

A unified theory of balance in the extratropics

Article

Published Version

Saujani, S. and Shepherd, T. G. ORCID:
<https://orcid.org/0000-0002-6631-9968> (2006) A unified theory
of balance in the extratropics. Journal Of Fluid Mechanics,
569. pp. 447-464. ISSN 0022-1120 doi:
10.1017/S0022112006002783 Available at
<https://centaur.reading.ac.uk/32070/>

It is advisable to refer to the publisher's version if you intend to cite from the
work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1017/S0022112006002783>

To link to this article DOI: <http://dx.doi.org/10.1017/S0022112006002783>

Publisher: Cambridge University Press

All outputs in CentAUR are protected by Intellectual Property Rights law,
including copyright law. Copyright and IPR is retained by the creators or other
copyright holders. Terms and conditions for use of this material are defined in
the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

A unified theory of balance in the extratropics

By SIMAL SAUJANI AND THEODORE G. SHEPHERD

Department of Physics, University of Toronto, Toronto, Ontario M5S 1A7, Canada

(Received 19 August 2005 and in revised form 12 June 2006)

Many physical systems exhibit dynamics with vastly different time scales. Often the different motions interact only weakly and the slow dynamics is naturally constrained to a subspace of phase space, in the vicinity of a slow manifold. In geophysical fluid dynamics this reduction in phase space is called balance. Classically, balance is understood by way of the Rossby number R or the Froude number F ; either $R \ll 1$ or $F \ll 1$.

We examined the shallow-water equations and Boussinesq equations on an f -plane and determined a dimensionless parameter ϵ , small values of which imply a time-scale separation. In terms of R and F ,

$$\epsilon = \frac{RF}{\sqrt{R^2 + F^2}}.$$

We then developed a unified theory of (extratropical) balance based on ϵ that includes all cases of small R and/or small F . The leading-order systems are ensured to be Hamiltonian and turn out to be governed by the quasi-geostrophic potential-vorticity equation. However, the height field is not necessarily in geostrophic balance, so the leading-order dynamics are more general than in quasi-geostrophy. Thus the quasi-geostrophic potential-vorticity equation (as distinct from the quasi-geostrophic dynamics) is valid more generally than its traditional derivation would suggest. In the case of the Boussinesq equations, we have found that balanced dynamics generally implies hydrostatic balance without any assumption on the aspect ratio; only when the Froude number is not small and it is the Rossby number that guarantees a time-scale separation must we impose the requirement of a small aspect ratio to ensure hydrostatic balance.

1. Introduction

The concept of balance is a central theme in geophysical fluid dynamics (e.g. Daley 1991). Although rotating stratified fluids allow for both vortical and inertia-gravity-wave motions, we often observe the former to play a dominant role in the time evolution of the system. That is, we find the dependent variables are constrained by diagnostic equations, and we say the dynamics is balanced. There has been much effort to derive dynamical equations to model geophysical flows that remain balanced, or nearly so, for all time. Such models are an important means of developing theoretical understanding and diagnostic-analysis methods for the large-scale atmosphere and oceans.

Recent research has shown that for non-trivial flows the concept of balance is not exact (e.g. Warn 1997; Ford, McIntyre & Norton 2000; Wirosoetisno & Shepherd 2000),

and most mathematical approaches to balance are based on an asymptotic analysis. However, practical applications of balance show that it is surprisingly accurate even when the asymptotic parameters are not particularly small (McIntyre & Norton 2000).

Classically, there are two main approaches corresponding to two distinct balanced regimes. The small-Rossby-number limit (Charney 1948) corresponds to a rapidly rotating flow and, as is well known, yields the quasi-geostrophic equations. The small-Froude-number limit (Charney 1963) corresponds to a strongly stratified flow and was originally developed for the tropics, where the Rossby number is unbounded. It has been argued that this regime is also relevant to the extratropical mesoscale, where the Rossby number is $O(1)$ or larger (Lilly 1983). In both cases, small values of a dimensionless parameter characterize a particular balanced dynamics; the parameter may be considered as a dimensionless amplitude of the motion (relative to rotation or stratification). Assuming an asymptotic expansion in this parameter, we may write a hierarchy of dynamical equations. The lowest-order equations capture the essential effects of the small parameter; higher-order equations may be solved as needed. Care must be taken, however, that the parameter is sufficiently small to ensure that the analysis and resulting models are valid. In particular, attempting to allow for, say, $O(1)$ Rossby-number terms by including higher-order terms in a small-Rossby-number expansion is inconsistent. It would also seem to be quite unnecessary since the basis for balance, in this case, should be a small Froude number.

The derivations of these two classical models are completely independent. Although there are *ad hoc* methods to combine them *a posteriori* (e.g. McWilliams 1985; Vallis 1996), a more general theory would incorporate both cases systematically and simultaneously. Moreover, since the governing equations are Hamiltonian, the leading-order balance equations would themselves be Hamiltonian. This is important for the development of general theory (Shepherd 2003). Combining models in an *ad hoc* fashion tends to spoil the Hamiltonian structure.

Fundamentally, balance is possible when the dynamics admits motions with vastly different time scales (van Kampen 1985). For geophysical flows, balance requires the time scale of the vortical motion to be significantly larger than that of the inertia-gravity waves. Clearly, if the dynamics is to be determined by a subset of the variables (the ‘slow’ variables) then we would expect little or no interaction with the remaining (‘fast’) variables. By demanding different time scales we minimize interactions due to frequency matching through nonlinearity (Errico 1982; Saujani & Shepherd 2002). Furthermore, by distinguishing slow and fast variables we can avoid secular growth in the asymptotic analysis (Warn *et al.* 1995; Wirosotetisno, Shepherd & Temam 2002). We shall therefore define a dimensionless time-scale-separation parameter ϵ , small values of which will characterize balance rather than a particular balanced state. This simple definition makes ϵ the correct asymptotic parameter for a general theory of balance. It turns out that

$$\epsilon = \frac{RF}{\sqrt{R^2 + F^2}}, \quad (1.1)$$

where R is the Rossby number and F is the Froude number. It follows that $\epsilon \sim \min[R, F]$. Hence small ϵ includes the small-Rossby-number and small-Froude-number cases – the latter even when $R \rightarrow \infty$. Moreover, no further assumptions are necessary. In particular, the balance of terms in the equations of motion is not considered.

For a nonlinear system we infer the dominant time scales of motion by considering the dispersion relation for the linearized system and comparing the resulting time scales with the nonlinear (advective) time scale. We expect the linearized frequencies

to correspond to peaks in the frequency spectrum, broadened due to nonlinearity. The geophysical models to be considered possess just two time scales, one slow and one fast. By considering their ratio we may choose the dimensionless product ϵ in such a way that the limit $\epsilon \rightarrow 0$ forces this ratio to be large. Consequently, we expect some kind of balance with ϵ as our asymptotic parameter.

Explicit in our approach is the assumption of a reference state. This is what allows us to define a time-scale separation and thus develop the asymptotic theory that leads to the balance constraints, without any further assumptions. A complementary approach (e.g. Salmon 1983) is to impose the balance constraints on the basis of empirical knowledge, in which case there is no need to introduce a reference state.

We focus on the rotating shallow-water equations and their Boussinesq counterpart. The Coriolis parameter is taken to be constant (not excluding zero). Our analysis incorporates the slaving technique of Warn *et al.* (1995). This leads to unified balanced dynamical equations that apply equally well to both the small-Rossby-number and small-Froude-number regimes.

2. Analysis of the shallow-water equations

2.1. Fast and slow variables

The rotating shallow-water equations in Cartesian coordinates are

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - f v + g \frac{\partial h}{\partial x} = 0, \quad (2.1)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + f u + g \frac{\partial h}{\partial y} = 0, \quad (2.2)$$

$$\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + v \frac{\partial h}{\partial y} + h \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0, \quad (2.3)$$

where (u, v) is the horizontal fluid velocity, f is the (constant) Coriolis parameter, h is the surface height and g is the (constant) acceleration due to gravity. For simplicity, we shall assume a domain D with doubly periodic boundary conditions.

First we linearize about a rest state $(u, v, h) = (0, 0, H)$, where H is constant. Assuming normal-mode solutions of the form

$$\{u, v, h\} = \{\tilde{u}, \tilde{v}, \tilde{h}\} e^{i(kx + ly - \omega t)} \quad (2.4)$$

the dispersion relation is found to be

$$\omega^3 - (f^2 + gH\kappa^2)\omega = 0, \quad (2.5)$$

where $\kappa^2 \equiv k^2 + l^2$. The solutions are

$$\omega_0 = 0, \quad \omega_{\pm} = \pm \sqrt{f^2 + gH\kappa^2}. \quad (2.6)$$

Under nonlinear frequency broadening by advection these become, to a good approximation,

$$\omega_0 \sim U\kappa, \quad \omega_{\pm} \sim U\kappa \pm \sqrt{f^2 + gH\kappa^2}, \quad (2.7)$$

where U is a characteristic flow speed. Now consider the ratio of the slow time scale and the fast time scale, that is, the ratio of the fast frequency and the slow. We have

$$\frac{\omega_+}{\omega_0} \sim 1 + \frac{\sqrt{R^2 + F^2}}{RF}, \quad (2.8)$$

where $R = U\kappa/f$ is the Rossby number and $F = U/\sqrt{gH}$ is the Froude number.

Two important points are borne from the above simple analysis. First, the dispersion relation has three solutions: one slow vortical mode and two fast inertia-gravity waves. Thus there are two time scales of the motion and we expect to find suitable variables that split the dynamics into these separate time scales. Second, the time scales are distinct if and only if the ratio ω_+/ω_0 is large, i.e. if and only if

$$\epsilon \equiv \frac{RF}{\sqrt{R^2 + F^2}} \ll 1. \quad (2.9)$$

This ϵ is our time-scale separation parameter. As $\epsilon \rightarrow 0$ the motion may be considered as weak (relative to rotation and gravity) and we are ensured disparate time scales in the dynamics. This necessary condition satisfied, we are then justified to hope for, if not expect, a balanced system by setting the fast dynamics to zero.

Notice that

$$\frac{\min[R, F]}{\sqrt{2}} \leq \epsilon \leq \min[R, F]. \quad (2.10)$$

As an asymptotic parameter for balance, ϵ is equivalent to $\min[R, F]$ and yet has been determined solely from the physics. Both the traditional approximations, the small-Rossby-number limit ($R \rightarrow 0$, $B \equiv F/R$ fixed) and the small-Froude-number limit ($F \rightarrow 0$, R fixed), imply $\epsilon \rightarrow 0$. However, $\epsilon \rightarrow 0$ is more general; it includes, for example, planetary geostrophy ($R \rightarrow 0$, F fixed). We will consider the time-scale separation limit $\epsilon \rightarrow 0$ without specifying R or F , thereby obtaining a unified scaling.

Let us now set about the task of finding suitable variables. Our goal is to rewrite the shallow-water equations in terms of slow and fast variables that decouple the slow motion from the fast. Following Warn *et al.* (1995) the desired form is

$$\frac{\partial s}{\partial t} = \mathcal{S}(s, \tilde{f}; \epsilon), \quad (2.11)$$

$$\frac{\partial \tilde{f}}{\partial t} + \frac{\Lambda \tilde{f}}{\epsilon} = \mathcal{F}(s, \tilde{f}; \epsilon). \quad (2.12)$$

Here s and $\tilde{f}$ refer to slow and fast variables respectively, Λ is a linear invertible operator, $\mathcal{S}$ and $\mathcal{F}$ are nonlinear operators and ϵ is a small asymptotic parameter. Obviously, s evolves on an $O(1)$ time scale, while $\tilde{f}$ naturally evolves on a time scale of $O(\epsilon) \ll 1$. The eigenvalues of Λ are the eigenfrequencies of the fast waves on the fast time t/ϵ .

First we must non-dimensionalize the equations. We use the characteristic flow speed U , the horizontal length scale L , the mean depth H and, as we are seeking slow solutions, the advective time scale $T = L/U$. Noting that all variables are now dimensionless, the equations become

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - \frac{v}{R} + \frac{1}{F^2} \frac{\partial h}{\partial x} = 0, \quad (2.13)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + \frac{u}{R} + \frac{1}{F^2} \frac{\partial h}{\partial y} = 0, \quad (2.14)$$

$$\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + v \frac{\partial h}{\partial y} + h \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0. \quad (2.15)$$

Since $\epsilon \ll 1$ corresponds to weak motions, it is natural to assume small surface displacements by taking $h = 1 + \epsilon \eta$. This, however, neglects the possibility of large length scales over which small surface displacements can accumulate to produce $O(1)$

effects. In the light of this $R \rightarrow 0$ limit, we choose to scale with $h = 1 + F\eta$. Also, it is convenient to rewrite these equations in terms of a velocity potential and stream function: $(u, v) = \nabla\phi + \hat{z} \times \nabla\psi$. This allows us to replace (2.13), (2.14) with equations for the scalars ϕ and ψ . Finally, choosing the rotational Froude number $B = F/R$ (the shallow-water analogue of the inverse square root of the Burger number) and ϵ instead of the Rossby and Froude numbers as our dimensionless parameters, we obtain

$$\begin{aligned} \frac{\partial}{\partial t} \nabla^2 \phi - \frac{1}{\epsilon} \left(\frac{B}{\sqrt{1+B^2}} \nabla^2 \psi - \frac{1}{\sqrt{1+B^2}} \nabla^2 \eta \right) \\ + J(\phi, \nabla^2 \psi) - 2J \left(\frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \right) + \frac{1}{2} \nabla^2 [\nabla\phi \cdot \nabla\phi - 2J(\phi, \psi)] = 0, \end{aligned} \quad (2.16)$$

$$\frac{\partial}{\partial t} \nabla^2 \psi + \frac{1}{\epsilon} \frac{B}{\sqrt{1+B^2}} \nabla^2 \phi + \nabla \cdot (\nabla^2 \psi \nabla\phi) + J(\psi, \nabla^2 \psi) = 0, \quad (2.17)$$

$$\frac{\partial \eta}{\partial t} + \nabla\phi \cdot \nabla\eta + J(\psi, \eta) + \left(\frac{1}{\sqrt{1+B^2\epsilon}} + \eta \right) \nabla^2 \phi = 0. \quad (2.18)$$

Here $J(f, g) = \det(\partial(f, g)/\partial(x, y))$. As written, the equations naturally suggest the fast variables

$$\delta \equiv \nabla^2 \phi, \quad (2.19)$$

$$\Gamma \equiv b \nabla^2 \psi - \sqrt{1-b^2} \nabla^2 \eta, \quad (2.20)$$

where $b \equiv B/\sqrt{1+B^2} \in [0, 1)$. (We allow $b=0$ so as to include the case $f=0$.) Since the potential vorticity is an exact Lagrangian invariant, it evolves on the advective time scale and thus makes an obvious choice for the slow variable. Here it takes the form

$$Q = \frac{b\sqrt{1-b^2} + \epsilon\sqrt{1-b^2} \nabla^2 \psi}{\sqrt{1-b^2} + \epsilon\eta}. \quad (2.21)$$

However, since we are interested in disturbances to the rest state, rather than potential vorticity the disturbance potential vorticity is the appropriate quantity. Let us call it q . We must have

$$Q = \frac{b\sqrt{1-b^2} + \epsilon\sqrt{1-b^2} \nabla^2 \psi}{\sqrt{1-b^2} + \epsilon\eta} = b + \frac{\epsilon}{\sqrt{1-b^2}} q \quad (2.22)$$

$$\iff q = \frac{(1-b^2)\nabla^2 \psi - b\sqrt{1-b^2}\eta}{\sqrt{1-b^2} + \epsilon\eta}. \quad (2.23)$$

Note that the perturbation potential vorticity is scaled to be an $O(\epsilon)$ quantity, except when $b \rightarrow 1$ ($R \rightarrow 0$) and it becomes $O(1)$. Exactly as for η , its value must be allowed to accrue over large length scales in this limit. We have then, finally,

$$\frac{\partial q}{\partial t} = -\nabla\phi \cdot \nabla q - J(\psi, q), \quad (2.24)$$

$$\frac{\partial \delta}{\partial t} - \frac{\Gamma}{\epsilon} = -J(\phi, \nabla^2 \psi) + 2J \left(\frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \right) - \frac{1}{2} \nabla^2 [\nabla\phi \cdot \nabla\phi + 2J(\psi, \phi)], \quad (2.25)$$

$$\frac{\partial \Gamma}{\partial t} + \frac{A\delta}{\epsilon} = \sqrt{1-b^2} \nabla^2 [\nabla \cdot (\eta \nabla\phi) + J(\psi, \eta)] - b[\nabla \cdot (\nabla^2 \psi \nabla\phi) + J(\psi, \nabla^2 \psi)], \quad (2.26)$$

where $A \equiv b^2 - (1 - b^2)\nabla^2$. That we have succeeded in writing the equations in the desired (Warn *et al.* 1995) form (2.11), (2.12) justifies our choice of fast and slow variables. Here b ranges between 0 (inclusive) and 1 (exclusive), and there is no corresponding singularity in these equations nor in the variables defined by (2.19), (2.20) and (2.23). Thus ϵ is our sole singular perturbation parameter; special limits can subsequently be considered by treating b (i.e. B) as a regular perturbation parameter.

In finding asymptotic solutions we shall repeatedly encounter Poisson's equation, and it is convenient to use the notation $(\nabla^2)^{-1}$. Technically, the Laplacian with periodic boundary conditions has a unique solution only to within an arbitrary constant. For us, this constant can always be shown to be zero or to be attached to a velocity potential or stream function. In the latter case it has no physical significance and we assign it to be zero for simplicity. The distinction will be made clear as necessary. The operator A is strictly invertible for $b \neq 0$ and reduces to the Laplacian when $b = 0$.

2.2. Asymptotic models

Following van Kampen (1985) and Warn *et al.* (1995), we assume the slow variable q to be exact and expand all other variables in our asymptotic parameter ϵ (see Wirosoetisno *et al.* 2002 for a justification). In particular, the fast variables have asymptotic expansions

$$\delta = \delta_0 + \epsilon\delta_1 + \epsilon^2\delta_2 + \cdots \quad (2.27)$$

$$\Gamma = \Gamma_0 + \epsilon\Gamma_1 + \epsilon^2\Gamma_2 + \cdots \quad (2.28)$$

Any other variables are expanded similarly and are necessarily written in terms of q , δ and Γ . Provided $F = \epsilon/\sqrt{1 - b^2} \ll 1$, we may expand q according to

$$q = \sqrt{1 - b^2}\nabla^2\psi_0 - b\eta_0 + \epsilon \left(\sqrt{1 - b^2}\nabla^2\psi_1 - b\eta_1 - \eta_0\nabla^2\psi_0 + \frac{b}{\sqrt{1 - b^2}}\eta_0^2 \right) + O(\epsilon^2). \quad (2.29)$$

Since this equation must hold at all orders,

$$q = \sqrt{1 - b^2}\nabla^2\psi_0 - b\eta_0. \quad (2.30)$$

Now at $O(\epsilon^{-1})$, (2.25) and (2.26) give us

$$-\Gamma_0 = -b\nabla^2\psi_0 + \sqrt{1 - b^2}\nabla^2\eta_0 = 0, \quad (2.31)$$

$$A\delta_0 = [b^2 - (1 - b^2)\nabla^2]\nabla^2\phi_0 = 0. \quad (2.32)$$

Immediately we have $\nabla^2\phi_0 = 0$, whence we may take $\phi_0 = 0$. Using (2.30), (2.31) we can solve for ψ_0 and η_0 in terms of q alone:

$$\psi_0 = -\sqrt{1 - b^2}A^{-1}q, \quad (2.33)$$

$$\eta_0 = -bA^{-1}q. \quad (2.34)$$

Thus, from (2.24), our leading-order evolution equation for the balanced dynamics is

$$\frac{\partial q}{\partial t} = -J(\psi_0, q), \quad (2.35)$$

with (2.33) or, equivalently,

$$q = \sqrt{1 - b^2}\nabla^2\psi_0 - \frac{b^2}{\sqrt{1 - b^2}}\psi_0, \quad (2.36)$$

$$= \frac{1}{\sqrt{1 + B^2}}(\nabla^2\psi_0 - B^2\psi_0). \quad (2.37)$$

This is the leading-order balanced-dynamics model and is guaranteed to be Hamiltonian. The potential-vorticity evolution equation (2.35), (2.36) is identical to the quasi-geostrophic (QG) equation for shallow-water flow (e.g. Pedlosky 1987), but our derivation shows that the equation is valid more generally than its traditional derivation would suggest; no assumption of $R \ll 1$ has been made, only that $\epsilon \ll 1$. As such, different balanced regimes are possible.

The standard QG limit is $R \rightarrow 0$ and $F \rightarrow 0$ with b (i.e. B) held fixed. In this case (2.33), (2.34) gives geostrophic balance and (2.35) gives the corresponding QG evolution equation.

In the small-Froude-number limit $F \rightarrow 0$ with R fixed (but arbitrary), we have $b \rightarrow 0$ (i.e. $B \rightarrow 0$) and $q = \nabla^2 \psi_0$. Here δ_0 satisfies Laplace's equation and so must be constant. Integrating δ_0 over the domain and applying the periodic boundary conditions reveals this constant to be zero. Therefore (2.35) reduces to the two-dimensional Euler equation. This regime holds even if $f = 0$ ($R = \infty$); there is no singularity, we simply have $\epsilon = F \rightarrow 0$. Clearly, quasi-geostrophy is not implied in this limit since (2.31) implies that η_0 is constant. As η_0 is defined to be the departure from the mean depth this constant must be zero. Hence height is a fast variable rather than being related to the stream function through geostrophic balance. Note that it was not necessary to rescale the perturbation height h . All this confirms that we have chosen the correct asymptotic parameter and correct variables to describe the motion.

The third special case of physical interest is planetary geostrophy (PG) or Type II QG scaling; here $R \rightarrow 0$ with F fixed implying $b \rightarrow 1$ (i.e. $B \rightarrow \infty$). Returning to definition (2.23) and noting that $\sqrt{1 - b^2} = \epsilon/F$ we have now

$$q = -\frac{\eta_0}{1 + F\eta_0} + \epsilon \left[\frac{\nabla^2 \psi_0 - F\eta_1}{F(1 + F\eta_0)} + \frac{F\eta_0\eta_1}{(1 + F\eta_0)^2} \right] + O(\epsilon^2), \quad (2.38)$$

whence, since q is exact,

$$q = -\frac{\eta_0}{1 + F\eta_0}. \quad (2.39)$$

Of course, (2.31), (2.32) still apply; we find $\nabla^2 \psi_0 = \nabla^2 \phi_0 = 0$ and the leading-order dynamics is trivial, $\partial q / \partial t = 0$. The height perturbations are, however, non-trivial, being slaved to q via

$$\eta_0 = -\frac{q}{1 + Fq}. \quad (2.40)$$

The key point is that the time-evolution equation (2.35) together with the diagnostic relations (2.31), (2.32) (and the definition (2.30)) represent a unified theory of balance based upon an assumed time-scale separation between the vortical motion and inertia-gravity waves. In particular, there is no assumption of geostrophy, nor any particular assumption on R or F .

Continuing to $O(1)$ in (2.25), (2.26) we find

$$-\Gamma_1 = -b \nabla^2 \psi_1 + \sqrt{1 - b^2} \nabla^2 \eta_1 = 2J \left(\frac{\partial \psi_0}{\partial x}, \frac{\partial \psi_0}{\partial y} \right), \quad (2.41)$$

$$A\delta_1 = [b^2 - (1 - b^2)\nabla^2] \nabla^2 \phi_1 = -bJ(\psi_0, \nabla^2 \psi_0). \quad (2.42)$$

Since q is exact, and first assuming $F \ll 1$, we have from (2.29) that

$$\sqrt{1 - b^2} \nabla^2 \psi_1 - b\eta_1 - \frac{\eta_0 q}{\sqrt{1 - b^2}} = 0. \quad (2.43)$$

The system is linear in ψ_1 , ϕ_1 and η_1 and it is straightforward to solve for these variables in terms of q alone (through ψ_0 and η_0). This yields

$$\psi_1 = -(A \nabla^2)^{-1} \left[2bJ \left(\frac{\partial \psi_0}{\partial x}, \frac{\partial \psi_0}{\partial y} \right) + \nabla^2(\eta_0 q) \right], \quad (2.44)$$

$$\phi_1 = -b(A \nabla^2)^{-1} [J(\psi_0, \nabla^2 \psi_0)], \quad (2.45)$$

$$\eta_1 = -A^{-1} \left[2\sqrt{1-b^2} J \left(\frac{\partial \psi_0}{\partial x}, \frac{\partial \psi_0}{\partial y} \right) + \frac{b}{\sqrt{1-b^2}} \eta_0 q \right]. \quad (2.46)$$

Thus, from (2.24), the evolution equation for the balanced dynamics with $O(\epsilon)$ correction is now

$$\frac{\partial q}{\partial t} = -J(\psi_0, q) - \epsilon [\nabla \phi_1 \cdot \nabla q + J(\psi_1, q)], \quad (2.47)$$

with (2.30), (2.31) and (2.41)–(2.43).

As with the leading-order system we can now consider various special cases. In the QG limit the equations hold as written with b (i.e. B) regarded as an arbitrary parameter. Setting $b = 1/\sqrt{2}$ (i.e. $B = 1$), which amounts to a specific choice of the length scale L , our model reduces to (40)–(42) of Warn *et al.* (1995). (Note that their ϵ is equivalent to our R and $B = 1$ implies $R = \sqrt{2}\epsilon$ for our R and ϵ . With this caveat the equations are indeed identical.)

For $F \rightarrow 0$ with R fixed we have $\epsilon = Rb \rightarrow 0$ as $b \rightarrow 0$. From (2.42) we see that δ_1 satisfies Laplace's equation and, just as was argued for δ_0 , must also vanish. From (2.43) it follows that $\psi_1 = 0$ and there is no correction to the leading-order flow at this order. There is, however, a non-trivial correction to the surface displacement given by (2.41),

$$\nabla^2 \eta_1 = 2J \left(\frac{\partial \psi_0}{\partial x}, \frac{\partial \psi_0}{\partial y} \right), \quad (2.48)$$

and also, from (2.30) and (2.31), an $O(b)$ correction to η_0 , namely

$$\nabla^2 \eta_0 = bq. \quad (2.49)$$

Thus the $O(\epsilon)$ slow dynamics can be written

$$\frac{\partial q}{\partial t} = -J(\psi_0, q), \quad (2.50)$$

$$\nabla^2 \eta = \epsilon \frac{q}{R} + 2\epsilon J \left(\frac{\partial \psi_0}{\partial x}, \frac{\partial \psi_0}{\partial y} \right), \quad (2.51)$$

where $\eta = \eta_0 + \epsilon \eta_1$. The evolution equation remains the two-dimensional Euler equation, but now we have (ageostrophic) nonlinear Charney balance providing a slaving relation for the height field.

For $R \rightarrow 0$, (2.31) now implies an $O(\sqrt{1-b^2})$ correction to the leading-order velocity field, namely

$$-\nabla^2 \psi_0 + \frac{\epsilon}{F} \nabla^2 \eta_0 = 0. \quad (2.52)$$

Accordingly, to this order of approximation, (2.41) and (2.42) imply $\psi_1 = \phi_1 = 0$. Using (2.38) with (2.39) we obtain

$$\nabla^2 \psi_0 - F \eta_1 (1 + Fq) = 0 \quad (2.53)$$

and it is consistent to take $\eta_1 = 0$ at this order. We have now a non-zero velocity field but still trivial slow dynamics. (A non-trivial slow dynamics does emerge two orders higher.)

The system of equations (2.24)–(2.26) is indeed quite general. As noted by Warn *et al.* (1995), systems of this form allow balanced dynamics that can be extended to arbitrarily high order; in particular, the invertibility of the linear operators on the left-hand sides of (2.25) and (2.26) guarantees slaving relations to all orders. The main feature of the present analysis is that two theories of balance, previously considered disjoint, are shown to be special cases of one general theory.

3. Analysis of the Boussinesq equations

3.1. Fast and slow variables

We now consider the (non-hydrostatic) Boussinesq equations

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - f v + \frac{1}{\rho_s} \frac{\partial P}{\partial x} = 0, \quad (3.1)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + f u + \frac{1}{\rho_s} \frac{\partial P}{\partial y} = 0, \quad (3.2)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} - \frac{\Theta}{\theta_s} g + \frac{1}{\rho_s} \frac{\partial P}{\partial z} = 0, \quad (3.3)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (3.4)$$

$$\frac{\partial \Theta}{\partial t} + u \frac{\partial \Theta}{\partial x} + v \frac{\partial \Theta}{\partial y} + w \frac{\partial \Theta}{\partial z} = 0, \quad (3.5)$$

where (u, v, w) is the fluid velocity, f is the (constant) Coriolis parameter, g is the (constant) acceleration due to gravity, ρ_s is the (constant) reference density, θ_s is the (constant) reference potential temperature, Θ is the potential-temperature departure from θ_s and P is the pressure departure from a reference pressure. For simplicity, we again choose periodic boundary conditions on a domain D , except that now they will be applied to departures from a stably stratified rest state with potential temperature $\Theta_0(z)$ and pressure $P_0(z)$ in hydrostatic balance. The Brunt–Väisälä frequency $N = \sqrt{(g/\theta_s)d\Theta_0/dz}$ is assumed constant.

We proceed much as before. Linearizing about $(u, v, w, P, \Theta) = (0, 0, 0, P_0(z), \Theta_0(z))$ and assuming normal-mode solutions of the form

$$\{u, v, w, P - P_0, \Theta - \Theta_0\} = \{\tilde{u}, \tilde{v}, \tilde{w}, \tilde{P}, \tilde{\Theta}\} e^{i(kx + ly + mz - \omega t)}, \quad (3.6)$$

we obtain the dispersion relation

$$\omega^3 - \left(\frac{f^2 m^2 + N^2 \kappa^2}{\kappa^2 + m^2} \right) \omega = 0, \quad (3.7)$$

where again $\kappa = k^2 + l^2$. We find solutions

$$\omega_0 = 0, \quad \omega_{\pm} = \sqrt{\frac{f^2 m^2 + N^2 \kappa^2}{\kappa^2 + m^2}}. \quad (3.8)$$

Under a nonlinear frequency broadening by advection we have, to a good approximation,

$$\omega_0 \sim U\kappa, \quad \omega_{\pm} \sim U\kappa \pm \sqrt{\frac{f^2 m^2 + N^2 \kappa^2}{\kappa^2 + m^2}}, \quad (3.9)$$

where U is a characteristic flow speed. The ratio of the slow and fast time scales is

$$\frac{\omega_+}{\omega_0} \sim 1 + \frac{1}{\sqrt{1 + \alpha^2}} \frac{\sqrt{R^2 + F^2}}{RF}. \quad (3.10)$$

Here $R = U\kappa/f$ is the Rossby number, $F = Um/N$ is the Froude number for a stratified flow and $\alpha = \kappa/m$ is the aspect ratio of the length scales. Our condition for balance, i.e. a separation in time scales, holds if and only if

$$\epsilon = \frac{RF}{\sqrt{R^2 + F^2}} \ll \frac{1}{\sqrt{1 + \alpha^2}}. \quad (3.11)$$

Typically, in geophysical applications $\alpha \leq 1$. As ϵ is an asymptotic parameter we may replace the right-hand side of (3.11) with unity. Then this condition is identical to that for the shallow-water equations.

To continue we must non-dimensionalize the equations. We use the characteristic flow speed U , the horizontal length scale L , the vertical length scale H and the (horizontal) advective time scale $T = L/U$. For the potential temperature we use $H d\Theta_0/dz = H\theta_s N^2/g$, and for the pressure we assume a scaling Π , to be determined. Furthermore, we shall assume displacements from the background stratification $\Theta_0(z)$ by setting $\Theta = \Theta_0(z) + F\theta$ and $P = P_0(z) + Fp$. These perturbations are always small except in the PG limit when F is held fixed. Recall that this is the long-length-scale limit over which small perturbations may accumulate. As before, we prefer $b = B/\sqrt{1 + B^2} = F/\sqrt{R^2 + F^2}$ and ϵ instead of R and F for our dimensionless parameters; we have now

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - \frac{b}{\epsilon} v + \frac{\epsilon \Pi}{\sqrt{1 - b^2} \rho_s U^2} \frac{\partial p}{\partial x} = 0, \quad (3.12)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + \frac{b}{\epsilon} u + \frac{\epsilon \Pi}{\sqrt{1 - b^2} \rho_s U^2} \frac{\partial p}{\partial y} = 0, \quad (3.13)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} - \frac{\sqrt{1 - b^2}}{\alpha^2 \epsilon} \theta + \frac{\epsilon \Pi}{\alpha^2 \sqrt{1 - b^2} \rho_s U^2} \frac{\partial p}{\partial z} = 0, \quad (3.14)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (3.15)$$

$$\frac{\partial \theta}{\partial t} + u \frac{\partial \theta}{\partial x} + v \frac{\partial \theta}{\partial y} + w \frac{\partial \theta}{\partial z} + \frac{\sqrt{1 - b^2}}{\epsilon} w = 0. \quad (3.16)$$

Note that here Θ_0 is dimensionless and, by our choice of scaling, $d\Theta_0/dz = 1$. Observe also that hydrostatic balance emerges naturally upon taking the limit $\alpha \rightarrow 0$ with no assumption on ϵ .

We may now decide upon an appropriate value for Π . From the momentum equations it follows that

$$\Pi \sim \rho_s U^2 \max \left[\frac{\sqrt{1 - b^2}}{\epsilon}, \frac{b\sqrt{1 - b^2}}{\epsilon^2}, \frac{1 - b^2}{\epsilon^2} \right]. \quad (3.17)$$

Since we are interested in balanced dynamics for which $\epsilon \ll 1$ this suggests the scaling $\Pi = \sqrt{1-b^2}\rho_s U^2/\epsilon^2$. Thus we have

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - \frac{b}{\epsilon} v + \frac{1}{\epsilon} \frac{\partial p}{\partial x} = 0, \quad (3.18)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + \frac{b}{\epsilon} u + \frac{1}{\epsilon} \frac{\partial p}{\partial y} = 0, \quad (3.19)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} - \frac{\sqrt{1-b^2}}{\alpha^2 \epsilon} \theta + \frac{1}{\alpha^2 \epsilon} \frac{\partial p}{\partial z} = 0, \quad (3.20)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (3.21)$$

$$\frac{\partial \theta}{\partial t} + u \frac{\partial \theta}{\partial x} + v \frac{\partial \theta}{\partial y} + w \frac{\partial \theta}{\partial z} + \frac{\sqrt{1-b^2}}{\epsilon} w = 0. \quad (3.22)$$

It is worth remarking that as $\epsilon \rightarrow 0$ (3.20) yields

$$\sqrt{1-b^2} \theta = \frac{\partial p}{\partial z} \quad (3.23)$$

to leading order. Thus, unless $b \rightarrow 1$, balanced dynamics implies hydrostatic balance. No assumption of small aspect ratio ($\alpha \ll 1 \iff H \ll L$) is required.

Now we exploit the incompressibility constraint in order to reduce the number of equations. Formally, we shall consider w a functional of u and v and tacitly remember (3.21) as part of the system. As it is only $\partial w/\partial z$ that (3.21) determines, this necessarily introduces some ambiguity, and we shall occasionally refer back to the full set of governing equations when seeking solutions. Choosing dependent variables δ, ζ instead of u, v , we have

$$\frac{\partial \delta}{\partial t} - \frac{b\zeta - \nabla^2 p}{\epsilon} = -\mathbf{u} \cdot (\nabla \delta - \nabla^2 \mathbf{u}) - \frac{1}{2} \nabla^2 |\mathbf{u}|^2 + \zeta^2 - \frac{\partial \mathbf{u}}{\partial z} \cdot \nabla w - w \frac{\partial \delta}{\partial z}, \quad (3.24)$$

$$\frac{\partial \zeta}{\partial t} + \frac{b\delta}{\epsilon} = -\mathbf{u} \cdot \nabla \zeta - \delta \zeta - w \frac{\partial \zeta}{\partial z} + \frac{\partial}{\partial z} (\hat{\mathbf{z}} \times \mathbf{u}) \cdot \nabla w, \quad (3.25)$$

$$\frac{\partial \theta}{\partial t} + \frac{\sqrt{1-b^2}}{\epsilon} w = -\mathbf{u} \cdot \nabla \theta - w \frac{\partial \theta}{\partial z}, \quad (3.26)$$

where $\delta = \partial u/\partial x + \partial v/\partial y$ is the horizontal divergence, $\zeta = \partial v/\partial x - \partial u/\partial y$ is the vertical component of vorticity, $\mathbf{u} = (u, v)$ and $\nabla = (\partial/\partial x, \partial/\partial y)$. Again, as in the shallow-water analysis, we have a system with three prognostic variables. This is consistent with the third-order dispersion relation (3.7), which implies one slow and two fast variables. Pressure is a diagnostic variable and it is useful to have an explicit expression for it:

$$\mathcal{L} p = b\zeta + \frac{\sqrt{1-b^2}}{\alpha^2} \frac{\partial \theta}{\partial z} + \epsilon \left(\mathbf{u} \cdot \nabla^2 \mathbf{u} - \frac{1}{2} \nabla^2 |\mathbf{u}|^2 + \zeta^2 - 2 \frac{\partial \mathbf{u}}{\partial z} \cdot \nabla w - \delta^2 \right), \quad (3.27)$$

where $\mathcal{L} \equiv \nabla^2 + (1/\alpha^2)(\partial^2/\partial z^2)$.

As written, the equations naturally suggest fast variables δ and $\tilde{\Gamma} = b\zeta - \nabla^2 p$. Using (3.27) it follows that

$$\begin{aligned} \tilde{\Gamma} = b\zeta - \nabla^2 \mathcal{L}^{-1} \left[b\zeta + \frac{\sqrt{1-b^2}}{\alpha^2} \frac{\partial \theta}{\partial z} \right. \\ \left. + \epsilon \left(\mathbf{u} \cdot \nabla^2 \mathbf{u} - \frac{1}{2} \nabla^2 |\mathbf{u}|^2 + \zeta^2 - 2 \frac{\partial \mathbf{u}}{\partial z} \cdot \nabla w - \delta^2 \right) \right]. \end{aligned} \quad (3.28)$$

It is reasonable then to take

$$\begin{aligned} \Gamma \equiv \alpha^2 \mathcal{L} \left[b(1 - \nabla^2 \mathcal{L}^{-1}) \zeta - \frac{\sqrt{1-b^2}}{\alpha^2} \nabla^2 \mathcal{L}^{-1} \frac{\partial \theta}{\partial z} \right] \\ = b \frac{\partial^2 \zeta}{\partial z^2} - \sqrt{1-b^2} \nabla^2 \frac{\partial \theta}{\partial z} \end{aligned} \quad (3.29)$$

as our second fast variable, since (apart from application of the operator $\alpha^2 \mathcal{L}$) Γ and $\tilde{\Gamma}$ differ by a quantity of $O(\epsilon)$. Notice that in determining Γ we have replaced the pressure field using (3.27), an equation involving $\partial \theta / \partial z$ rather than θ itself. Necessarily, we have lost physical information and, as is the case for w , this is resolved by invoking the original equations when needed.

We have also an exact Lagrangian invariant, the Boussinesq potential vorticity

$$\begin{aligned} Q = \frac{\sqrt{1-b^2}}{R} + \sqrt{1-b^2} \zeta + b \frac{\partial \theta}{\partial z} \\ + \epsilon \left[\left(\alpha^2 \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) \frac{\partial \theta}{\partial x} + \left(\frac{\partial u}{\partial z} - \alpha^2 \frac{\partial w}{\partial x} \right) \frac{\partial \theta}{\partial y} + \zeta \frac{\partial \theta}{\partial z} \right], \end{aligned} \quad (3.30)$$

where we have non-dimensionalized with $U(\rho_s L \sqrt{1-b^2})^{-1} d\theta_0/dz$. Wishing to focus on the disturbance fields, we choose for the slow variable

$$\begin{aligned} q = Q - \frac{\sqrt{1-b^2}}{R} \\ = \sqrt{1-b^2} \zeta + b \frac{\partial \theta}{\partial z} + \epsilon \left[\left(\alpha^2 \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) \frac{\partial \theta}{\partial x} + \left(\frac{\partial u}{\partial z} - \alpha^2 \frac{\partial w}{\partial x} \right) \frac{\partial \theta}{\partial y} + \zeta \frac{\partial \theta}{\partial z} \right]. \end{aligned} \quad (3.31)$$

Our equations (3.24)–(3.26) can now be cast in the desired (Warn *et al.* 1995) form (2.11), (2.12) as follows:

$$\frac{\partial q}{\partial t} = -\mathbf{u} \cdot \nabla q - w \frac{\partial q}{\partial z}, \quad (3.32)$$

$$\begin{aligned} \frac{\partial \delta}{\partial t} - \frac{\mathcal{L}^{-1} \Gamma}{\alpha^2 \epsilon} = -\mathbf{u} \cdot (\nabla \delta - \nabla^2 \mathbf{u}) - \frac{1}{2} \nabla^2 |\mathbf{u}|^2 + \zeta^2 - \frac{\partial \mathbf{u}}{\partial z} \cdot \nabla w - w \frac{\partial \delta}{\partial z} \\ - \nabla^2 \mathcal{L}^{-1} \left(\mathbf{u} \cdot \nabla^2 \mathbf{u} - \frac{1}{2} \nabla^2 |\mathbf{u}|^2 + \zeta^2 - 2 \frac{\partial \mathbf{u}}{\partial z} \cdot \nabla w - \delta^2 \right), \end{aligned} \quad (3.33)$$

$$\begin{aligned} \frac{\partial \Gamma}{\partial t} + \frac{\mathcal{A} \delta}{\epsilon} = \sqrt{1-b^2} \frac{\partial}{\partial z} \nabla^2 \left(\mathbf{u} \cdot \nabla \theta + w \frac{\partial \theta}{\partial z} \right) \\ - b \frac{\partial^2}{\partial z^2} \left(\mathbf{u} \cdot \nabla \zeta + \delta \zeta + w \frac{\partial \zeta}{\partial z} - \frac{\partial}{\partial z} (\hat{\mathbf{z}} \times \mathbf{u}) \cdot \nabla w \right), \end{aligned} \quad (3.34)$$

where $\mathcal{A} = (1-b^2) \nabla^2 + b^2 \partial^2 / \partial z^2$ is an elliptic operator. Here b appears as a regular perturbation parameter; once again there is no singularity for $b \in (0, 1)$. Formally, we

allow for $b=0$ so as to include the possibility of no rotation. As in the shallow-water analysis, certain differential operators are not generally invertible; the ambiguity causes no problems and we write their inverses formally for convenience. This will be justified as needed.

3.2. Asymptotic models

Proceeding with the analysis, we expand all variables in ϵ except the slow variable q . The first-order system is simple:

$$q = \sqrt{1-b^2}\zeta_0 + b\frac{\partial\theta_0}{\partial z}, \quad (3.35)$$

$$\Gamma_0 = b\frac{\partial^2\zeta_0}{\partial z^2} - \sqrt{1-b^2}\nabla^2\frac{\partial\theta_0}{\partial z} = 0, \quad (3.36)$$

$$\mathcal{A}\delta_0 = \left[(1-b^2)\nabla^2 + b^2\frac{\partial^2}{\partial z^2}\right]\delta_0 = 0. \quad (3.37)$$

The last of these implies that δ_0 is constant. Integrating δ_0 over a horizontal cross-section and applying the periodic boundary conditions reveals this constant to be zero. Using $\nabla^2\psi_0 \equiv \zeta_0$ the first two equations yield

$$\nabla^2\psi_0 = \sqrt{1-b^2}\mathcal{A}^{-1}\nabla^2q, \quad (3.38)$$

$$\frac{\partial\theta_0}{\partial z} = b\mathcal{A}^{-1}\frac{\partial^2q}{\partial z^2}. \quad (3.39)$$

The inverse of $\mathcal{A}$ introduces constants of integration. Unique solutions are determined by the boundary conditions, namely

$$\int (1-b^2)\mathcal{A}^{-1}\nabla^2q \, dA = 0, \quad (3.40)$$

$$\int b^2\mathcal{A}^{-1}\frac{\partial^2q}{\partial z^2} \, dz = 0. \quad (3.41)$$

Here the first integral is over any horizontal cross-section, the second over any vertical column. From (3.20) we see that $\sqrt{1-b^2}\theta_0 = \partial p_0/\partial z$. With (3.38), (3.39) we can find p_0 (to within a constant) from (3.27) and so have θ_0 well determined. Now (3.22) shows that $w_0=0$ and so, by (3.21), the leading-order velocity field is two-dimensional and non-divergent. Thus $\mathbf{u}_0 = \hat{\mathbf{z}} \times \nabla\psi_0$ and the leading-order balanced dynamics is

$$\frac{\partial q}{\partial t} = -J(\psi_0, q), \quad (3.42)$$

with (3.38) or, equivalently,

$$q = \sqrt{1-b^2}\nabla^2\psi_0 + \frac{b^2}{\sqrt{1-b^2}}\frac{\partial^2\psi_0}{\partial z^2} \quad (3.43)$$

$$= \frac{1}{\sqrt{1+B^2}} \left(\nabla^2\psi_0 + B^2\frac{\partial^2\psi_0}{\partial z^2} \right). \quad (3.44)$$

This model is Hamiltonian. As with (2.35), (2.36) for the shallow-water system, the potential-vorticity evolution equation is identical to the QG equation. Note again, however, that no explicit restriction has been placed on R , only that ϵ be sufficiently small. In particular, geostrophic balance is not required. Thus, we have again shown that the QG potential vorticity equation – as distinct from QG dynamics – is valid more generally than its traditional derivation would suggest.

In the small- F limit $b = \epsilon/R \rightarrow 0$, or even if $b = 0$, we have

$$q = \zeta_0. \quad (3.45)$$

To this order (3.37) implies $\delta_0 = \delta_0(z, t)$. Integrating over a horizontal cross-section and applying the boundary conditions forces $\delta_0 = 0$ identically. Yet these equations are insufficient to determine θ_0 uniquely. In fact, θ_0 remains an undetermined periodic function of z .

To see this we appeal to the full set of equations (3.18)–(3.22) with $b = 0$. Clearly $\nabla p_0 = 0$, implying that p_0 is a function of z, t alone. Since

$$\theta_0 = \frac{\partial p_0}{\partial z}, \quad (3.46)$$

thus so also is θ_0 a function of z, t alone. Additionally, to leading order we find that $w_0 = 0$. For more information we require the $O(\epsilon)$ horizontal divergence and knowledge of the vertical velocity w_1 . From (3.34), (3.21) and the above we have

$$\mathcal{A} \delta_1 = \nabla^2 \delta_1 = \frac{\partial}{\partial z} \nabla^2 \left(\mathbf{u}_0 \cdot \nabla \theta_0 + w_0 \frac{\partial \theta_0}{\partial z} \right) = 0, \quad (3.47)$$

$$\delta_1 + \frac{\partial w_1}{\partial z} = 0. \quad (3.48)$$

Again the periodic boundary conditions force δ_1 to be identically zero. It follows that w_1 depends only on x, y, t . Now (3.22) becomes

$$\frac{\partial \theta_0}{\partial t}(z, t) + w_1(x, y, t) = 0, \quad (3.49)$$

and it follows that both $\partial \theta_0 / \partial t$ and w_1 are functions of t alone. Returning to (3.46), we differentiate once with respect to time, integrate vertically and invoke the periodic boundary conditions to conclude that θ_0 is independent of time.

Therefore $\theta_0 = \theta_0(z)$ and otherwise goes undetermined, except that it must be periodic and satisfy (3.46). We may consider it part of the initial conditions, the horizontal average of θ at $t = 0$. However, it plays no role in the dynamics at this order. The leading-order flow is two-dimensional and we may write $\mathbf{u}_0 = \hat{\mathbf{z}} \times \nabla \psi_0$. We have then the leading-order balanced dynamics

$$\frac{\partial q}{\partial t} = -J(\psi_0, q), \quad (3.50)$$

with (3.45). The slow dynamics decouples in the vertical into two-dimensional Euler dynamics for each value of z (e.g. Vallis 1996).

For the small- R limit, $b = \sqrt{1 - (\epsilon/F)^2} \rightarrow 1$; now

$$q = \frac{\partial \theta_0}{\partial z}. \quad (3.51)$$

From (3.36), (3.37) and the periodic boundary conditions we readily find that both δ_0 and ζ_0 are independent of z . With (3.21) it follows that δ_0 must vanish. From (3.18)–(3.20) we obtain $\hat{\mathbf{z}} \times (\mathbf{u}_0, v_0) + \nabla p_0 = 0$ and that p_0 is independent of z . This is exactly the situation of the Taylor–Proudman theorem, which tells us to expect horizontal components of velocity independent of z and a horizontal divergence equal to zero. In the classical exposition, rigid-boundary conditions force the vertical component of velocity to be equal to zero throughout the flow. In our case we have no information about w_0 except that it depends on x, y, t only. The fluid moves in columns, each with its own vertical velocity profile.

More can be said, however, if we are allowed to assume a small aspect ratio. For example, suppose $f \lesssim N$; this is the regime relevant to large-scale geophysical flows. Then $R \rightarrow 0$ with all else fixed is certainly the long-wave limit and we have

$$\alpha = \frac{H}{L} = \frac{R}{F} \frac{f}{N} \lesssim \frac{\epsilon}{F} \rightarrow 0. \quad (3.52)$$

From the governing equations we can now determine w_0 . We have

$$\frac{\theta_0}{F} - \frac{\partial p_1}{\partial z} = 0, \quad (3.53)$$

$$\frac{\partial w_0}{\partial z} = 0, \quad (3.54)$$

$$\frac{\partial \theta_0}{\partial t} + u_0 \frac{\partial \theta_0}{\partial x} + v_0 \frac{\partial \theta_0}{\partial y} + w_0 \left(q + \frac{1}{F} \right) = 0. \quad (3.55)$$

The first of these implies the z -average of θ_0 to be zero. For a statically stable fluid we are assured that $q + 1/F \neq 0$ and, since u_0, v_0, w_0 are independent of z , we can integrate the last to conclude that $w_0 = 0$ identically.

Continuing the general analysis to the next order, (3.31), (3.33) and (3.34) together with previous work give

$$\sqrt{1-b^2} \zeta_1 + b \frac{\partial \theta_1}{\partial z} = \nabla \left(\frac{\partial \psi_0}{\partial z} \right) \cdot \nabla \theta_0 - \nabla^2 \psi_0 \frac{\partial \theta_0}{\partial z}, \quad (3.56)$$

$$-b \frac{\partial^2 \zeta_1}{\partial z^2} + \sqrt{1-b^2} \nabla^2 \frac{\partial \theta_1}{\partial z} = \frac{\partial^2}{\partial z^2} \left(\nabla \psi_0 \cdot \nabla (\nabla^2 \psi_0) - \frac{1}{2} \nabla^2 |\nabla \psi_0|^2 + (\nabla^2 \psi_0)^2 \right), \quad (3.57)$$

$$\left((1-b^2) \nabla^2 + b^2 \frac{\partial^2}{\partial z^2} \right) \delta_1 = \sqrt{1-b^2} \frac{\partial}{\partial z} \nabla^2 J(\psi_0, \theta_0) - b \frac{\partial^2}{\partial z^2} \left(J(\psi_0, \nabla^2 \psi_0) \right). \quad (3.58)$$

These equations are linear in $\zeta_1, \theta_1, \delta_1$ and straightforward to solve; using $\nabla^2 \psi_1 \equiv \zeta_1$ and $\nabla^2 \phi_1 \equiv \delta_1$, we find

$$\begin{aligned} \nabla^2 \psi_1 = \mathcal{A}^{-1} \left[\sqrt{1-b^2} \nabla^2 \left(\nabla \left(\frac{\partial \psi_0}{\partial z} \right) \cdot \nabla \theta_0 - \nabla^2 \psi_0 \frac{\partial \theta_0}{\partial z} \right) \right. \\ \left. - b \frac{\partial^2}{\partial z^2} \left(\nabla \psi_0 \cdot \nabla (\nabla^2 \psi_0) - \frac{1}{2} \nabla^2 |\nabla \psi_0|^2 + (\nabla^2 \psi_0)^2 \right) \right], \end{aligned} \quad (3.59)$$

$$\begin{aligned} \frac{\partial \theta_1}{\partial z} = \mathcal{A}^{-1} \frac{\partial^2}{\partial z^2} \left[b \left(\nabla \left(\frac{\partial \psi_0}{\partial z} \right) \cdot \nabla \theta_0 - \nabla^2 \psi_0 \frac{\partial \theta_0}{\partial z} \right) \right. \\ \left. - \sqrt{1-b^2} \frac{\partial^2}{\partial z^2} \left(\nabla \psi_0 \cdot \nabla (\nabla^2 \psi_0) - \frac{1}{2} \nabla^2 |\nabla \psi_0|^2 + (\nabla^2 \psi_0)^2 \right) \right], \end{aligned} \quad (3.60)$$

$$\nabla^2 \phi_1 = \mathcal{A}^{-1} \left[\sqrt{1-b^2} \frac{\partial}{\partial z} \nabla^2 J(\psi_0, \theta_0) - b \frac{\partial^2}{\partial z^2} J(\psi_0, \nabla^2 \psi_0) \right]. \quad (3.61)$$

The apparent ambiguity in θ_1 is resolved using (3.27) and then (3.20), exactly as was done for θ_0 . We determine the vertical velocity field w_1 from (3.22). Our $O(\epsilon)$ slow

dynamics is now

$$\frac{\partial q}{\partial t} = -J(\psi_0, q) - \epsilon \left[\nabla \phi_1 \cdot \nabla q + J(\psi_1, q) + w_1 \frac{\partial q}{\partial z} \right], \quad (3.62)$$

with (3.35)–(3.37) and (3.56)–(3.58). This equation is quite general and, if so desired, could be extended to arbitrarily high order.

Both small-Rossby-number and small-Froude-number balance regimes are found by considering special limits for b . We simply note that, for the small- F limit, (3.56) is replaced with

$$\zeta_1 + q \frac{\partial \theta_0}{\partial z} - \nabla \left(\frac{\partial \psi_0}{\partial z} \right) \cdot \nabla \theta_0 = 0, \quad (3.63)$$

which determines ζ_1 , since at this order θ_0 is slaved to q by way of (3.36). Neglecting terms of $O(b)$, (3.57) and (3.58) reduce to linear partial differential equations for θ_1 and δ_1 , with everything on the right-hand sides known. In the limit of small R , (3.56) becomes

$$q \zeta_0 + \frac{\partial \theta_1}{\partial z} - \nabla \left(\frac{\partial \psi_0}{\partial z} \right) \cdot \nabla \theta_0 = 0, \quad (3.64)$$

which determines θ_1 ; now it is ζ_0 that is slaved to q through (3.36). Neglecting terms of $O(\sqrt{1-b^2})$, (3.57) and (3.58) reduce to linear partial differential equations for ζ_1 and δ_1 .

Bokhove (1997) applied the same approach to the hydrostatic Boussinesq equations on an f -plane but using the Rossby number for his asymptotic parameter and assuming QG scaling. Furthermore, he considered a geometry unbounded in the horizontal and with rigid vertical boundaries. The restriction to QG scaling permits an $O(1)$ variation in $N(z)$, something we cannot allow because we are making explicit use of the Froude number. In the QG limit, our analysis implies hydrostatic balance whereas Bokhove assumes it from the outset; the resulting systems of balanced equations differ only in that Bokhove chooses a different slow variable, the QG (or linearized) potential vorticity. In consequence, with b now held fixed, the right-hand side of (3.56) is replaced with zero while (3.57), (3.58) remain unchanged, and so his slow-dynamics equation is not simply pure advection.

4. Conclusions

Our purpose here has been to consider balanced dynamics from a general point of view. Rather than focus on specific equations we have asked what is required for a system evolving in time to remain balanced. Our answer is physically motivated: the dominant frequencies of motion should split into two groups, fast and slow, and there should be little interaction between them – the point being to limit the potential for frequency matching through nonlinearity. With this idea we have determined a dimensionless quantity which, in some sense, measures the strength of possible coupling between the fast and slow motions. Using this as our asymptotic parameter, we have found balanced-dynamics equations that are quite general. Our approach unifies the notions of small-Rossby-number and small-Froude-number balance, and, of course, preserves the Hamiltonian structure at leading order.

For the shallow-water equations and Boussinesq equations on an f -plane, we have learned that a time-scale separation in the dynamics is a necessary and sufficient condition for all classic forms of balance based on $R \ll 1$ or $F \ll 1$. Moreover, for the

Boussinesq equations without any assumption on the aspect ratio, balanced dynamics generally implies hydrostatic balance. Only when the Froude number is not small and it is the Rossby number that guarantees a time-scale separation must we impose a small aspect ratio to ensure hydrostatic balance.

The analysis has been presented on an f -plane for simplicity only. Including the effects of a variable Coriolis parameter on a (mid-latitude) β -plane is straightforward. The slow-motion solution of the dispersion relation is no longer trivial. Now we have Rossby waves with frequency $\omega_R \sim \beta L \sim fL/a$, where a is the radius of the earth. Since the inertia-gravity-wave frequency ω_G is bounded from below by f , we have a time-scale separation $\omega_R/\omega_G \lesssim L/a \ll 1$ provided that L/a is sufficiently small. On the β -plane L/a is an independent dimensionless parameter. If the vortical frequency $\omega_V = U/L$, we have $\omega_R/\omega_V \sim \beta L^2/U \sim (L/a)/R$ and this ratio may be small, large or otherwise. While the situation is somewhat more complicated than the analysis above, the separation of time scales can still be made to hold. In the tropics the situation is quite different, for there the Rossby-wave frequency cannot be bounded away from that of the fast motion; $\omega_R \lesssim \sqrt{\beta c}/2$ but the fast-motion frequency cutoff is $\sqrt{\beta c}$, where c is the gravity-wave speed.

For the Boussinesq equations we chose to assume a constant Brunt–Väisälä frequency $N(z) = N$ for the background state. Evidently there is no simple physical parameter to control this term if we relax this assumption. Nevertheless, the present variables suffice to distinguish fast and slow motions if we include in our hypothesis that N be slowly varying, i.e. $dN/dz \sim \epsilon$. The equations are modified accordingly and the analysis follows through.

A restriction of our approach is to assume a non-zero stratification and, in particular, $f \lesssim N$. Associated with this is the assumption $\alpha \leq 1$. This is appropriate for large-scale geophysical flows. In the case of unstratified rotating flow, where a Taylor–Proudman-like balance is possible for which $\alpha \gg 1$, a modified analysis would be required.

Another restriction is the use of periodic boundary conditions. Rigid boundaries allow additional slow dynamics (Bokhove 2000; Ren & Shepherd 1997), while unbounded domains allow Lighthill radiation (Ford, McIntyre & Norton 2000). Again, for these cases, a modified analysis would be required.

Finding balanced models uniformly valid on the sphere would seem to be the final frontier; the complications that arise near the equator are quite challenging. The technique presented here, however, has met with some success. A full discussion of the equatorial β -plane equations appears in Saujani (2005).

REFERENCES

- BOKHOVE, O. 1997 Slaving principles, balanced dynamics, and the hydrostatic Boussinesq equations. *J. Atmos. Sci.* **54**, 1662–1674.
- BOKHOVE, O. 2000 On hydrostatic flows in isentropic coordinates. *J. Fluid Mech.* **402**, 291–310.
- CHARNEY, J. G. 1948 On the scale of atmospheric motions. *Geophys. Publ.* **17**, No. 2, 251–265.
- CHARNEY, J. G. 1963 A note on large scale motions in the tropics. *J. Atmos. Sci.* **20**, 607–609.
- DALEY, R. 1991 *Atmospheric Data Analysis*. Cambridge University Press.
- ERRICO, R. M. 1982 Normal mode initialization and the generation of gravity waves by quasi-geostrophic forcing. *J. Atmos. Sci.* **39**, 573–586.
- FORD, R., MCINTYRE, M. E. & NORTON, W. A. 2000 Balance and the slow quasimanifold: some explicit results. *J. Atmos. Sci.* **57**, 1236–1234.
- VAN KAMPEN, N. G. 1985 Elimination of fast variables. *Phys. Rep.* **124**, 69–160.

- LILLY, D. K. 1983 Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.* **40**, 746–761.
- MCINTYRE, M. E. & NORTON, W. A. 2000 Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.* **57**, 1214–1235.
- MCWILLIAMS, J. C. 1985 A uniformly valid model spanning the regimes of geostrophic and isotropic, stratified turbulence: balanced turbulence. *J. Atmos. Sci.* **42**, 1773–1774.
- PEDLOSKY, J. 1987 *Geophysical Fluid Dynamics*, 2nd edn. Springer.
- REN, S. & SHEPHERD, T. G. 1997 Lateral boundary contributions to wave-activity invariants and nonlinear stability theorems for balanced dynamics. *J. Fluid Mech.* **345**, 287–305.
- SALMON, R. 1983 Practical use of Hamilton's principle. *J. Fluid Mech.* **132**, 431–444.
- SAUJANI, S. & SHEPHERD, T. G. 2002 Comments on 'Balance and the slow quasimanifold: some explicit results.' *J. Atmos. Sci.* **59**, 2874–2877.
- SAUJANI, S. 2005 Towards a unified theory of balanced dynamics. PhD thesis, Department of Physics, University of Toronto, 71 pp.
- SHEPHERD, T. G. 2003 Hamiltonian dynamics. In *Encyclopedia of Atmospheric Sciences* (ed. J. R. Holton *et al.*), pp. 929–939. Academic.
- VALLIS, G. K. 1996 Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Q. J. R. Met. Soc.* **122**, 291–322.
- WARN, T. 1997 Nonlinear balance and quasi-geostrophic sets. *Atmos.-Ocean* **35**, 135–145.
- WARN, T., BOKHOVE, O., SHEPHERD, T. G. & VALLIS, G. K. 1995 Rossby number expansions, slaving principles and balance dynamics. *Q. J. R. Met. Soc.* **121**, 723–739.
- WIROSOETISNO, D. & SHEPHERD, T. G. 2000 Averaging, slaving and balance dynamics in a simple atmospheric model. *Physica D* **141**, 37–53.
- WIROSOETISNO, D., SHEPHERD, T. G. & TEMAM, R. M. 2002 Free gravity waves and balanced dynamics. *J. Atmos. Sci.* **59**, 3382–3398.