows are closer to the isotropic slope of $-5/3$.

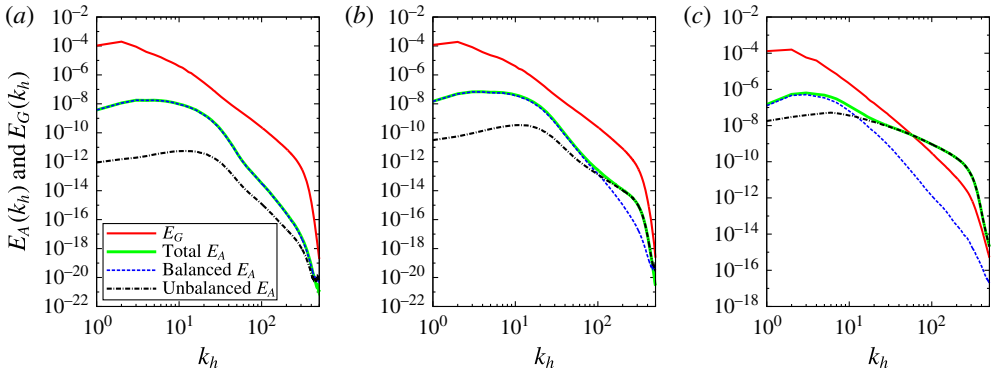


FIGURE 18. (Colour online) The balanced and unbalanced part of the ageostrophic energy spectrum. Geostrophic energy is shown for reference. The initial Ro values are: (a) $Ro = 0.01$, (b) $Ro = 0.02$ and (c) $Ro = 0.07$.

5.5. *The persistence of balance*

In this section, we try to examine the spontaneous imbalance generated in initially balanced flow as time grows. To that end we take the output of the Boussinesq simulation after $t = 10\tau$ and implement Baer–Tribbia initialization again to derive the balanced part of the motion. Then, we subtract the balanced field from the total field to obtain the unbalanced part of the motion.

Figure 18 shows the balanced and unbalanced, as well as the total ageostrophic spectrum. For the lowest Ro , the rotation is very strong, and the flow stays in balance. For this reason the balanced and total spectra lie on top of each other and the unbalanced spectrum is negligibly small. For the middle Ro , a kink emerges in the total ageostrophic spectrum. The balanced spectrum retains its steep slope. However, the shallow unbalanced part grows in amplitude and intersects the balanced spectrum. Hence, the total ageostrophic spectrum, which is the sum of the two, displays a kink near their intersection. At the highest Ro , the unbalanced spectrum grows further in magnitude and the crossover wavenumber moves to larger scales. Since a shallow range of the total ageostrophic spectrum emerges at larger scales, we can see that the geostrophic and ageostrophic spectra intersect as well.

To see how the onset of breakdown scales with Ro , we plotted the wavenumber at which balanced and unbalanced ageostrophic spectra cross ($k_{h,cross}$) as a function of Ro in figure 19. This was extracted for a number of Ro and at two different resolutions. It clearly shows that $k_{h,cross}$ scales with Ro^{-2} , just as the total balanced ageostrophic energy.

Considering figure 18 along with figure 19, we can elucidate the mechanism of balance breakdown. The ageostrophic component consists of a balanced and unbalanced part

$$A_k = A_k^{balanced} + A_k^{unbalanced}. \tag{5.1}$$

The balanced spectrum has a steep slope (close to -6) while the unbalanced slope is shallow. We speculate that the unbalanced slope is close to $-5/3$, given sufficient resolution, or equivalently high enough Reynolds numbers. As rotation weakens, the shallow unbalanced part of the spectrum increases in amplitude and hence extends to larger scales. The kink therefore appears at lower wavenumbers. This shallow part of the ageostrophic energy spectrum may intersect the steeper geostrophic spectrum and

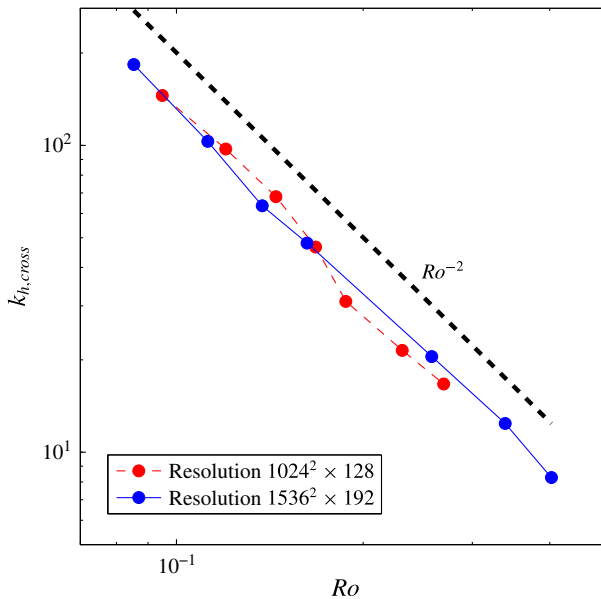


FIGURE 19. (Colour online) The wavenumber at which the balanced and unbalanced part of the ageostrophic energy spectrum cross versus (vorticity-based) Ro .

can hardly be neglected in the dynamics at smaller scales. Therefore, the total energy consists of a steeper spectrum at large scales, dominated by geostrophic dynamics, and a shallow spectrum at small scales, dominated by unbalanced ageostrophic dynamics. This is consistent with atmospheric data (Gage 1979) and the previous unbalanced simulations of Bartello (2010). Further below this intersection scale, geostrophic modes also transition to a shallow spectrum as the linear decomposition loses its meaning.

The scaling of figure 19 can be explained by considering (3.5), where the ageostrophic modes are expanded in terms of Ro and kept to first order. Hence, it can be concluded that $A^{balanced}/A^{unbalanced} = O(Ro^{-1})$ in our initialization scheme. Therefore, the balanced fraction of the ageostrophic energy is $O(Ro^{-2})$ larger than the unbalanced part. Hence, the change in the crossover wavenumber is expected to be proportional to Ro^{-2} as well. A similar argument can be made for the wavenumber at which geostrophic and ageostrophic spectra cross. It should be noted that other definitions of balance (at higher order, for example) produce different scalings for $A^{balanced}/A^{unbalanced}$ at large Ro .

A visualization of total, balanced and unbalanced parts of the vertical vorticity is displayed in figure 20. The steep spectra of balanced ageostrophic modes and the shallow spectra of unbalanced modes are reflected in the real-space plots of panels 20(b) and 20(c), respectively. The unbalanced part of the ageostrophic modes appears as small-scale structures surrounding the larger-scale vortices, whereas the balanced part embodies smooth vortical cores similar in scale to those of the total flow. These ageostrophic centres do not necessarily have the same signs as those of the total flow, which are predominantly geostrophic. Note that in figure 20(c) there are also some large-scale extrema corresponding to those of the total flow. This could be due to the shortcoming of our balance model, which is only second order in Ro . It may be that the vortical cores appearing there are corrections of even higher order to balance.

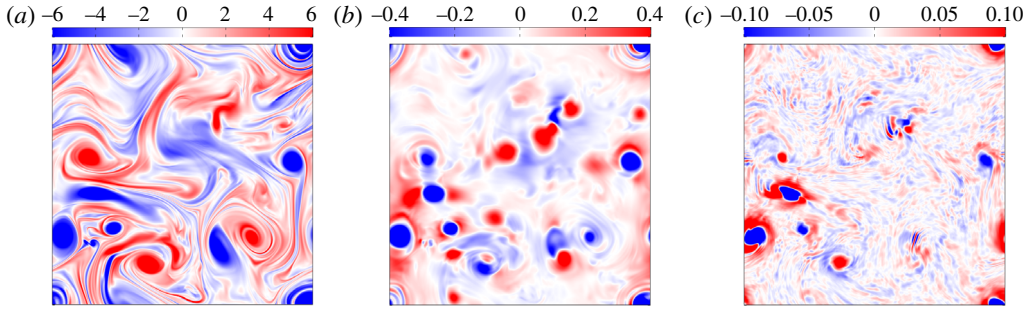


FIGURE 20. (Colour online) Horizontal slices of vertical vorticity ($\zeta = \nabla \times \mathbf{u}(x, y, z) \cdot \hat{\mathbf{z}}$) for (a) the total flow, ζ_{total} , (b) the balanced ageostrophic modes, $\zeta_{balanced\ ageo}$, and (c) the unbalanced ageostrophic modes, $\zeta_{unbalanced\ ageo}$. The total flow has the initial Rossby number of 0.035.

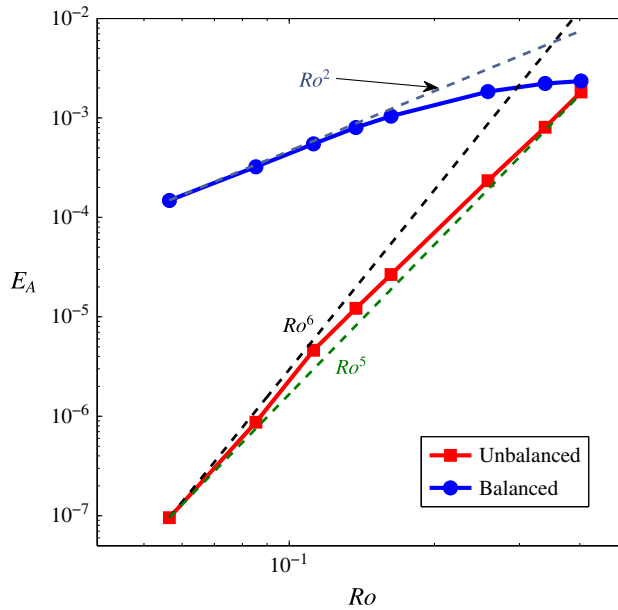


FIGURE 21. (Colour online) The unbalanced and balanced part of ageostrophic energy as a function of (vorticity-based) Ro at $t = 10\tau$.

An example of a dipole can be seen in the lower left part of the domain. There is a strong bundle of waves in the unbalanced panel travelling with this dipole. This is similar to the slow-moving inertia-gravity wave packets formed along dipoles in the studies of Snyder *et al.* (2007) and Viúdez (2007, 2008). Note that these studies considered isolated dipoles, while we are considering a large distribution of vortices in homogeneous turbulence.

Figure 21 presents the balanced and unbalance ageostrophic energy summed over all wavenumbers as a function of Ro . According to (3.7a), when the flow is balanced we expect $A_k \sim O(Ro)$. Hence, the ageostrophic energy asymptotes to Ro^2 as $Ro \rightarrow 0$. Figure 21 also shows that the unbalanced ageostrophic energy increases faster as rotation gets weak. The unbalanced ageostrophic energy scales as Ro^α , with α

between 5 and 6, whereas the balanced energy scales as Ro^2 at low Ro , and becomes even weaker at larger Ro . At small Rossby numbers, different balance definitions should asymptote to the same behaviour as in figure 21. However, as Ro increases the scaling of unbalanced ageostrophic energy is subject to the error in the definition of balance.

6. Conclusion

We have described the evolution of nonlinearly balanced initial conditions under the non-hydrostatic Boussinesq equations. To that end, we used the Baer–Tribbia scheme, which is a nonlinear normal mode initialization method enforcing balance at $O(Ro^2)$ by calculating ageostrophic modes that eliminate fast time derivatives, at least initially.

We investigated how initially balanced and energetically equivalent unbalanced flows differ in their evolution in time and modal interplay. At strong rotation, the time series of large-scale average quantities of initially balanced and unbalanced simulations do not show large differences. Nevertheless, there is a considerable difference in the frequency spectra over the range of their linear inertia-gravity wave frequencies. This shows that the initialization can lower the amplitude of these wave modes while not substantially affecting slower subinertial frequencies associated with the large-scale quasigeostrophic flow. Of course the effectiveness of this procedure is reduced as Ro increases.

The initialization produces an ageostrophic spectrum that is much steeper than the (already steep) geostrophic spectrum. Hence, one can conclude that balance occurs in conjunction with steep spectra and that balance dynamics cannot explain the observed subdeformation scale $-5/3$ spectra in the atmosphere and ocean. If Ro is small enough, ageostrophic modes maintain their steep balanced spectrum at a given resolution and over a given integration period. As Ro is increased, a shallow tail develops in the ageostrophic spectrum after a short time, which we have identified as signalling the growth of unbalanced dynamics. The onset of this unbalanced range occurs at lower wavenumbers when Ro is increased further. It also becomes increasingly shallow, and asymptotes to a slope consistent with $k^{-5/3}$, if the Reynolds number is high enough. In fact, we found the slope of the shallow range to be somewhat sensitive to viscosity and diffusion at our resolutions.

Our results for strong rotation confirm and complete the picture drawn by Dritschel & McKiver (2015) for the maintenance of balance. At very small Rossby numbers, the ageostrophic spectrum stays balanced and steep in the resolutions considered in their study. However, our numerous simulations at different Ro and Re hint that any initially balanced rotating stratified flow develops an unbalanced shallow tail at sufficiently large times if the resolution is large enough (or equivalently dissipation is weak enough) to permit it. This is indeed an open question that can be addressed with higher-resolution simulations. Even if future investigations show the existence of a shallow tail at very small Ro , it will clearly be at very small scales. Considering that its energy will then be very low, unbalanced energy is much smaller than balanced energy, once again consistent with Dritschel & McKiver (2015).

The shallow spectrum resulting from the growth of unbalanced ageostrophic modes is clearly consistent with the atmospheric data (Gage & Nastrom 1986). However, this similarity should be considered within the limitation of our idealized boundary conditions. Internal and external boundaries have been shown to play an important role in the breakdown of balance. Nevertheless, characteristics of unbalanced modes generated spontaneously can be analysed in a periodic configuration such as ours, as a first step in understanding at least the internal dynamics far from boundaries.

When Ro is large enough such that the ageostrophic spectrum admits an unbalanced shallow range, the corresponding frequency spectrum does not exhibit large peaks in the frequency band between N and f . This signals that the developing unbalanced part of the dynamics is not composed of quasilinear high-frequency inertia-gravity waves, but a transition to a more general form of turbulence involving both stratification and rotation.

Acknowledgements

As well as the three reviewers, the authors would like to thank E. Atallah, J. McWilliams and J. Tribbia for their valuable input. Financial support was received from the Natural Sciences and Engineering Research Council and computer resources were generously provided by Compute Canada/Calcul Québec.

Appendix A. Numerical implementation of the Baer–Tribbia scheme

To describe our algorithm succinctly, we use the notation/terminology of linear and nonlinear operators in phase space interchangeably with the vector form modal variables. The evolution of the state variable \mathbf{X} can be expressed by

$$\frac{\partial \mathbf{X}}{\partial t} = i\mathbf{L}\mathbf{X} + \mathcal{N}(\mathbf{X}). \tag{A 1}$$

This equation is the equivalent of (2.1), where all the linear terms are collected in the matrix, \mathbf{L} , and the quadratic terms are presented with $\mathcal{N}(\mathbf{X})$, a nonlinear vector function of the state vector \mathbf{X} .

Using the pseudospectral technique, $\mathcal{N}(\mathbf{X})$ can be efficiently computed by transforming the modal variables to real space, carrying the multiplication for physical variables and then transforming the multiplied terms back to spectral space. The operator $\mathcal{N}(\mathbf{X})$ can take any other vector with the same dimension as \mathbf{X} and output the convolution sum fast and efficiently. In addition to this operator, we consider the following linear projections

$$\zeta = \mathcal{G}\mathbf{X}, \quad \eta = \mathcal{A}\mathbf{X}, \quad \mathbf{X} = \mathcal{P}^{-1}(\zeta, \eta), \tag{A 2a-c}$$

where \mathcal{G} and \mathcal{A} are the projections on geostrophic and ageostrophic manifolds. \mathcal{P}^{-1} does the inverse transform, by taking the projected geostrophic and ageostrophic components and outputting the original state variable. Note that all operators in (A 2) are linear, hence easy to compute.

Using these we derive the nonlinear terms in (2.4) when ageostrophic modes are set to zero

$$\hat{\eta} = \Phi(G, 0) = \mathcal{A}\mathcal{N}(\mathcal{P}^{-1}(\mathcal{G}\mathbf{X}, 0)), \quad \hat{\zeta} = \Psi(G, 0) = \mathcal{G}\mathcal{N}(\mathcal{P}^{-1}(\mathcal{G}\mathbf{X}, 0)). \tag{A 3a,b}$$

In the next step, all coefficients in (3.8) are derived as follows

$$\Psi_k^{0,0} = \hat{\eta}, \tag{A 4}$$

$$\Psi_k^{1,0} = \mathcal{A}\mathcal{N}(\mathcal{P}^{-1}(\mathcal{G}\mathbf{X}, \hat{\eta})) - \mathcal{A}\mathcal{N}(\mathcal{P}^{-1}(0, \hat{\eta})) - \mathcal{A}\mathcal{N}(\mathcal{P}^{-1}(\mathcal{G}\mathbf{X}, 0)), \tag{A 5}$$

$$\begin{aligned}
\Psi_k^{0,1} &= \sum_{kp+q} \psi_{GG} \left(G_p + \sum_{p=m+n} \phi_{GG} G_m G_n \right) \left(G_q + \sum_{q=r+s} \phi_{GG} G_r G_s \right) \\
&\quad - \sum_{k=p+q} \psi_{GG} G_p G_q - \sum_{k=p+q} \psi_{GG} \left(\sum_{p=m+n} \phi_{GG} G_m G_n \right) \left(\sum_{q=r+s} \phi_{GG} G_r G_s \right) \\
&= \mathcal{A} \mathcal{N}(\mathcal{P}^{-1}(\hat{\zeta} + \mathcal{G}\mathbf{X}, 0)) - \mathcal{A} \mathcal{N}(\mathcal{P}^{-1}(\mathcal{G}\mathbf{X}, 0)) - \mathcal{A} \mathcal{N}(\mathcal{P}^{-1}(\hat{\zeta}, 0)). \quad (\text{A } 6)
\end{aligned}$$

REFERENCES

- BAER, F. & TRIBBIA, J. J. 1977 On complete filtering of gravity modes through nonlinear initialization. *Mon. Weath. Rev.* **105** (12), 1536–1539.
- BARTELLO, P. 1995 Geostrophic adjustment and inverse cascades in rotating stratified turbulence. *J. Atmos. Sci.* **52** (24), 4410–4428.
- BARTELLO, P. 2010 Quasigeostrophic and stratified turbulence in the atmosphere. In *IUTAM Symposium on Turbulence in the Atmosphere and Oceans*, pp. 117–130. Springer.
- BARTELLO, P., MÉTAIS, O. & LESIEUR, M. 1996 Geostrophic versus wave eddy viscosities in atmospheric models. *J. Atmos. Sci.* **53** (4), 564–571.
- BARTELLO, P. & TOBIAS, S. M. 2013 Sensitivity of stratified turbulence to the buoyancy Reynolds number. *J. Fluid Mech.* **725**, 1–22.
- BRETHOUWER, G., BILLANT, P., LINDBORG, E. & CHOMAZ, J. M. 2007 Scaling analysis and simulation of strongly stratified turbulent flows. *J. Fluid Mech.* **585**, 343–368.
- CALLIES, J., FERRARI, R. & BÜHLER, O. 2014 Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl Acad. Sci. USA* **111** (48), 17033–17038.
- CHARNEY, J. G. 1971 Geostrophic turbulence. *J. Atmos. Sci.* **28** (6), 1087–1095.
- COLE, J. D. 1968 *Perturbation Methods in Applied Mathematics*. Blaisdell.
- DEUSEBIO, E., VALLGREN, A. & LINDBORG, E. 2013 The route to dissipation in strongly stratified and rotating flows. *J. Fluid Mech.* **720**, 66–103.
- DRITSCHEL, D. G. & MCKIVER, W. J. 2015 Effect of Prandtl's ratio on balance in geophysical turbulence. *J. Fluid Mech.* **777**, 569–590.
- DRITSCHEL, D. G. & VIÚDEZ, A. 2003 A balanced approach to modelling rotating stably stratified geophysical flows. *J. Fluid Mech.* **488**, 123–150.
- EVANS, K. J., LAURITZEN, P. H., MISHRA, S. K., NEALE, R. B., TAYLOR, M. A. & TRIBBIA, J. J. 2013 AMIP simulation with the CAM4 Spectral Element Dynamical Core. *J. Clim.* **26** (3), 689–709.
- FORD, R., MCINTYRE, M. E. & NORTON, W. A. 2000 Balance and the slow quasimanifold: some explicit results. *J. Atmos. Sci.* **57** (9), 1236–1254.
- GAGE, K. S. 1979 Evidence for a $k^{-5/3}$ law inertial range in mesoscale two-dimensional turbulence. *J. Atmos. Sci.* **36** (10), 1950–1954.
- GAGE, K. S. & NASTROM, G. D. 1986 Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.* **43** (7), 729–740.
- HAKIM, G. J. 2000 Climatology of coherent structures on the extratropical tropopause. *Mon. Weath. Rev.* **128** (2), 385–406.
- HAMILTON, K., TAKAHASHI, Y. O. & OHFUCHI, W. 2008 Mesoscale spectrum of atmospheric motions investigated in a very fine resolution global general circulation model. *J. Geophys. Res.* **113** (D18), D18110.
- HOSKINS, B. J. & HODGES, K. I. 2002 New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.* **59** (6), 1041–1061.
- KREISS, H.-O. & LORENZ, J. 1994 On the existence of slow manifolds for problems with different timescales. *Phil. Trans. R. Soc. Lond. A* **346** (1679), 159–171.

- LEITH, C. E. 1980 Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.* **37** (5), 958–968.
- LINDBORG, E. 2006 The energy cascade in a strongly stratified fluid. *J. Fluid Mech.* **550**, 207–242.
- LORENZ, E. N. 1980 Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.* **37** (8), 1685–1699.
- LORENZ, E. N. 1986 On the existence of a slow manifold. *J. Atmos. Sci.* **43** (15), 1547–1558.
- LORENZ, E. N. 1992 The slow manifold-what is it? *J. Atmos. Sci.* **49** (24), 2449–2451.
- LORENZ, E. N. & KRISHNAMURTHY, V. 1987 On the nonexistence of a slow manifold. *J. Atmos. Sci.* **44** (20), 2940–2950.
- MARINO, R., MININNI, P. D., ROSENBERG, D. & POUQUET, A. 2013 Inverse cascades in rotating stratified turbulence: fast growth of large scales. *Europhys. Lett.* **102** (4), 44006.
- MCKIVER, W. J. & DRITSCHEL, D. G. 2008 Balance in non-hydrostatic rotating stratified turbulence. *J. Fluid Mech.* **596**, 201–219.
- MCWILLIAMS, J. C. 1989 Statistical properties of decaying geostrophic turbulence. *J. Fluid Mech.* **198**, 199–230.
- MCWILLIAMS, J. C., WEISS, J. B. & YAVNEH, I. 1994 Anisotropy and coherent vortex structures in planetary turbulence. *Science* **264** (5157), 410–413.
- MCWILLIAMS, J. C., WEISS, J. B. & YAVNEH, I. 1999 The vortices of homogeneous geostrophic turbulence. *J. Fluid Mech.* **401**, 1–26.
- MELANDER, M. V., ZABUSKY, N. J. & MCWILLIAMS, J. C. 1988 Symmetric vortex merger in two dimensions: causes and conditions. *J. Fluid Mech.* **195**, 303–340.
- NADIGA, B. T. 2014 Nonlinear evolution of a baroclinic wave and imbalanced dissipation. *J. Fluid Mech.* **756**, 965–1006.
- POLVANI, L. M. 1991 Two-layer geostrophic vortex dynamics. Part 2. Alignment and two-layer V-states. *J. Fluid Mech.* **225**, 241–270.
- POLVANI, L. M., ZABUSKY, N. J. & FLIERL, G. R. 1989 Two-layer geostrophic vortex dynamics. Part 1. Upper-layer V-states and merger. *J. Fluid Mech.* **205**, 215–242.
- POUQUET, A. & MARINO, R. 2013 Geophysical turbulence and the duality of the energy flow across scales. *Phys. Rev. Lett.* **111** (23), 234501.
- REZNIK, G. M., ZEITLIN, V. & BEN JELLOUL, M. 2001 Nonlinear theory of geostrophic adjustment. Part 1. Rotating shallow-water model. *J. Fluid Mech.* **445**, 93–120.
- SKAMAROCK, W. C. 2004 Evaluating mesoscale nwp models using kinetic energy spectra. *Mon. Weath. Rev.* **132** (12), 3019–3032.
- SNYDER, C., MURAKI, D. J., PLOUGONVEN, R. & ZHANG, F. 2007 Inertia-gravity waves generated within a dipole vortex. *J. Atmos. Sci.* **64** (12), 4417–4431.
- TULLOCH, R. & SMITH, K. S. 2006 A theory for the atmospheric energy spectrum: depth-limited temperature anomalies at the tropopause. *Proc. Natl Acad. Sci. USA* **103** (40), 14690–14694.
- VALLGREN, A., DEUSEBIO, E. & LINDBORG, E. 2011 Possible explanation of the atmospheric kinetic and potential energy spectra. *Phys. Rev. Lett.* **107** (26), 268501.
- VANNESTE, J. 2013 Balance and spontaneous wave generation in geophysical flows. *Annu. Rev. Fluid Mech.* **45**, 147–172.
- VANNESTE, J. & YAVNEH, I. 2004 Exponentially small inertia-gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.* **61** (2), 211–223.
- VIÚDEZ, Á. 2007 The origin of the stationary frontal wave packet spontaneously generated in rotating stratified vortex dipoles. *J. Fluid Mech.* **593**, 359–383.
- VIÚDEZ, Á. 2008 The stationary frontal wave packet spontaneously generated in mesoscale dipoles. *J. Phys. Oceanogr.* **38** (1), 243–256.
- VIÚDEZ, Á. & DRITSCHEL, D. G. 2004 Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.* **521**, 343–352.
- WAITE, M. L. & BARTELLO, P. 2004 Stratified turbulence dominated by vortical motion. *J. Fluid Mech.* **517**, 281–308.
- WAITE, M. L. 2011 Stratified turbulence at the buoyancy scale. *Phys. Fluids* **23** (6), 066602.
- WAITE, M. L. & SNYDER, C. 2009 The mesoscale kinetic energy spectrum of a baroclinic life cycle. *J. Atmos. Sci.* **66** (4), 883–901.
- WARN, T. 1997 Nonlinear balance and quasi-geostrophic sets. *Atmos.-Ocean* **35** (2), 135–145.

- WARN, T., BOKHOVE, O., SHEPHERD, T. G. & VALLIS, G. K. 1995 Rossby number expansions, slaving principles, and balance dynamics. *Q. J. R. Meteorol. Soc.* **121** (523), 723–739.
- WARN, T. & MENARD, R. 1986 Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model. *Tellus A* **38** (4), 285–294.
- WHITEHEAD, J. P. & WINGATE, B. A. 2014 The influence of fast waves and fluctuations on the evolution of the dynamics on the slow manifold. *J. Fluid Mech.* **757**, 155–178.
- ZEITLIN, V., REZNIK, G. M. & JELLOUL, M. 2003 Nonlinear theory of geostrophic adjustment. Part 2. Two-layer and continuously stratified primitive equations. *J. Fluid Mech.* **491**, 207–228.