Variational treatment of inertia-gravity waves interacting with a quasigeostrophic mean flow

By Rick Salmon

Scripps Institution of Oceanography, University of California San Diego
La Jolla CA 92093-0213, USA email address: rsalmon@ucsd.edu

(Received 24 August 2016)

The equations for three-dimensional, hydrostatic, Boussinesq dynamics are equivalent to a variational principle that is closely analogous to the variational principle for classical electrodynamics. Inertia-gravity waves are analogous to electromagnetic waves, and available potential vorticity (i.e. the amount by which the potential vorticity exceeds the potential vorticity of the rest state) is analogous to electric charge. The Lagrangian can be expressed as the sum of three parts. The first part corresponds to quasigeostrophic dynamics in the absence of inertia-gravity waves. The second part corresponds to inertia-gravity waves in the absence of quasigeostrophic flow. The third part represents a coupling between the inertial-gravity waves and quasigeostrophic motion. This formulation provides the basis for a general theory of inertia-gravity waves interacting with a quasigeostrophic mean flow.

1. Introduction

This paper offers a new variational principle that seems especially useful for studying the interactions between inertia-gravity waves and quasigeostrophic flow. The variational principle is developed for shallow water dynamics in Section 2 and for hydrostatic Boussinesq dynamics in Section 3. The latter is of primary interest. The Lagrangian for Boussinesq dynamics,

$$L = L_1[\psi, \gamma] + L_2[\psi, \gamma, \alpha, \beta], \tag{1.1}$$

depends on the four fields ψ, γ, α, and β. These have the following physical interpretations: ψ is the streamfunction and γ_t is the velocity potential corresponding to the 'thickness weighted' velocity, i.e. the velocity times the local isopycnal separation. α and β are particle labels that track and measure the available potential vorticity (APV), defined as the difference between the potential vorticity and its value in the rest state. Variations of ψ yield the equation that relates the velocity to the APV. Variations of γ yield the divergence equation. Variations of α and β yield the equation for the advection of APV. The importance of APV to wave-mean theory is strongly emphasized by Wagner & Young (2015).

If the flow is initially at rest then the APV vanishes and $L_2 \equiv 0$ does not contribute. In Section 4 we use $\delta L_1 = 0$ to calculate the $\mathrm{O}(a^2)$ mean flow that arises when inertia-gravity waves of amplitude a propagate into a region that is initially at rest. This is the classic problem treated by Bretherton (1969). If the phase-averaged wave amplitude varies slowly enough, we obtain (4.31), in which the curl of the pseudomomentum represents a source of quasigeostrophic potential vorticity.

Section 5 considers the case of nonvanishing APV. In this case L_2 does not vanish, and in Section 5 we develop approximations to L_2 that are analogous to those developed for L_1 in Section 4. Section 6 combines the results of Sections 4 and 5 to obtain a Lagrangian (6.12) that depends on the streamfunction $\bar{\psi}$ for the Lagrangian mean velocity, two variables $\tilde{\alpha}$ and $\tilde{\beta}$ that track and label the available potential vorticity $\tilde{Q}$ that would occur if the APV were advected by the mean flow alone, and the fluid particle displacements (ξ, η) caused by the waves. The Lagrangian (6.12) is our primary result. It consists of three parts. The first part, (6.13), corresponds to quasigeostrophic motion in the absence of inertia-gravity waves. The second part, (6.14), corresponds to inertia-gravity waves in the absence of a mean flow. The third part, (6.15), represents a coupling between the inertia-gravity waves and the quasigeostrophic mean flow.

The resulting dynamics always includes the equation (7.3) governing the advection of $\tilde{Q}$ by the Lagrangian mean streamfunction $\bar{\psi}$, but the equation (7.31) relating $\tilde{Q}$ to $\bar{\psi}$ contains a wave contribution q^w whose precise form depends on the strength of our assumptions about the wave field. The weakest such assumption is that the waves and the mean flow are separated only by their timescales. This was the assumption made by Wagner & Young (2015), and in Section 7 and Appendix B we show that our q^w is equivalent to theirs.

A somewhat stronger assumption is that the waves and the mean flow are separated both in timescale and in vertical lengthscale. This is the case for near-inertial waves, which oscillate rapidly in t and in z, but may have horizontal lengthscales comparable to those of the quasigeostrophic mean flow. Near-inertial motion was analyzed by Young & Ben Jelloul (1997) and Xie & Vanneste (2015), and in Section 7 we obtain wave-mean equations closely related to theirs.

The strongest assumption is the 'WKB' assumption that the waves and the mean flow have a scale separation with respect to all four of the independent variables x, y, z, t. In this case we recover results similar to those of Bühler & McIntyre (1998).

Our ability to recover results such as these, which were obtained by powerful methods that seem to have little in common, hints at the synthetic potential of our method.

Throughout the present paper we refer to an analogy between our formulation and the standard Lagrangian formulation of classical electrodynamics. This analogy, noted by Salmon (2014, hereafter S14), largely inspired the present work. S14 treated a simplified shallow water dynamics (for which the analogy is stronger) and did not consider coordinate-system rotation. In S14, the potential vorticity was assumed to be concentrated in point vortices, which are analogous to electrons. In the present paper we relax all the unrealistic assumptions of S14. (However, we continue to assume that the Coriolis parameter is a constant.) The electrodynamic analogy still seems quite useful, and it is explained more fully in Appendix A. However, readers who find it unhelpful, or even annoying, are invited to ignore it completely. The present paper is entirely understandable in its own terms.

2. Shallow water dynamics

The shallow water equations are

$$\mathbf{u}_t + \nabla(c^2\hat{h} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}) = (\zeta + f_0)(v, -u), \tag{2.1}$$

$$\hat{h}_t + \nabla\cdot(\hat{h}\mathbf{u}) = 0, \tag{2.2}$$

where $\mathbf{u}(\mathbf{x}, t) = (u, v)$ is the fluid velocity at location $\mathbf{x} = (x, y)$ and time t, $\hat{h} = h/h_0$ is the depth divided by its constant mean value h_0, $c^2 = gh_0$ with gravity constant g,

$\zeta = \nabla \times \mathbf{u} \equiv v_x - u_y$ is the vertical component of vorticity, and f_0 is the constant Coriolis parameter. Subscripts denote partial derivatives and $\nabla = (\partial_x, \partial_y)$. For definiteness, we suppose that the fluid is horizontally unbounded and quiescent at infinity. The vorticity and divergence equations corresponding to (2.1)-(2.2) are

$$\zeta_t = -\nabla \cdot ((\zeta + f_0)\mathbf{u}) \tag{2.3}$$

and

$$\frac{\partial}{\partial t}\nabla \cdot \mathbf{u} + \nabla^2(c^2\hat{h} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}) = \nabla \times ((\zeta + f_0)\mathbf{u}). \tag{2.4}$$

The shallow water equations conserve energy in the form,

$$H = \iint d\mathbf{x}\ \tfrac{1}{2}(h\mathbf{u}\cdot\mathbf{u} + gh^2) = h_0 \iint d\mathbf{x}\ \tfrac{1}{2}(\hat{h}\mathbf{u}\cdot\mathbf{u} + c^2\hat{h}^2), \tag{2.5}$$

and the potential vorticity,

$$q \equiv (\zeta + f_0)/\hat{h}, \tag{2.6}$$

on fluid particles,

$$q_t + \mathbf{u}\cdot\nabla q = 0. \tag{2.7}$$

In this section, we generalize the variational principle of S14 to full shallow water dynamics (2.1)-(2.2) with continuously distributed vorticity and constant Coriolis parameter f_0. We start by representing the physical variables

$$\hat{h} = 1 - \nabla^2\gamma, \tag{2.8}$$

$$\hat{h}u = -\psi_y + \gamma_{xt}, \tag{2.9}$$

$$\hat{h}v = \psi_x + \gamma_{yt}, \tag{2.10}$$

in terms of the potentials $\psi(\mathbf{x}, t)$ and $\gamma(\mathbf{x}, t)$. Compared to the method followed in S14, the representation (2.8)-(2.10) corresponds to the immediate adoption of the Coulomb gauge, as explained in Appendix A. For present purposes, it is sufficient to note that (2.8)-(2.10) is a general representation that automatically satisfies the shallow-water continuity equation (2.2).

We motivate the derivation of the variational principle by first imposing—and then gradually lifting—a strong constraint: We assume that the potential vorticity (2.6) is uniform. If the potential vorticity is uniform, then

$$\zeta + f_0 = f_0\hat{h}. \tag{2.11}$$

By (2.8)-(2.10), (2.11) is equivalent to

$$\frac{\partial}{\partial x}\left(\frac{\psi_x + \gamma_{yt}}{1 - \nabla^2\gamma}\right) + \frac{\partial}{\partial y}\left(\frac{\psi_y - \gamma_{xt}}{1 - \nabla^2\gamma}\right) = -f_0\nabla^2\gamma. \tag{2.12}$$

Both (2.12) and the corresponding form of (2.4), namely,

$$\left(\frac{-\psi_y + \gamma_{xt}}{1 - \nabla^2\gamma}\right)_{xt} + \left(\frac{\psi_x + \gamma_{yt}}{1 - \nabla^2\gamma}\right)_{yt} + \nabla^2\left(c^2\hat{h} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) = f_0\nabla^2\psi, \tag{2.13}$$

result from the variational principle $\delta L_1[\psi, \gamma] = 0$ where

$$L_1 = \iiint d\mathbf{x}dt\ \left(\tfrac{1}{2}\frac{(-\psi_y + \gamma_{xt})^2}{(1 - \nabla^2\gamma)} + \tfrac{1}{2}\frac{(\psi_x + \gamma_{yt})^2}{(1 - \nabla^2\gamma)} + f_0\nabla\psi\cdot\nabla\gamma - \tfrac{1}{2}c^2(\nabla^2\gamma)^2\right). \tag{2.14}$$

Now we partly relax the constraint: We allow the potential vorticity to be non-uniform,

but we assume that the departures from (2.11) are concentrated in delta functions. That is, we assume that

$$\zeta + f_0 = f_0\hat{h} + \sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)). \tag{2.15}$$

This is a consistent restriction on shallow water dynamics in the sense that (2.15) is compatible with the conservation of total vorticity, $\iint d\mathbf{x}\ hq$. The sum in (2.15) is the product of the fluid depth $\hat{h}$ and what Wagner & Young (2015) call the 'available potential vorticity'—the difference between the potential vorticity defined by (2.6) and its uniform value in the state of rest. This difference is, like the potential vorticity itself, conserved on fluid particles. In a slight departure the terminology of Wagner & Young (2015), we shall refer to the sum itself as the available potential vorticity (APV). This sum obeys the flux-form conservation law (2.40).

Since the delta functions in (2.15) represent singularities in potential vorticity, they must, according to (2.7), move at the fluid velocity $\mathbf{u}$. That is, the velocities of the point vortices must be given by $\dot{\mathbf{x}}_i(t) = \mathbf{u}(\mathbf{x}_i(t), t)$, or, using (2.8)-(2.10),

$$(1 - \nabla^2\gamma)\dot{\mathbf{x}}_i(t) = (-\psi_y + \gamma_{tx},\ \psi_x + \gamma_{ty}), \tag{2.16}$$

where the overdot denotes time differentiation. The Lagrangian corresponding to (2.16) is

$$L_2[\mathbf{x}_i(t)] = \int dt \sum_i \Gamma_i \left(-x_i\dot{y}_i + \psi(\mathbf{x}_i, t) - \gamma_y(\mathbf{x}_i, t)\dot{x}_i + \gamma_x(\mathbf{x}_i, t)\dot{y}_i\right) \tag{2.17}$$

That is, $\delta L_2/\delta \mathbf{x}_i = 0$ implies (2.16).

We obtain the Lagrangian for shallow water dynamics by combining (2.14) and (2.17) in the form

$$\begin{aligned}
&L[\psi(\mathbf{x}, t), \gamma(\mathbf{x}, t), \mathbf{x}_i(t)] \\
&= L_1[\psi(\mathbf{x}, t), \gamma(\mathbf{x}, t)] + L_2[\psi(\mathbf{x}, t), \gamma(\mathbf{x}, t), \mathbf{x}_i(t)] \\
&= L_1[\psi(\mathbf{x}, t), \gamma(\mathbf{x}, t)] + \\
&\iiint dt d\mathbf{x} \sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \left(-x\dot{y}_i + \psi(\mathbf{x}, t) - \gamma_y(\mathbf{x}, t)\dot{x}_i + \gamma_x(\mathbf{x}, t)\dot{y}_i\right).
\end{aligned} \tag{2.18}$$

As we shall show, shallow water dynamics with point vortices of APV is equivalent to the requirement that (2.18) be stationary with respect to arbitrary independent variations of the fields $\psi(\mathbf{x}, t)$ and $\gamma(\mathbf{x}, t)$ and the vortex locations $\mathbf{x}_i(t)$. The last line of (2.18) is an alternative way of writing (2.17) that is useful for performing the field variations. The form (2.17) is preferable for performing the vortex-location variations. Before demonstrating the equivalence between (2.18) and shallow water dynamics, we fully relax the constraint on potential vorticity.

Suppose that the APV is not confined to delta functions but varies continuously in space. It must still be advected at the fluid velocity $\mathbf{u}$. We replace the ansatz (2.15) by

$$\zeta + f_0 = f_0\hat{h} + \frac{\partial(\alpha, \beta)}{\partial(x, y)}, \tag{2.19}$$

where $\alpha(\mathbf{x}, t)$ and $\beta(\mathbf{x}, t)$ are a set of fluid particle labels that measure APV and move with the fluid. Thus,

$$(1 - \nabla^2\gamma)\boldsymbol{\alpha}_t + (-\psi_y + \gamma_{tx})\boldsymbol{\alpha}_x + (\psi_x + \gamma_{ty})\boldsymbol{\alpha}_y = 0, \tag{2.20}$$

where $\boldsymbol{\alpha} = (\alpha, \beta)$. Compare (2.20) to (2.16). The labels α and β replace the subscript i

on Γ_i in (2.15). Just as the last term in (2.15) integrates to $\sum_i \Gamma_i$, the last term in (2.19) integrates to $\iint d\alpha d\beta$; the total APV is a constant.

We obtain the general variational principle for shallow water dynamics by replacing

$$\sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \to \frac{\partial(\alpha,\beta)}{\partial(x,y)} \tag{2.21}$$

and $\dot{\mathbf{x}}_i \to \partial\mathbf{x}/\partial\tau$ in (2.18), where $\mathbf{x}(\alpha,\beta,\tau)$ is the location at time τ of the fluid particle labeled (α,β). Then

$$L_2 = \iiint dt d\mathbf{x} \frac{\partial(\alpha,\beta)}{\partial(x,y)} \left(-x\frac{\partial y}{\partial\tau} + \psi(\mathbf{x},t) - \gamma_y(\mathbf{x},t)\frac{\partial x}{\partial\tau} + \gamma_x(\mathbf{x},t)\frac{\partial y}{\partial\tau} \right). \tag{2.22}$$

By the chain rule,

$$\frac{\partial}{\partial\tau}\boldsymbol{\alpha}(\mathbf{x},t) = \frac{\partial\boldsymbol{\alpha}}{\partial t} + \left(\frac{\partial\mathbf{x}}{\partial\tau}\cdot\nabla\right)\boldsymbol{\alpha} = 0. \tag{2.23}$$

Solving (2.23) for $\partial\mathbf{x}/\partial\tau$ we obtain

$$\frac{\partial(\alpha,\beta)}{\partial(x,y)}\frac{\partial x}{\partial\tau} = \frac{\partial(\alpha,\beta)}{\partial(y,t)}, \tag{2.24}$$

$$\frac{\partial(\alpha,\beta)}{\partial(x,y)}\frac{\partial y}{\partial\tau} = \frac{\partial(\alpha,\beta)}{\partial(t,x)}. \tag{2.25}$$

Then using (2.24)-(2.25) to eliminate $\partial x/\partial\tau$ and $\partial y/\partial\tau$ from (2.22), and noting that the first term in (2.22) simplifies as

$$\begin{aligned}\iiint dt d\mathbf{x} \frac{\partial(\alpha,\beta)}{\partial(x,y)} \left(-x\frac{\partial y}{\partial\tau}\right) &= \iiint dt d\mathbf{x} \left(-x\frac{\partial(\alpha,\beta)}{\partial(t,x)}\right) \\ &= \iiint dt d\mathbf{x} \left(\alpha\frac{\partial(x,\beta)}{\partial(t,x)}\right) = \iiint dt d\mathbf{x}\ (-\alpha\beta_t),\end{aligned} \tag{2.26}$$

we obtain the generalization of (2.18) to the case of continuously varying vorticity, namely

$$\begin{aligned}&L[\psi(\mathbf{x},t),\gamma(\mathbf{x},t),\boldsymbol{\alpha}(\mathbf{x},t)] = L_1[\psi(\mathbf{x},t),\gamma(\mathbf{x},t)] + L_2[\psi(\mathbf{x},t),\gamma(\mathbf{x},t),\alpha(\mathbf{x},t),\beta(\mathbf{x},t)] \\ &= \iiint d\mathbf{x}dt\ \left(\tfrac{1}{2}\frac{(-\psi_y+\gamma_{xt})^2}{(1-\nabla^2\gamma)} + \tfrac{1}{2}\frac{(\psi_x+\gamma_{yt})^2}{(1-\nabla^2\gamma)} + f_0\nabla\psi\cdot\nabla\gamma - \tfrac{1}{2}c^2(\nabla^2\gamma)^2\right) \\ &+ \iiint dt d\mathbf{x} \left(-\alpha\beta_t + \psi\frac{\partial(\alpha,\beta)}{\partial(x,y)} - \gamma_y\frac{\partial(\alpha,\beta)}{\partial(y,t)} + \gamma_x\frac{\partial(\alpha,\beta)}{\partial(t,x)}\right).\end{aligned} \tag{2.27}$$

We shall demonstrate that the requirement that (2.27) be stationary with respect to arbitrary variations in its four fields is equivalent to shallow water dynamics.

The equations resulting from the requirement that (2.27) be stationary are:

$$\delta\psi: \quad \frac{\partial}{\partial x}\left(\frac{\psi_x+\gamma_{yt}}{1-\nabla^2\gamma}\right) + \frac{\partial}{\partial y}\left(\frac{\psi_y-\gamma_{xt}}{1-\nabla^2\gamma}\right) + f_0\nabla^2\gamma = Q, \tag{2.28}$$

$$\delta\gamma: \quad \left(\frac{-\psi_y+\gamma_{xt}}{1-\nabla^2\gamma}\right)_{xt} + \left(\frac{\psi_x+\gamma_{yt}}{1-\nabla^2\gamma}\right)_{yt} + \nabla^2\left(c^2\hat{h} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) - f_0\nabla^2\psi = \nabla\times\mathbf{J}, \tag{2.29}$$

$$\delta\boldsymbol{\alpha}: \quad (1-\nabla^2\gamma)\boldsymbol{\alpha}_t + (-\psi_y+\gamma_{tx})\boldsymbol{\alpha}_x + (\psi_x+\gamma_{ty})\boldsymbol{\alpha}_y = 0, \tag{2.30}$$

where $\hat{h}$ and $\mathbf{u}$ are given by (2.8)-(2.10),

$$Q \equiv \frac{\partial(\alpha,\beta)}{\partial(x,y)} \tag{2.31}$$

is the APV, and

$$\mathbf{J} \equiv (J^x, J^y) \equiv \left(\frac{\partial(\alpha,\beta)}{\partial(y,t)}, \frac{\partial(\alpha,\beta)}{\partial(t,x)} \right). \tag{2.32}$$

From (2.24)-(2.25) and (2.31) it follows that

$$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \tau}\, Q = \mathbf{u} Q. \tag{2.33}$$

Thus $\mathbf{J}$ is the flux of APV. In terms of the physical variables defined by (2.8)-(2.10), (2.28) and (2.29) take the forms

$$\zeta + f_0 - f_0 \hat{h} = Q \tag{2.34}$$

and

$$\nabla \cdot \mathbf{u}_t + \nabla^2 \left(c^2 \hat{h} + \tfrac{1}{2} \mathbf{u} \cdot \mathbf{u} \right) - f_0 (\hat{h} v)_x + f_0 (\hat{h} u)_y = (vQ)_x - (uQ)_y. \tag{2.35}$$

Eliminating Q between (2.34) and (2.35), we obtain the shallow water divergence equation (2.4). To see that (2.28)-(2.30) imply the shallow water vorticity equation (2.3), we make use of the identity

$$\frac{\partial}{\partial t} \frac{\partial(\alpha,\beta)}{\partial(x,y)} + \frac{\partial}{\partial x} \frac{\partial(\alpha,\beta)}{\partial(y,t)} + \frac{\partial}{\partial y} \frac{\partial(\alpha,\beta)}{\partial(t,x)} = 0, \tag{2.36}$$

which holds for any two functions $\alpha(x,y,t)$ and $\beta(x,y,t)$. By (2.31)-(2.33), (2.36) is equivalent to

$$Q_t + \nabla \cdot (Q \mathbf{u}) = 0. \tag{2.37}$$

Substituting (2.34) into (2.37) and using (2.2), we obtain (2.3). This concludes the proof that stationarity of (2.27) is equivalent to shallow water dynamics.

Before proceeding, we briefly return to the Lagrangian (2.18) for point vortices of APV. The equations resulting from the requirement that (2.18) be stationary with respect to variations of $\psi(\mathbf{x},t)$, $\gamma(\mathbf{x},t)$ and $\mathbf{x}_i(t)$ are (2.28) with

$$Q = \sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)), \tag{2.38}$$

(2.29) with

$$\mathbf{J} = \sum_i \Gamma_i \delta(\mathbf{x} - \mathbf{x}_i(t)) \dot{\mathbf{x}}_i, \tag{2.39}$$

and (2.16). The definitions (2.38) and (2.39) imply

$$Q_t + \nabla \cdot \mathbf{J} = 0. \tag{2.40}$$

If the fluid depth is nearly constant, then we may replace the denominators $1 - \nabla^2 \gamma$ in (2.14) by unity, obtaining the quadratic approximation,

$$L_1^0 = \iiint d\mathbf{x} dt \, \left(\tfrac{1}{2} \nabla\psi \cdot \nabla\psi + \tfrac{1}{2} \nabla\gamma_t \cdot \nabla\gamma_t + f_0 \nabla\psi \cdot \nabla\gamma - \tfrac{1}{2} c^2 (\nabla^2 \gamma)^2 \right), \tag{2.41}$$

to L_1. If $f_0 = 0$ then the Lagrangian $L_1^0 + L_2$, where L_2 is given by (2.17), is equivalent to the Lagrangian given in S14, which was shown in that paper to be closely analogous to the Lagrangian for classical electrodynamics. In this approximation the waves are linear non-dispersive waves that interact with the point vortices but not with each other. The point vortices are analogous to electrons. Q is analogous to electric charge density and $\mathbf{J}$ is analogous to electric current density. Equation (2.16) is analogous to the Lorentz force law. Further details of the electrodynamic analogy are given in Appendix A.

In the present paper we have greater use for the continuous-vorticity Lagrangian (2.27) than for its discrete-vorticity counterpart (2.18). The latter remains a valuable thinking tool (as illustrated in Section 5) and may be useful for numerical studies. However, it is worth pointing out that point-vortex dynamics corresponding to (2.18) cannot be obtained as a special case of (2.27). This is obvious from the fact that, in applying point vortex dynamics, we must omit the influence of each vortex on itself. Furthermore, any attempt to derive the continuous-vorticity dynamics by averaging over point vortices would introduce terms analogous to the molecular viscosity terms that arise when one attempts to derive ideal fluid dynamics by averaging over molecular motions.

The continuous-vorticity Lagrangian (2.27) seems a good starting point for theories describing the interactions between inertia-gravity waves and quasi-geostrophic flow. Three reasons for this are readily apparent. First, the Eulerian average of (2.9) and (2.10) corresponds to the Lagrangian mean velocity, often denoted $\bar{\mathbf{u}}^L$. Thus the average values of ψ and γ_t represent the stream function and velocity potential for the Lagrangian mean flow. The automatic satisfaction of (2.2) by the representation (2.8)-(2.10) corresponds to the fact that, in the Lagrangian mean, no Reynolds fluxes occur in the average of (2.2).

Second, the Lagrangian L_1, by itself, describes the fluid motion that results when waves propagate into a region that is initially at rest, and in which the APV therefore vanishes. That is, (2.14) is the Lagrangian for nonlinear inertia-gravity waves uncontaminated by APV. If L_1 is approximated as (2.41), then the equations resulting from $\delta L_1^0 = 0$ combine to yield the Klein-Gordon equation,

$$\nabla^2\gamma_{tt} + f_0^2\nabla^2\gamma - c^2\nabla^2\nabla^2\gamma = 0, \tag{2.42}$$

for $\nabla^2\gamma$, which describes linear, shallow-water inertia-gravity waves. The fully nonlinear form (2.14) will be used in Section 4 to determine the mean flow that arises when inertia-gravity waves propagate into an initially quiescent region.

Third, there is an intimate connection between (2.27) and the Lagrangian for quasi-geostrophic motion. To see this most directly, suppose that all the γ-terms except $f_0\nabla\psi\cdot\nabla\gamma$ and $c^2(\nabla^2\gamma)^2$ are simply dropped from (2.27). (This can be justified by scaling the variables in the manner appropriate for geostrophic flow.) The resulting Lagrangian is

$$\iiint dtd\mathbf{x}\left(\tfrac{1}{2}\nabla\psi\cdot\nabla\psi + f_0\nabla\psi\cdot\nabla\gamma - \tfrac{1}{2}c^2(\nabla^2\gamma)^2 - \alpha\beta_t + \psi J(\alpha,\beta)\right), \tag{2.43}$$

where $J(\alpha,\beta)\equiv\alpha_x\beta_y - \beta_x\alpha_y$. If we use

$$\delta\gamma: \qquad \nabla^2\left(f_0\psi + c^2\nabla^2\gamma\right) = 0, \tag{2.44}$$

hence

$$\psi = -c^2/f_0\nabla^2\gamma = g/f_0\ (h - h_0), \tag{2.45}$$

to eliminate γ from (2.43), we obtain the Lagrangian

$$L_{QG}[\psi,\alpha,\beta] = \iiint dtd\mathbf{x}\left(\tfrac{1}{2}\nabla\psi\cdot\nabla\psi + \frac{1}{2}\frac{f_0^2}{c^2}\psi^2 - \alpha\beta_t + \psi J(\alpha,\beta)\right) \tag{2.46}$$

for quasi-geostrophic dynamics. The condition (2.45) is a statement of geostrophic bal-

ance. Stationarity of (2.46) implies

$$\delta\psi: \quad \nabla^2\psi - \frac{f_0^2}{c^2}\psi = J(\alpha,\beta), \tag{2.47}$$

$$\delta\beta: \quad \alpha_t + J(\psi,\alpha) = 0, \tag{2.48}$$

$$\delta\alpha: \quad \beta_t + J(\psi,\beta) = 0. \tag{2.49}$$

Using the Jacobi identity,

$$J(J(\alpha,\beta),\psi) + J(J(\beta,\psi),\alpha) + J(J(\psi,\alpha),\beta) = 0, \tag{2.50}$$

we find that (2.47)-(2.49) imply the quasi-geostrophic potential vorticity equation,

$$(\nabla^2\psi - \psi/{r_d}^2)_t + J(\psi, \nabla^2\psi - \psi/{r_d}^2) = 0, \tag{2.51}$$

where $r_d = c/f_0$ is the deformation radius.

A more systematic treatment of (2.27) yields theories that describe the interactions between motions that obey (2.51) at leading order and those that obey (2.42) at leading order. The use of the Lagrangian formulation guarantees that all conservation laws will be automatically maintained. However, instead of applying this strategy to shallow water dynamics, we generalize our results to hydrostatic Boussinesq dynamics, which are of much greater interest to meteorologists and oceanographers. That such a generalization must exist is obvious from the close correspondence between the shallow water equations and the hydrostatic Boussinesq equations in buoyancy coordinates.

3. Hydrostatic Boussinesq dynamics

In buoyancy coordinates, the hydrostatic Boussinesq equations take the form

$$\mathbf{u}_t + \nabla\left(B + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) = (\zeta + f_0)(v,-u), \tag{3.1}$$

$$z = -B_\theta, \tag{3.2}$$

$$z_{\theta t} + (uz_\theta)_x + (vz_\theta)_y = 0. \tag{3.3}$$

Here, $\mathbf{u}(x,y,\theta,t) = \mathbf{u}(\mathbf{x},\theta,t) = (u,v)$ is the horizontal velocity, θ is the buoyancy, $z(\mathbf{x},\theta,t)$ is the height of the iso-buoyancy surface corresponding to θ at the horizontal location $\mathbf{x}$, and $B = p - z\theta$ with p equal to the pressure divided by the constant mass density. As before $\zeta = \nabla\times\mathbf{u} \equiv v_x - u_y$ but now all horizontal derivatives are taken with θ fixed. For simplicity we consider a horizontally unbounded fluid with flat rigid boundaries at z_1 and $z_2 = z_1 + h_0$. At these boundaries, the buoyancy is (and remains) uniform at the values θ_1 and θ_2. Thus the boundary conditions on (3.1)-(3.3) are

$$B_\theta(\mathbf{x},\theta_1,t) = -z_1, \qquad B_\theta(\mathbf{x},\theta_2,t) = -z_2. \tag{3.4}$$

The associated vorticity and divergence equations are

$$\zeta_t = -\nabla\cdot((\zeta+f_0)\mathbf{u}) \tag{3.5}$$

and

$$\frac{\partial}{\partial t}\nabla\cdot\mathbf{u} + \nabla^2(B + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}) = \nabla\times((\zeta+f_0)\mathbf{u}), \tag{3.6}$$

where now $\nabla = (\partial_x,\partial_y)$ is the horizontal gradient operator with θ held fixed. The potential vorticity equation is

$$\frac{D}{Dt}\left(\frac{\zeta+f_0}{z_\theta}\right) = 0, \tag{3.7}$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}. \tag{3.8}$$

For derivations of these equations, see, for example, Salmon (1998, pp. 105-107).

We seek a variational principle analogous to that found for shallow water dynamics. First, as in Section 2, we suppose that the fluid was, at some time, in the state of rest. In this state the potential vorticity depends only on θ. But since the potential vorticity and buoyancy are conserved on fluid particles, it must be true that

$$\frac{\zeta + f_0}{z_\theta} = F(\theta) \tag{3.9}$$

at all times, for some function F. Evaluating (3.9) in the state of rest, we find that $F(\theta) = f_0 {N_0}^2(\theta)$ where $N_0(\theta)$ is the Vaisala frequency associated with the state of rest. Thus

$$\zeta + f_0 = f_0 {N_0}^2(\theta) z_\theta. \tag{3.10}$$

We introduce the potential representation

$$ {N_0}^2(\theta)\ z_\theta = 1 - \nabla^2\gamma, \tag{3.11}$$

$$ {N_0}^2(\theta)\ z_\theta u = -\psi_y + \gamma_{xt}, \tag{3.12}$$

$$ {N_0}^2(\theta)\ z_\theta v = \psi_x + \gamma_{yt}. \tag{3.13}$$

Just as (2.8)-(2.10) automatically satisfies (2.2), the representation (3.11)-(3.13) satisfies (3.3). The prescribed function ${N_0}^2(\theta)$ is analogous to h_0^{-1} in Section 2. In terms of ψ and γ, the ansatz (3.10) takes the same form,

$$\frac{\partial}{\partial x}\left(\frac{\psi_x + \gamma_{yt}}{1 - \nabla^2\gamma}\right) + \frac{\partial}{\partial y}\left(\frac{\psi_y - \gamma_{xt}}{1 - \nabla^2\gamma}\right) = -f_0\nabla^2\gamma, \tag{3.14}$$

as in shallow water dynamics. The θ-derivative of (3.2) is

$$B_{\theta\theta} = -N_0^{-2}(\theta)(1 - \nabla^2\gamma). \tag{3.15}$$

Let

$$B = B_0(\theta) + \hat{B}(x, y, \theta, t), \tag{3.16}$$

where $B_0(\theta)$ is associated with the state of rest. Then the hydrostatic equation (3.15) becomes

$$\hat{B}_{\theta\theta} = N_0^{-2}(\theta)\nabla^2\gamma, \tag{3.17}$$

and the divergence equation (3.6) becomes

$$\left(\frac{-\psi_y + \gamma_{xt}}{1 - \nabla^2\gamma}\right)_{xt} + \left(\frac{\psi_x + \gamma_{yt}}{1 - \nabla^2\gamma}\right)_{yt} + \nabla^2\left(\hat{B} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) = f_0\nabla^2\psi. \tag{3.18}$$

Equations (3.14), (3.17), and (3.18) result from the variational principle $\delta L_1[\psi, \gamma, \hat{B}] = 0$, where

$$L_1 = \iiiint d\mathbf{x}d\theta dt \left(\tfrac{1}{2}\hat{B}_\theta^2 + \frac{1}{N_0^2}\left[\tfrac{1}{2}\frac{(-\psi_y + \gamma_{xt})^2}{(1 - \nabla^2\gamma)} + \tfrac{1}{2}\frac{(\psi_x + \gamma_{yt})^2}{(1 - \nabla^2\gamma)} + f_0\nabla\psi\cdot\nabla\gamma + \hat{B}\nabla^2\gamma\right]\right). \tag{3.19}$$

The Lagrangian (3.19) is analogous to (2.14).

Before proceeding we slightly adjust the vertical coordinate. We define a new buoyancy coordinate $Z(\theta)$ by

$$dZ/d\theta = N_0^{-2}(\theta), \qquad Z(\theta_1) = z_1. \tag{3.20}$$

Thus $Z(\theta)$ is the height of buoyancy surface θ in the state of rest. With this definition (3.19) takes the form

$$L_1 = \iiiint d\mathbf{x}dZdt \left(\tfrac{1}{2} N_0^{-2} \hat{B}_Z^2 + \tfrac{1}{2} \frac{(-\psi_y + \gamma_{xt})^2}{(1-\nabla^2\gamma)} + \tfrac{1}{2} \frac{(\psi_x + \gamma_{yt})^2}{(1-\nabla^2\gamma)} + f_0 \nabla\psi\cdot\nabla\gamma + \hat{B}\nabla^2\gamma \right). \tag{3.21}$$

The integration limits on Z are z_1 and z_2, and we henceforth regard $N_0 = N_0(Z)$. From now on we work in the (x, y, Z, t) system. Equation (3.10) takes the form

$$\zeta + f_0 = f_0 \; dz/dZ, \tag{3.22}$$

and (3.17) takes the form

$$(N_0^{-2}\hat{B}_Z)_Z = \nabla^2\gamma. \tag{3.23}$$

Following the same path as in Section 2 (but skipping the step corresponding to point vortices), we generalize the ansatz (3.22) to

$$\zeta + f_0 = f_0 \; \frac{dz}{dZ} + \frac{\partial(\alpha,\beta)}{\partial(x,y)}, \tag{3.24}$$

where (α, β) are labels that measure APV and move along buoyancy surfaces at the horizontal velocity $\mathbf{u}$. Thus

$$(\alpha,\beta)_t + \mathbf{u}\cdot\nabla(\alpha,\beta) = 0. \tag{3.25}$$

We define

$$L_2[\psi,\gamma,\alpha,\beta] = \iiiint d\mathbf{x}dZdt \; \left(-\alpha\beta_t + \psi\frac{\partial(\alpha,\beta)}{\partial(x,y)} - \gamma_y \frac{\partial(\alpha,\beta)}{\partial(y,t)} + \gamma_x \frac{\partial(\alpha,\beta)}{\partial(t,x)} \right). \tag{3.26}$$

The Lagrangian (3.21) is analogous to (2.14), and (3.26) is analogous to the last line of (2.27). As we now demonstrate, the variational principle $\delta(L_1 + L_2) = 0$ is equivalent to hydrostatic Boussinesq dynamics.

The variational principle $\delta(L_1 + L_2) = 0$ implies

$$\delta\psi: \quad \frac{\partial}{\partial x}\left(\frac{\psi_x + \gamma_{yt}}{1-\nabla^2\gamma}\right) + \frac{\partial}{\partial y}\left(\frac{\psi_y - \gamma_{xt}}{1-\nabla^2\gamma}\right) + f_0\nabla^2\gamma = \frac{\partial(\alpha,\beta)}{\partial(x,y)}, \tag{3.27}$$

$$\delta\gamma: \quad \left(\frac{-\psi_y + \gamma_{xt}}{1-\nabla^2\gamma}\right)_{xt} + \left(\frac{\psi_x + \gamma_{yt}}{1-\nabla^2\gamma}\right)_{yt} + \nabla^2\left(\hat{B} + \tfrac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) - f_0\nabla^2\psi$$
$$= \frac{\partial}{\partial x}\frac{\partial(\alpha,\beta)}{\partial(t,x)} - \frac{\partial}{\partial y}\frac{\partial(\alpha,\beta)}{\partial(y,t)}, \tag{3.28}$$

$$\delta\hat{B}: \quad (N_0^{-2}\hat{B}_Z)_Z = \nabla^2\gamma, \tag{3.29}$$

$$\delta\beta: \quad \alpha_t + u\alpha_x + v\alpha_y = 0, \tag{3.30}$$

$$\delta\alpha: \quad \beta_t + u\beta_x + v\beta_y = 0, \tag{3.31}$$

and the boundary conditions $\hat{B}_Z = 0$ at $Z = z_1$ and $Z = z_2$, where, by (3.11)-(3.13),

$$\mathbf{u} = \frac{(-\psi_y + \gamma_{tx},\; \psi_x + \gamma_{ty})}{1-\nabla^2\gamma}. \tag{3.32}$$

Solving (3.30) and (3.31) for u and v we obtain

$$u = \frac{1}{Q}\frac{\partial(\alpha,\beta)}{\partial(y,t)}, \qquad v = \frac{1}{Q}\frac{\partial(\alpha,\beta)}{\partial(t,x)}, \tag{3.33}$$

where

$$Q = \frac{\partial(\alpha, \beta)}{\partial(x, y)} \tag{3.34}$$

is the APV. Then (3.27) may be rewritten as

$$\zeta + f_0 \nabla^2 \gamma = Q, \tag{3.35}$$

and (3.28) may be rewritten as

$$\frac{\partial}{\partial t} \nabla \cdot \mathbf{u} + \nabla^2 (\hat{B} + \tfrac{1}{2} \mathbf{u} \cdot \mathbf{u}) - f_0 \nabla^2 \psi = \nabla \times (Q \mathbf{u}). \tag{3.36}$$

By (3.32) and (3.35), the rhs of (3.36) is

$$\nabla \times \big((\zeta + f_0) \mathbf{u} - f_0 (1 - \nabla^2 \gamma) \mathbf{u} \big) = \nabla \times ((\zeta + f_0) \mathbf{u}) - f_0 \nabla^2 \psi. \tag{3.37}$$

Thus (3.36) is equivalent to the divergence equation (3.6). By (3.33) and (3.34), the mathematical identity (2.36) takes the same form,

$$Q_t + \nabla \cdot (Q \mathbf{u}) = 0, \tag{3.38}$$

as in Section 2. Substituting (3.32) and (3.35) into (3.38), we obtain the vorticity equation (3.5). This concludes the proof that $\delta(L_1 + L_2) = 0$ is equivalent to hydrostatic Boussinesq dynamics.

In the rest of this paper we replace Z by z to simplify notation, but it is important to remember that z is a disguised buoyancy coordinate. It is in fact the same coordinate introduced by Young (2012).

To further simplify matters, we adopt the standard field-theory convention that all variations vanish at the extremities of the fluid, both in time and in space. That is, we do not attempt to incorporate boundary conditions (which are in some sense arbitrary) into the variational principle. Our focus is on the equations themselves.

4. Bretherton flow

As a first example of the use of the variational principle discovered in the previous section, we consider the flow that develops when waves enter a region that is initially at rest. Bühler & McIntyre (2005) call this 'Bretherton flow' after the pioneering work of Bretherton (1969). If the fluid is initially at rest, then the APV vanishes, and the dynamics is completely described by $\delta L_1[\psi, \gamma, \hat{B}] = 0$ where L_1 is given by (3.21). If the waves are sufficiently weak we may write

$$L_1 = L_1^0 + L_{1C}, \tag{4.1}$$

where

$$L_1^0 = \iiiint d\mathbf{x} dz dt \left(\tfrac{1}{2} N_0^{-2} \hat{B}_z^2 + \tfrac{1}{2} \nabla \psi \cdot \nabla \psi + \tfrac{1}{2} \nabla \gamma_t \cdot \nabla \gamma_t + f_0 \nabla \psi \cdot \nabla \gamma + \hat{B} \nabla^2 \gamma \right) \tag{4.2}$$

contains the quadratic terms in L_1, and

$$L_{1C} = \iiiint d\mathbf{x} dz dt \left(\tfrac{1}{2} \nabla \psi \cdot \nabla \psi + J(\psi, \gamma_t) + \tfrac{1}{2} \nabla \gamma_t \cdot \nabla \gamma_t \right) \left(\nabla^2 \gamma + (\nabla^2 \gamma)^2 + \cdots \right) \tag{4.3}$$

contains the higher order corrections. By (3.11), small $\nabla^2 \gamma$ corresponds to a nearly undisturbed stratification.

If the wave amplitude is small, then, to leading order, the waves are governed by $\delta L_1^0 = 0$, which implies

$$\delta\psi: \quad \nabla^2\psi + f_0\nabla^2\gamma = 0, \tag{4.4}$$

$$\delta\hat{B}: \quad (N_0^{-2}\hat{B}_z)_z = \nabla^2\gamma, \tag{4.5}$$

$$\delta\gamma: \quad \nabla^2\gamma_{tt} - f_0\nabla^2\psi + \nabla^2\hat{B} = 0. \tag{4.6}$$

Using (4.4) and (4.5) to eliminate ψ and γ, we obtain a single equation for $\hat{B}$,

$$(\partial_{tt} + f_0^2)(N_0^{-2}\hat{B}_z)_z + \nabla^2\hat{B} = 0. \tag{4.7}$$

If $N_0(z)$ varies slowly, then (4.7) supports inertia-gravity waves of the form $\hat{B} \propto \cos(kx + ly + mz - \omega t)$ with dispersion relation

$$\omega^2 = f_0^2 + K^2N_0^2/m^2 \equiv \omega_r(K, m, z)^2, \tag{4.8}$$

where $K^2 = k^2 + l^2$. The dispersion relation (4.8) is correct for hydrostatic inertia-gravity waves, which have frequencies much smaller than N_0.

If the wave amplitude a is small, then the waves, which are approximately governed by (4.4)-(4.6), induce an $\mathrm{O}(a^2)$ mean flow that may be computed from L_1. Let

$$\psi = \bar{\psi} + \psi', \quad \gamma = \bar{\gamma} + \gamma', \quad \hat{B} = \bar{B} + B', \tag{4.9}$$

where bar denotes the mean flow, and prime denotes the wave. We assume that the wave is slowly varying in the sense that its amplitude, wavenumbers and frequency all vary on scales that are large compared to its wavelength and period. We follow Whitham's (1965, 1974) procedure. The averaged Lagrangian is

$$\langle L_1\rangle = L_1^0[\bar{B}, \bar{\psi}, \bar{\gamma}] + \langle L_1^0[B', \psi', \gamma']\rangle + \langle L_{1C}[\bar{\psi} + \psi', \bar{\gamma} + \gamma']\rangle \tag{4.10}$$

where L_1^0 and L_{1C} are defined by (4.2) and (4.3). The angle brackets denote the average over wave phase. Thus $\langle\bar{\psi}\rangle = \bar{\psi}$ and $\langle\psi'\rangle = 0$. We approximate the last term in (4.10) by keeping only the cubic contributions to (4.3) that are quadratic in the primes. Then

$$\langle L_{1C}\rangle = \iiiint d\mathbf{x}dzdt\left(M\bar{\psi} + S\bar{\gamma}\right), \tag{4.11}$$

where

$$M = \langle J(\gamma_t', \nabla^2\gamma')\rangle - \nabla\cdot\langle\nabla^2\gamma'\nabla\psi'\rangle \tag{4.12}$$

and

$$\begin{aligned} S =& \nabla^2\left(\tfrac{1}{2}\langle\nabla\psi'\cdot\nabla\psi'\rangle + \langle J(\psi', \gamma_t')\rangle + \tfrac{1}{2}\langle\nabla\gamma_t'\cdot\nabla\gamma_t'\rangle\right) \\ &+ \langle J(\psi', \nabla^2\gamma')\rangle_t + \nabla\cdot\langle\nabla^2\gamma'\nabla\gamma_t'\rangle_t. \end{aligned} \tag{4.13}$$

The enormous advantage of Whitham's method is that terms in the Lagrangian may be simplified before the variations are taken to determine the equations. It turns out that all the wave-forcing terms on the rhs of (4.12) and (4.13) are negligible except the first term in (4.12). Let

$$\gamma' = A(x, y, z, t)\cos\phi(x, y, z, t), \tag{4.14}$$

where A is the slowly varying amplitude and ϕ is the rapidly varying phase. The frequency and wave numbers,

$$\omega = -\phi_t, \quad k = \phi_x, \quad l = \phi_y, \quad m = \phi_z, \tag{4.15}$$

all vary slowly. First we consider the two terms on the rhs of (4.12). At leading order,

$$\langle J(\gamma'_t, \nabla^2\gamma')\rangle = \langle \gamma'_{tx}\nabla^2\gamma'\rangle_y - \langle \gamma'_{ty}\nabla^2\gamma'\rangle_x = \nabla \times \left(\tfrac{1}{2}\omega K^2 A^2\, \mathbf{k}\right), \tag{4.16}$$

where $K^2 = k^2 + l^2$. For the second term we have

$$-\nabla \cdot \langle \nabla^2\gamma' \nabla\psi'\rangle = f_0 \nabla \cdot \langle \nabla^2\gamma' \nabla\gamma'\rangle, \tag{4.17}$$

where we have used the fact $\psi' = -f_0\gamma'$ at leading order; see (4.4). The phase average in (4.17) is over two factors, $\nabla^2\gamma'$ and $\nabla\gamma'$, that are 90° out of phase. Thus (4.17) is smaller than (4.16) by a factor ϵ, where $\epsilon << 1$ is the 'scale separation' parameter—the ratio of the wavelength to the length scale of slow variation. Similar remarks apply to all of the terms in (4.13): In contrast to (4.16), all involve at least two derivatives of the slowly varying wave amplitude, frequency and wavenumbers. Thus the approximation (4.11) becomes

$$\langle L_{1C}\rangle = \iiiint d\mathbf{x}dzdt\ \bar{\psi}\ \nabla \times \left(\tfrac{1}{2}\omega K^2 A^2\, \mathbf{k}\right), \tag{4.18}$$

and the complete Lagrangian (4.10) becomes

$$\begin{aligned}\langle L_1\rangle[\bar{B}, \bar{\psi}, \bar{\gamma}, A, \phi] = \\ \iiiint d\mathbf{x}dzdt\ \big[&\tfrac{1}{2}N_0^{-2}\bar{B}_z^2 + \tfrac{1}{2}\nabla\bar{\psi}\cdot\nabla\bar{\psi} + \tfrac{1}{2}\nabla\bar{\gamma}_t\cdot\nabla\bar{\gamma}_t + f_0\nabla\bar{\psi}\cdot\nabla\bar{\gamma} + \bar{B}\nabla^2\bar{\gamma} \\ &+ \tfrac{1}{4}K^2A^2(\omega^2 - \omega_r^2) + \bar{\psi}\ \nabla \times \left(\tfrac{1}{2}\omega K^2 A^2\, \mathbf{k}\right)\big],\end{aligned} \tag{4.19}$$

where the relative frequency ω_r is defined by (4.8). The equations resulting from $\delta\langle L_1\rangle = 0$ are:

$$\delta\bar{B}: \quad (N_0^{-2}\bar{B}_z)_z - \nabla^2\bar{\gamma} = 0, \tag{4.20}$$

$$\delta\bar{\psi}: \quad \nabla^2\bar{\psi} + f_0\nabla^2\bar{\gamma} = \nabla \times (E\mathbf{k}/\omega), \tag{4.21}$$

$$\delta\bar{\gamma}: \quad \nabla^2\bar{\gamma}_{tt} - f_0\nabla^2\bar{\psi} + \nabla^2\bar{B} = 0, \tag{4.22}$$

$$\delta A^2: \quad \omega^2 = f_0^2 + K^2N_0^2/m^2 + 2\omega\bar{\mathbf{U}}\cdot\mathbf{k}, \tag{4.23}$$

$$\delta\phi: \quad \left(\frac{E}{\omega}\left(1 - \frac{\bar{\mathbf{U}}\cdot\mathbf{k}}{\omega}\right)\right)_t + \nabla_3 \cdot \left(\frac{E}{\omega}\left(\frac{\omega_r}{\omega}\mathbf{c}_{gr} + \bar{\mathbf{U}}\right)\right) = 0, \tag{4.24}$$

where $E \equiv \frac{1}{2}K^2\omega^2A^2$, $\bar{\mathbf{U}} \equiv (-\bar{\psi}_y, \bar{\psi}_x)$, $\nabla_3 \equiv (\partial_x, \partial_y, \partial_z)$, and

$$\mathbf{c}_{gr} = \left(\frac{\partial\omega_r}{\partial k}, \frac{\partial\omega_r}{\partial l}, \frac{\partial\omega_r}{\partial m}\right) = \frac{1}{\sqrt{1 + f_0^2m^2/K^2N_0^2}}\frac{N_0}{m}\left(\frac{k}{K}, \frac{l}{K}, -\frac{K}{m}\right) \tag{4.25}$$

is the relative group velocity. The only forcing term in the equations (4.20)-(4.22) determining the mean flow is the curl of the pseudomomentum

$$\mathbf{p} \equiv E\mathbf{k}/\omega. \tag{4.26}$$

This $O(a^2)$ term induces an $O(a^2)$ mean flow that contributes the small final term to the dispersion relation (4.23). Eqn. (4.23) differs from the expected Doppler-shifted result $\omega = \omega_r + \mathbf{U}\cdot\mathbf{k}$ by an even smaller $O(a^4)$ term, namely $(\mathbf{U}\cdot\mathbf{k})^2$, because the approximation (4.18) excludes terms that are quadratic in the mean flow. Similar remarks apply to (4.24). If we keep only the largest terms in each equation, then (4.20)-(4.22) are unchanged, (4.23) reduces to $\omega = \omega_r$, and (4.24) takes the familiar form,

$$(E/\omega)_t + \nabla_3 \cdot (\mathbf{c}_g E/\omega) = 0, \tag{4.27}$$

of action conservation. In this limit we may consider the slowly varying wave train to be a prescribed solution of the linear equations (4.4)-(4.6). The only challenge is to determine the induced mean flow from (4.20)-(4.22). However, it is worth pointing out that (4.20)-(4.24) are consistent with all conservation laws (including total energy conservation) because they have been obtained from a variational principle, whereas *a posteriori* approximations like (4.27) generally destroy exact conservation.

The general solution of the mean flow equations (4.20)-(4.22) includes free inertia-gravity waves, but to include them would be redundant: We are solely interested in the forced response to $\nabla \times \mathbf{p}$. If $f_0 = 0$, then $\bar{\gamma}$ and $\bar{B}$ vanish, and

$$\nabla^2 \bar{\psi} = \nabla \times \mathbf{p} \tag{4.28}$$

determines the Lagrangian mean vorticity. If $f_0 \neq 0$ then (4.20)-(4.22) may be combined to give

$$(\partial_{tt} + f_0^2)(N_0^{-2}\bar{B}_z)_z + \nabla^2 \bar{B} = f_0 \nabla \times \mathbf{p}. \tag{4.29}$$

Suppose that the pseudomomentum $\mathbf{p}$ corresponds to a wavepacket. The wavepacket propagates at the group velocity corresponding to its 'carrier wavenumber' $\mathbf{k}$. The forced solution of (4.29) propagates at this same group velocity. Hence we may solve (4.29) in the reference frame moving with the packet by replacing $\partial_t \to -\mathbf{c}_{gr} \cdot \nabla_3$ in (4.29). If the group velocity is sufficiently small, then the ∂_{tt}-term in (4.29) is negligible, and (4.29) reduces to

$$\nabla^2 \bar{B} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N_0^2}\frac{\partial \bar{B}}{\partial z}\right) = f_0 \nabla \times \mathbf{p}. \tag{4.30}$$

In this same limit, (4.22) implies $\bar{B} = f_0 \bar{\psi}$ so that (4.30) is equivalent to

$$\nabla^2 \bar{\psi} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N_0^2}\frac{\partial \bar{\psi}}{\partial z}\right) = \nabla \times \mathbf{p}. \tag{4.31}$$

The lhs of (4.31) is the quasigeostrophic potential vorticity. Thus, if the inertia-gravity wavepacket propagates slowly enough, the mean flow response is quasigeostrophic.

The WKB ansatz (4.14), which assumes a scale separation in all four of the dependent variables (x, y, z, t), is overly restrictive, because inertia-gravity waves with frequencies near f_0 correspond to large horizontal length scales. If we assume that only the time scale separates the mean flow from the waves, then only the last two terms in (4.13) may be neglected; all the other terms in (4.12) and (4.13) must be retained. However, before proceeding with less restrictive assumptions we generalize the dynamics to the case of nonvanishing APV.

5. Nonvanishing available potential vorticity

If the APV does not vanish then the full Lagrangian includes L_2, which couples the inertia-gravity waves to the APV. The introduction of APV is analogous to the introduction of electric charge. In this section we develop approximations to L_2 analogous to those developed for L_1 in the previous section. We start by considering the form (2.17) of L_2 corresponding to shallow-water dynamics with point vortices of APV. Let

$$\mathbf{x}_i(t) = \bar{\mathbf{x}}_i(t) + \boldsymbol{\xi}_i(t), \tag{5.1}$$

where $\bar{\mathbf{x}}_i(t) = \langle \mathbf{x}_i(t) \rangle$ is the average location of the i-th point vortex and $\boldsymbol{\xi}_i(t) = (\xi_i(t), \eta_i(t))$ is the departure therefrom. Thus $\xi_i(t)$ and $\eta_i(t)$ are rapidly fluctuating vari-

ables like $\gamma'(\mathbf{x},t)$ and $\psi'(\mathbf{x},t)$. Substituting (5.1) into (2.17) and averaging we obtain

$$\langle L_2\rangle = \langle \int dt \sum_i \Gamma_i \big(-\bar{x}_i\dot{\bar{y}}_i - \xi_i\dot{\eta}_i + \psi(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t) \\ - \gamma_y(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)(\dot{\bar{x}}_i+\dot{\xi}_i) + \gamma_x(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)(\dot{\bar{y}}_i+\dot{\eta}_i)\big)\,\rangle. \tag{5.2}$$

We proceed by Taylor expansion of the functions ψ, γ_y, and γ_x about the argument $\bar{\mathbf{x}}_i$. Consider the single term

$$\langle \int dt \sum_i \Gamma_i \psi(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)\rangle = \langle \int dt \sum_i \Gamma_i \left(\bar{\psi}(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t) + \psi'(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)\right)\rangle. \tag{5.3}$$

The function $\bar{\psi}$ depends slowly on its argument no matter what the character of the argument itself. Similarly, the function ψ' depends rapidly on its argument. Proceeding with the Taylor expansion, and keeping only the quadratic terms in the fast variables, we obtain

$$\langle \int dt \sum_i \Gamma_i \psi(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)\rangle \approx \langle \int dt \sum_i \Gamma_i \left(\bar{\psi}(\bar{\mathbf{x}}_i,t) + \tfrac{1}{2}(\boldsymbol{\xi}_i\cdot\nabla)^2\bar{\psi}(\bar{\mathbf{x}}_i,t) + \boldsymbol{\xi}_i\cdot\nabla\psi'(\bar{\mathbf{x}}_i,t)\right)\rangle. \tag{5.4}$$

In the preceding section we approximated L_1 by keeping only the cubic terms of the form bar-prime-prime. Since the constant Γ_i counts a 'bar' variable, the middle term on the rhs of (5.4) is of the order bar-bar-prime-prime. Thus if we treat L_2 in the same manner as we treated L_1 in the previous section, we can consistently neglect this middle term. (We come back to this point in Section 8.) Our approximation becomes

$$\langle \int dt \sum_i \Gamma_i \psi(\bar{\mathbf{x}}_i+\boldsymbol{\xi}_i,t)\rangle \approx \int dt \sum_i \Gamma_i \left(\bar{\psi}(\bar{\mathbf{x}}_i,t) + \langle\boldsymbol{\xi}_i\cdot\nabla\psi'(\bar{\mathbf{x}}_i,t)\rangle\right). \tag{5.5}$$

The last two terms in (5.2) may be approximated in a similar manner, and the result is

$$\langle L_2\rangle = \int dt \sum_i \Gamma_i \big(-\bar{x}_i\dot{\bar{y}}_i + \bar{\psi}(\bar{\mathbf{x}}_i,t) - \bar{\gamma}_y(\bar{\mathbf{x}}_i,t)\dot{\bar{x}}_i + \bar{\gamma}_x(\bar{\mathbf{x}}_i,t)\dot{\bar{y}}_i \\ - \langle\xi_i\dot{\eta}_i\rangle + \langle\boldsymbol{\xi}_i\cdot\nabla\psi'(\bar{\mathbf{x}}_i,t)\rangle - \langle\gamma'_y(\bar{\mathbf{x}}_i,t)\dot{\xi}_i\rangle + \langle\gamma'_x(\bar{\mathbf{x}}_i,t)\dot{\eta}_i\rangle\big). \tag{5.6}$$

The $\bar{\gamma}$-terms in (5.6) involve only the mean flow and represent higher-order corrections to quasigeostrophy. Neglecting these terms, and defining $\boldsymbol{\xi}(\mathbf{x},t)$ to be the fluctuating displacement of the fluid particle currently located at $\mathbf{x}$ (whether or not there is any APV there), we rewrite (5.6) as

$$\langle L_2\rangle = \iiint dtd\mathbf{x} \sum_i \Gamma_i\delta(\mathbf{x}-\bar{\mathbf{x}}_i(t))\big(-x\dot{\bar{y}}_i + \bar{\psi} - \langle\xi\eta_t\rangle + \langle\boldsymbol{\xi}\cdot\nabla\psi'\rangle - \langle\gamma'_y\xi_t\rangle + \langle\gamma'_x\eta_t\rangle\big). \tag{5.7}$$

Because of the delta function, all of the terms in the integrand of (5.7) can be considered functions of $(\mathbf{x},t)$. We define

$$\tilde{Q}(x,y,t) \equiv \sum_i \Gamma_i\delta(\mathbf{x}-\bar{\mathbf{x}}_i(t)). \tag{5.8}$$

In contrast to the Q defined by (2.38), $\tilde{Q}$ depends slowly on time; it evolves with the average locations of the point vortices. However $\tilde{Q}$ is not the average of Q; in that way it resembles the quantity $\tilde{\rho}$ introduced by Bühler & McIntyre (1998). Just as their $\tilde{\rho}$ represents the mass density that would occur if the fluid moved at the average velocity, $\tilde{Q}$ represents the APV that would be present if the point vortices move at their average velocity $\dot{\bar{\mathbf{x}}}_i$.

We pass to the limit of continuous APV by replacing

$$\sum_i \Gamma_i \delta(\mathbf{x} - \bar{\mathbf{x}}_i(t)) \to \frac{\partial(\tilde{\alpha}, \tilde{\beta})}{\partial(x, y)}, \tag{5.9}$$

where $(\tilde{\alpha}, \tilde{\beta})$ are labeling variables for $\tilde{Q}$. Compare (5.9) to (2.21). Then, following steps similar to those between (2.21) and (2.27), we arrive at

$$\langle L_2 \rangle = \iiint dt d\mathbf{x} \left(-\tilde{\alpha}\tilde{\beta}_t + \frac{\partial(\tilde{\alpha}, \tilde{\beta})}{\partial(x, y)} \left(\bar{\psi} - \langle \xi \eta_t \rangle + \langle \boldsymbol{\xi} \cdot \nabla \psi' \rangle - \langle \gamma'_y \xi_t \rangle + \langle \gamma'_x \eta_t \rangle \right) \right) \tag{5.10}$$

for the shallow-water case. The $\langle L_2 \rangle$ for hydrostatic Boussinesq dynamics requires only an additional integration with respect to z and the re-interpretation of all the derivatives in (5.10) as derivatives with z held fixed.

The Lagrangian (5.10) depends on the four fast variables ψ', γ', ξ and η. However, we may use the equations

$$\delta\xi : \qquad \eta_t = \psi'_x + \gamma'_{ty}, \tag{5.11}$$

$$\delta\eta : \qquad \xi_t = -\psi'_y + \gamma'_{tx}, \tag{5.12}$$

to eliminate ψ' and γ' in favor of ξ and η. Note that the variations $\delta\boldsymbol{\xi}$ are performed inside the averaging symbol in (5.10), and that the time-integrations by parts needed to establish (5.11)-(5.12) ignore the slow time variations of the Jacobian in (5.10). Eqns (5.11)-(5.12) equate the time derivatives of $\boldsymbol{\xi}$ to the fluctuating velocity. Substituting (5.11)-(5.12) back into (5.10) (and adding the z-integration required for three-dimensional Boussinesq dynamics) we obtain

$$\langle L_2 \rangle = \iiiint dt d\mathbf{x} dz \left(-\tilde{\alpha}\tilde{\beta}_t + \frac{\partial(\tilde{\alpha}, \tilde{\beta})}{\partial(x, y)} \left(\bar{\psi} + \langle \xi \eta_t \rangle \right) \right) \tag{5.13}$$

for hydrostatic Boussinesq dynamics.

6. The complete Lagrangian

Now we combine the results of Sections 4 and 5 to obtain the complete Lagrangian. To obtain a form of $\langle L_1 \rangle$ that is compatible with (5.13), we must eliminate the variables ψ' and γ' from (4.10)-(4.13). Using (5.11)-(5.12) we obtain

$$\begin{aligned} \langle L_1 \rangle = \iiiint d\mathbf{x} dz dt \, \big(& \tfrac{1}{2} N_0^{-2} (\bar{B}_z)^2 + \tfrac{1}{2} \nabla\bar{\psi} \cdot \nabla\bar{\psi} + \tfrac{1}{2} \nabla\bar{\gamma}_t \cdot \nabla\bar{\gamma}_t + f_0 \nabla\bar{\psi} \cdot \nabla\bar{\gamma} + \bar{B} \nabla^2 \bar{\gamma} \\ & \langle \tfrac{1}{2} N_0^{-2} (B'_z)^2 + \tfrac{1}{2} \boldsymbol{\xi}_t \cdot \boldsymbol{\xi}_t + f_0 \xi \eta_t + B'(\xi_x + \eta_y) \rangle + \bar{\psi} M + \bar{\gamma} \nabla^2 T \big), \end{aligned} \tag{6.1}$$

where

$$M = -\nabla \times \langle (\xi_x + \eta_y) \boldsymbol{\xi}_t \rangle \tag{6.2}$$

and

$$T = \langle \tfrac{1}{2} \boldsymbol{\xi}_t \cdot \boldsymbol{\xi}_t \rangle. \tag{6.3}$$

In writing (6.1), we have dropped the last two terms in (4.13): Since the average is an average over the fast time dependence of the waves, we have $\partial_t \langle \rangle = 0$.

Now we embark on a series of simplifications. First, assuming that the mean flow response will be slow, we drop the $\frac{1}{2} \nabla\bar{\gamma}_t \cdot \nabla\bar{\gamma}_t$ term from (6.1). This neglect is equivalent

to the neglect of $\bar{B}_{tt}$ in (4.29), which led to (4.31). By *a priori* neglect of $\frac{1}{2}\nabla\bar{\gamma}_t\cdot\nabla\bar{\gamma}_t$, we prevent the mean flow from developing its own inertia-gravity waves.

Next, noting that the slow variables $\bar{B}$ and $\bar{\gamma}$, appear only in $\langle L_1\rangle$, we use the equation

$$\delta\bar{\gamma}: \quad -f_0\nabla^2\bar{\psi}+\nabla^2\bar{B}+\nabla^2 T=0 \tag{6.4}$$

to remove $\bar{B}$ and $\bar{\gamma}$ from (6.1). That is, substituting

$$\bar{B}=f_0\bar{\psi}-T \tag{6.5}$$

into the terms

$$\begin{aligned}&\iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}N_0^{-2}(\bar{B}_z)^2+f_0\nabla\bar{\psi}\cdot\nabla\bar{\gamma}+\bar{B}\nabla^2\bar{\gamma}+\bar{\gamma}\nabla^2 T\right)\\ &=\iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}N_0^{-2}(\bar{B}_z)^2+\nabla^2\bar{\gamma}(-f_0\bar{\psi}+\bar{B}+T)\right),\end{aligned} \tag{6.6}$$

we obtain

$$\iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}N_0^{-2}(\bar{B}_z)^2+0\right)\approx\iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}\frac{f_0^2}{N_0^2}(\bar{\psi}_z)^2-\frac{f_0}{N_0^2}\bar{\psi}_z T_z\right) \tag{6.7}$$

after neglecting terms that are quartic in the wave amplitudes. With these simplifications (6.1) takes the form

$$\begin{aligned}\langle L_1\rangle=\iiiint d\mathbf{x}dzdt\ \big(&\tfrac{1}{2}\nabla\bar{\psi}\cdot\nabla\bar{\psi}+\tfrac{1}{2}\frac{f_0^2}{N_0^2}(\bar{\psi}_z)^2-\frac{f_0}{N_0^2}\bar{\psi}_z T_z\\ &+\langle\tfrac{1}{2}N_0^{-2}(B'_z)^2+\tfrac{1}{2}\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t+f_0\xi\eta_t+B'\nabla^2(\xi_x+\eta_y)\rangle+\bar{\psi}M\big).\end{aligned} \tag{6.8}$$

Finally, we note that $\delta(\langle L_1\rangle+\langle L_2\rangle)=0$ implies

$$\delta\bar{\psi}: \qquad \nabla^2\bar{\psi}+\left(\frac{f_0^2}{N_0^2}\bar{\psi}_z\right)_z-M-\left(\frac{f_0}{N_0^2}T_z\right)_z=\frac{\partial(\tilde{\alpha},\tilde{\beta})}{\partial(x,y)}. \tag{6.9}$$

To consistent order we may use (6.9) to eliminate the Jacobian on the rhs of (5.13) from its product with $\langle\xi\eta_t\rangle$. Neglecting terms quartic in the wave amplitudes, we obtain

$$\langle L_2\rangle[\tilde{\alpha},\tilde{\beta},\bar{\psi},\xi,\eta]=\iiiint dtd\mathbf{x}dz\left(-\tilde{\alpha}\tilde{\beta}_t+\bar{\psi}\frac{\partial(\tilde{\alpha},\tilde{\beta})}{\partial(x,y)}+\langle\xi\eta_t\rangle\mathcal{L}\bar{\psi}\right), \tag{6.10}$$

where

$$\mathcal{L}\equiv\nabla^2+\partial_z\frac{f_0^2}{N_0^2}\partial_z \tag{6.11}$$

is the familiar potential vorticity operator.

Combining (6.8) and (6.10) we obtain the averaged Lagrangian for the complete system as

$$\langle L_1+L_2\rangle=L_{QG}[\bar{\psi},\tilde{\alpha},\tilde{\beta}]+L_{IG}[\xi,\eta,B']+L_C[\bar{\psi},\xi,\eta], \tag{6.12}$$

where

$$L_{QG}=\iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}\nabla\bar{\psi}\cdot\nabla\bar{\psi}+\tfrac{1}{2}\frac{f_0^2}{N_0^2}(\bar{\psi}_z)^2-\tilde{\alpha}\tilde{\beta}_t+\bar{\psi}\frac{\partial(\tilde{\alpha},\tilde{\beta})}{\partial(x,y)}\right) \tag{6.13}$$

is the Lagrangian for quasigeostrophic dynamics in the absence of inertia-gravity waves,

$$L_{IG}=\iiiint d\mathbf{x}dzdt\ \langle\tfrac{1}{2}N_0^{-2}(B'_z)^2+\tfrac{1}{2}\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t+f_0\xi\eta_t+B'(\xi_x+\eta_y)\rangle \tag{6.14}$$

is the Lagrangian for inertia-gravity waves in the absence of a quasigeostrophic mean flow, and

$$L_C = \iiiint d\mathbf{x}dzdt\ \bar{\psi}\left(M + (\frac{f_0}{N_0^2}T_z)_z + \mathcal{L}\langle\xi\eta_t\rangle\right) \tag{6.15}$$

is the Lagrangian that couples the inertia-gravity waves to the quasigeostrophic motion. The quadratic wave averages M and T are defined by (6.2) and (6.3). The terms involving M and T occur in (6.1) and would be present even if there were no APV. The last term in (6.15) represents the coupling between the waves and the APV.

The Lagrangian (6.12)-(6.15) is our fundamental result. To derive it, we have assumed that the quasigeostrophic motion, now represented by $(\bar{\psi}, \tilde{\alpha}, \tilde{\beta})$, evolves slowly in comparison to the inertia-gravity waves, now represented by (ξ, η, B'), and we have identified the average with an average over fast time. This assumption of a timescale separation between the mean flow and the waves would seem to be the weakest assumption on which to base any wave-mean theory: The waves cannot have frequencies less than f_0, and the quasigeostrophic flow evolves on a timescale longer than $1/f_0$. Stronger assumptions lead to simplifications of (6.15) that yield simpler equations. In the next section we explore a spectrum of possibilities.

7. Wave-mean equations

We work with the Lagrangian (6.12)-(6.15), which assumes that the waves and the mean flow are separated only by their timescales. We obtain the equations for the mean flow by varying $\bar{\psi}$, $\tilde{\alpha}$, and $\tilde{\beta}$. By the Jacobi identity (2.50) and the definition

$$\tilde{Q} = \frac{\partial(\tilde{\alpha}, \tilde{\beta})}{\partial(x, y)}, \tag{7.1}$$

the equations

$$\delta\tilde{\boldsymbol{\alpha}}: \qquad \tilde{\boldsymbol{\alpha}}_t + J(\bar{\psi}, \tilde{\boldsymbol{\alpha}}) = 0 \tag{7.2}$$

imply

$$\tilde{Q}_t + J(\bar{\psi}, \tilde{Q}) = 0. \tag{7.3}$$

The $\delta\bar{\psi}$-variation yields the defining equation for $\tilde{Q}$. The variations of ξ, η, and B' yield the equations for the waves. The precise form of these equations will depend on additional assumptions that we may wish to make. At leading order, $\delta L_{IG} = 0$ implies the linear dynamics

$$\delta\xi: \qquad \xi_{tt} - f_0\eta_t = -B'_x, \tag{7.4}$$

$$\delta\eta: \qquad \eta_{tt} + f_0\xi_t = -B'_y, \tag{7.5}$$

$$\delta B': \qquad \partial_z(N_0^{-2}B'_z) = \xi_x + \eta_y, \tag{7.6}$$

which are equivalent to (4.4)-(4.6). We may consistently use (7.4)-(7.6) to simplify the coupling Lagrangian (6.15).

The strongest possible assumption appears to be the WKB assumption of a separation in scale with respect to all four of the independent variables x, y, z, t. Then the averaging operator corresponds to an average over wave phase. Let ϵ be the scale-separation parameter, as in Section 4. Then the last two terms in (6.15) are $\mathrm{O}(\epsilon^2)$ because they involve two spatial derivatives of phase averages, whereas the $\bar{\psi}M$-term in (6.15) is $\mathrm{O}(\epsilon)$,

as already noted in Section 4. Thus, under WKB scaling,

$$L_C \approx \iiiint d\mathbf{x}dzdt\ \left(-\bar{\psi}\nabla\times\langle\boldsymbol{\xi}_t(\xi_x+\eta_y)\rangle\right)$$
$$\approx \iiiint d\mathbf{x}dzdt\ \left(-\bar{\psi}\nabla\times\langle(\xi_x\xi_t,\eta_y\eta_t)\rangle\right) = \iiiint d\mathbf{x}dzdt\ \left(\bar{\psi}\nabla\times\mathbf{p}\rangle\right), \tag{7.7}$$

where

$$\mathbf{p} = -(\langle\xi_x\xi_t\rangle,\langle\eta_y\eta_t\rangle). \tag{7.8}$$

The second approximation in (7.7) uses the fact that the variables ξ and η are 90° out of phase, as may be shown from (7.4)-(7.6). Using (7.4)-(7.6) and $\partial_x\langle\rangle = 0$, etc., we can rewrite the rhs of (7.8) in several ways, including the one preferred by Bühler & McIntyre (1998). In the case of a slowly varying wavetrain we find that $\mathbf{p} = E\mathbf{k}/\omega$ as in Section 4. The variational principle generalises the results obtained in Section 4. The equation

$$\delta\bar{\psi}: \qquad \nabla^2\bar{\psi} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N_0^2}\frac{\partial\bar{\psi}}{\partial z}\right) - \nabla\times\mathbf{p} = \tilde{Q} \tag{7.9}$$

replaces (4.31), and we now have the equation (7.3) for advection of APV. The dispersion relation (4.23) and the action equation (4.24) are unchanged. These results are consistent with Bühler & McIntyre (1998). They obtain (7.3) and (7.9), but their wave equations differ from (4.23) and (4.24) because they assume that mean flow is $O(1)$. For reasons that are further explained in Section 8, we have assumed that the mean flow is $O(a^2)$.

A weaker assumption than the WKB assumption is that the waves have frequencies near the inertial frequency f_0. Such waves vary rapidly in t and z, but not in x or y. Assuming that the average corresponds to an average over fast time and fast z, we can consistently neglect the z-derivatives of averages in the coupling Lagrangian (6.15). Then

$$L_C = \iiiint d\mathbf{x}dzdt\ \bar{\psi}\left(\nabla^2\langle\xi\eta_t\rangle - \nabla\times\langle(\xi_x+\eta_y)\boldsymbol{\xi}_t\rangle\right). \tag{7.10}$$

We use the equation (7.6) to eliminate B' from (6.14). Then (6.14) takes the form

$$L_{IG} = \iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}\langle\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t\rangle + f_0\langle\xi\eta_t\rangle - \tfrac{1}{2}N_0^2\langle\left(\int dz(\xi_x+\eta_y)\right)^2\rangle\right). \tag{7.11}$$

In near-inertial motion the last term in (7.11) is small. With this in mind we rewrite the complete Lagrangian as

$$\langle L_1+L_2\rangle = L_{QG}[\bar{\psi},\tilde{\alpha},\tilde{\beta}] + L_I[\xi,\eta] + L_{CN}[\bar{\psi},\xi,\eta], \tag{7.12}$$

where L_{QG} is given by (6.13),

$$L_I = \iiiint d\mathbf{x}dzdt\ \left(\tfrac{1}{2}\langle\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t\rangle + f_0\langle\xi\eta_t\rangle\right) \tag{7.13}$$

is the Lagrangian for inertia waves, and

$$L_{CN} = \iiiint d\mathbf{x}dzdt\ \left(\bar{\psi}\nabla^2\langle\xi\eta_t\rangle - \bar{\psi}\nabla\times\langle(\xi_x+\eta_y)\boldsymbol{\xi}_t\rangle - \tfrac{1}{2}N_0^2\langle\left(\int dz(\xi_x+\eta_y)\right)^2\rangle\right) \tag{7.14}$$

is the sum of (7.10) and the last term in (7.11). We shall treat all the terms in (7.14) as small terms.

Following Young & Ben Jelloul (1997) and Xie & Vanneste (2015, hereafter XV), we

set

$$\xi + i\eta = \chi_z e^{-if_0 t}. \tag{7.15}$$

The exponential factor in (7.15) represents the fast time dependence of inertial motion. The complex coefficient $\chi(x, y, z, t)$ varies slowly in x, y, t but rapidly in z. From (7.15) we have

$$\xi_t + i\eta_t = -if_0\chi_z e^{-if_0 t} + \chi_{zt} e^{-if_0 t} \tag{7.16}$$

and

$$\int dz(\xi_x + \eta_y) = \tfrac{1}{2}\left(\chi_x - i\chi_y\right) e^{-if_0 t} + \text{cc}, \tag{7.17}$$

where cc denotes the complex conjugate. We substitute (7.16) and (7.17) back into (7.13) and (7.14), computing the averages as averages over the fast time dependence of the exponential factors. In (7.13) we must keep terms of the first and second order. In (7.14) we keep only the leading order terms. The Coriolis parameter f_0 serves as a convenient ordering parameter. Thus in (7.13) we set

$$\tfrac{1}{2}\langle \boldsymbol{\xi}_t \cdot \boldsymbol{\xi}_t \rangle = \tfrac{1}{2} f_0^2 \langle \chi_z \chi_z^* \rangle + i f_0 \langle \chi_{zt} \chi_z^* \rangle \tag{7.18}$$

and

$$\langle \xi \eta_t \rangle = -\tfrac{1}{2} f_0 \langle \chi_z \chi_z^* \rangle - i\tfrac{1}{2} \langle \chi_{zt} \chi_z^* \rangle, \tag{7.19}$$

and in (7.14) we set

$$\langle \xi \eta_t \rangle = -\tfrac{1}{2} f_0 \langle \chi_z \chi_z^* \rangle, \tag{7.20}$$

$$\nabla \times \langle \boldsymbol{\xi}_t (\xi_x + \eta_y) \rangle = \tfrac{1}{2} i f_0 \langle J(\chi_z^*, \chi_z) \rangle - \tfrac{1}{4} f_0 \nabla^2 \langle \chi_z \chi_z^* \rangle, \tag{7.21}$$

and

$$\langle \left(\int dz(\xi_x + \eta_y) \right)^2 \rangle = \tfrac{1}{2} \langle \nabla\chi \cdot \nabla\chi^* \rangle. \tag{7.22}$$

Here * denotes complex conjugation and the averaging symbols enclose the rapid z-dependence of χ. In simplifying the terms we have freely used integrations by parts with respect to the rapid variations in z and t, and the fact that the averages are averages over these rapid variations. With these substitutions (7.13) takes the form

$$L_I = \iiiint d\mathbf{x} dz dt \; \tfrac{1}{2} i f_0 \langle \chi_{zt} \chi_z^* \rangle, \tag{7.23}$$

and (7.14) takes the form

$$L_{CN} = \iiiint d\mathbf{x} dz dt \; \left(\tfrac{1}{2} i f_0 \bar{\psi} \langle J(\chi_z, \chi_z^*) \rangle - \tfrac{1}{4} f_0 \bar{\psi} \nabla^2 \langle \chi_z \chi_z^* \rangle - \tfrac{1}{4} N_0^2 \langle \nabla\chi \cdot \nabla\chi^* \rangle \right). \tag{7.24}$$

The complete Lagrangian for near-inertial motion is

$$\begin{aligned} &L_{QG}[\bar{\psi}, \tilde{\alpha}, \tilde{\beta}] + L_I[\chi] + L_{CN}[\bar{\psi}, \chi] \\ &\quad = \iiiint d\mathbf{x} dz dt \; \big(-\tfrac{1}{2} \bar{\psi} \mathcal{L} \bar{\psi} - \tilde{\alpha} \tilde{\beta}_t + \bar{\psi} J(\tilde{\alpha}, \tilde{\beta}) + \frac{i f_0}{2} \langle \chi_{zt} \chi_z^* \rangle - \frac{N_0^2}{4} \langle \nabla\chi \cdot \nabla\chi^* \rangle \\ &\qquad\qquad + \frac{i f_0}{2} \bar{\psi} \langle J(\chi_z, \chi_z^*) \rangle - \frac{f_0}{4} \bar{\psi} \nabla^2 \langle \chi_z \chi_z^* \rangle \big). \end{aligned} \tag{7.25}$$

The variations of $\tilde{\alpha}$ and $\tilde{\beta}$ lead to (7.3), as in general. The variations of $\bar{\psi}$ and χ yield

$$\delta\bar{\psi}: \qquad \mathcal{L}\bar{\psi} + \frac{if_0}{2}\langle J(\chi_z^*, \chi_z)\rangle + \frac{f_0}{4}\nabla^2\langle\chi_z\chi_z^*\rangle = \tilde{Q} \tag{7.26}$$

and

$$\delta\chi^*: \qquad \chi_{zzt} + \frac{iN_0^2}{2f_0}\nabla^2\chi + J(\bar{\psi}, \chi_z)_z + \frac{i}{2}\left(\chi_z\nabla^2\bar{\psi}\right)_z = 0. \tag{7.27}$$

(The variations $\delta\chi$ and $\delta\chi^*$ may be taken independently, but they lead to the same result.) Note that the terms in (7.27) vary rapidly in z. The equations (7.3), (7.26) and (7.27) are very close to the equations derived by XV. In our notation the XV equations comprise (7.3),

$$\mathcal{L}\bar{\psi} + \frac{if_0}{2}\langle J(\chi_z^*, \chi_z)\rangle + \frac{f_0}{4}\langle 2\nabla\chi_z\cdot\nabla\chi_z^* - \chi_{zz}\nabla^2\chi^* - \chi_{zz}^*\nabla^2\chi\rangle = \tilde{Q} \tag{7.28}$$

and

$$\chi_{zzt} + \frac{iN_0^2}{2f_0}\nabla^2\chi + J(\bar{\psi}, \chi_z)_z + \frac{i}{2}\left(\bar{\psi}_{zz}\nabla^2\chi + \chi_{zz}\nabla^2\bar{\psi} - 2\nabla\bar{\psi}_z\cdot\nabla\chi_z\right) = 0. \tag{7.29}$$

The equation (7.29) was previously derived by Young & Ben Jelloul (1997), who did not consider the effect of the waves on the mean flow. The XV equations correspond to a Lagrangian, namely

$$\begin{aligned} L_{XV} = \iiint d\mathbf{x}dzdt\, \big(&-\tfrac{1}{2}\bar{\psi}\mathcal{L}\bar{\psi} - \tilde{\alpha}\tilde{\beta}_t + \bar{\psi}J(\tilde{\alpha}, \tilde{\beta}) + \frac{if_0}{2}\langle\chi_{zt}\chi_z^*\rangle - \frac{N_0^2}{4}\langle\nabla\chi\cdot\nabla\chi^*\rangle \\ &+ \frac{if_0}{2}\bar{\psi}\,\langle J(\chi_z, \chi_z^*)\rangle - \frac{f_0}{4}\bar{\psi}\,\nabla^2\langle\chi_z\chi_z^*\rangle + \frac{f_0}{4}\bar{\psi}\,\langle\chi_z\nabla^2\chi^* + \chi_z^*\nabla^2\chi\rangle_z\big), \end{aligned} \tag{7.30}$$

that differs from (7.25) by a term with a factor of the form $\partial_z\langle\rangle$. This term is negligible if, as we have assumed, there is a vertical scale separation between the quasigeostrophic mean flow and the near-inertial waves. Similarly, the differences between (7.26) and (7.28), and between (7.27) and (7.29), are insignificant under this assumption. The important point is that both sets of equations—ours and those of XV—being equivalent to variational principles, correspond to a consistent set of conservation laws. This fact was used to great advantage by XV.

Finally, we consider the case in which we make no additional assumptions beyond the separation in timescales used to derive (6.12)-(6.15). Then all of the terms in (6.15) must be kept. If we keep all of the terms in (6.15), then the requirement that (6.12) be stationary with respect to variations in $\bar{\psi}$ yields the generalization,

$$\mathcal{L}\bar{\psi} + q^w = \tilde{Q}, \tag{7.31}$$

of (7.9) and (7.26), where

$$\begin{aligned} q^w &= -M - \partial_z\frac{f_0}{N_0^2}T_z - \mathcal{L}\langle\xi\eta_t\rangle \\ &= \nabla\times\langle(\xi_x + \eta_y)\boldsymbol{\xi}_t\rangle - \nabla^2\langle\xi\eta_t\rangle - \partial_z\frac{f_0}{N_0^2}\partial_z\left(f_0\langle\xi\eta_t\rangle + \tfrac{1}{2}\langle\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t\rangle\right). \end{aligned} \tag{7.32}$$

Using (7.4)-(7.6) and $\partial_t\langle\rangle = 0$ we find that

$$\tfrac{1}{2}\langle\boldsymbol{\xi}_t\cdot\boldsymbol{\xi}_t\rangle = -\tfrac{1}{2}\langle\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{tt}\rangle = -f_0\langle\xi\eta_t\rangle + \tfrac{1}{2}\langle\boldsymbol{\xi}\cdot\nabla B'\rangle. \tag{7.33}$$

By similar manipulations,

$$\nabla\times\langle(\xi_x + \eta_y)\boldsymbol{\xi}_t\rangle - \nabla^2\langle\xi\eta_t\rangle = \langle J(\xi_t, \xi)\rangle + \langle J(\eta_t, \eta)\rangle + \nabla\cdot\langle(\eta_x - \xi_y)\boldsymbol{\xi}_t\rangle. \tag{7.34}$$

Thus (7.32) may be written

$$q^w = \langle J(\xi_t,\xi)\rangle + \langle J(\eta_t,\eta)\rangle + \nabla\cdot\langle(\eta_x-\xi_y)\boldsymbol{\xi}_t\rangle - \left(\frac{f_0}{2N_0^2}\langle\boldsymbol{\xi}\cdot\nabla B'\rangle_z\right)_z \tag{7.35}$$

Assuming only the separation in timescales, Wagner & Young (2015, hereafter WY) derive (7.31) with

$$q^w = \langle J(\xi_t,\xi)\rangle + \langle J(\eta_t,\eta)\rangle + f_0\langle J(\xi,\eta)\rangle + \frac{f_0}{2}\langle\xi_i\xi_j\rangle_{,i,j}. \tag{7.36}$$

We must show that our (7.35) agrees with their (7.36). WY work in Cartesian coordinates with fluid particle displacements $(\xi_1,\xi_2,\xi_3)=(\xi,\eta,\zeta)$. (The vertical displacement ζ must not be confused with the relative vorticity denoted by the same symbol earlier in the paper.) Although the first three terms on the rhs of (7.36) involve only x- and y-derivatives, the last term in (7.36) contains index summations from 1 to 3. Thus the connection between (7.35) and (7.36) is not obvious. In Appendix B we show that (7.35) and (7.36) are in fact equivalent.

8. Discussion

This paper extends the familiar analogy between electric charge and quasigeostrophic potential vorticity to include the effects of inertia-gravity waves. In this analogy quasigeostrophic dynamics corresponds to a kind of Coulomb dynamics in which slowly moving charges interact via action-at-a-distance forces. The introduction of inertia-gravity waves alters this picture in two ways. First, nonlinear interactions between the waves—a feature not present in electrodynamics—can temporarily create a virtual electric charge, i.e. Bretherton flow. This charge is captured by the mean flow if dissipation removes the waves, but even if no dissipation occurs the virtual charge interacts with other charges in the fluid, as beautifully illustrated in the examples analyzed by Bühler & McIntyre (2003). Second, the Stokes drift of the waves contributes to the advection of the charges. It is remarkable that both of these effects, here computed separately as coupling contributions to L_1 and L_2, appear as contributions to q^w in (7.31).

One could extend our theory by including more of the terms that occur in the expansion (4.3) of L_1 and the Taylor expansion of L_2 as given by (5.2). In these expansions we have kept only cubic terms of the form 'bar-prime-prime'. This approximation, also introduced via scaling assumptions by WY and XV, assumes that the quasigeostrophic flow is weak compared to the inertia-gravity waves. Although this is certainly correct for Bretherton flow, oceanographers would be happy to see terms of the form 'bar-bar-prime-prime'. Unfortunately the number of such terms is daunting. Besides the desire to avoid complexity, our choice of coupling terms was made in the hope of recovering the results of WY and XV. In the WKB limit, we recover results very similar to those of Bühler & McIntyre (1998) despite their scaling assumption of a stronger mean flow. All three of these papers are challenging papers whose precise methods seem to have little in common. Our ability to recover similar results suggests that the present paper represents a valid synthesis.

Despite the agreement with other work, our methods of approximating the Lagrangian will strike many readers as ad hoc. The development of a more rigorous approach faces two fundamental difficulties. First, there seem to be no dependable rules determining when it is legal to substitute the results of a variational principle back into the variational principle itself. Second, it is useless to apply formal scaling theory to the terms in the Lagrangian because each such term typically contributes to several equations. The

importance of this contribution varies from equation to equation and thus cannot be gauged from the size of the term in the Lagrangian.

The most accepted method in our field is asymptotic expansion applied directly to the fluid equations, but there is a sense in which that method too is ad hoc. Virtually all interesting solutions of the fluid equations share the property of turbulence that small differences in initial conditions lead to completely different solutions after a finite time. It must equally be true that the small changes introduced by even the best approximations lead to solutions that eventually differ significantly from the exact solution. The most that can be hoped for is that the approximate solutions resemble the exact solution in a statistical sense. And this is likely to be true only if the approximation maintains conservation laws. Thus we should distinguish between an approximation that exactly conserves an approximate form of energy and an approximation that only approximately conserves anything. The energy of the latter is likely to increase indefinitely, leading to an enormous statistical divergence from the exact solution. As ad hoc as they may appear to be, Lagrangian approximation methods are the champions at maintaining conservation laws. So long as the corresponding symmetry property is not disturbed, the conservation law survives.

Finally, it should be emphasized that all theories of the general type discussed in this paper are 'low-energy' theories in the sense that only at low energy can the flow even approximately be considered to consist of quasigeostrophic dynamics weakly coupled to free inertia-gravity waves. At higher system energy the two interact in ways that change the fundamental character of the flow. Typically the free inertia-gravity waves are replaced by a balanced dynamics much more complicated than quasigeostrophy, in which much of the flow field is slaved to the potential vorticity; see especially McIntyre & Norton (2000).

Acknowledgements

This work was inspired by the elegant papers of Bühler & McIntyre (1998), Xie & Vanneste (2015), and Wagner & Young (2015). I thank G. L. Wagner, W. R. Young, J.-H. Xie, and three anonymous referees for valuable comments.

Appendix A. Analogy with electrodynamics

This appendix clarifies the connection between S14 and the present work. S14 considered a simplified shallow-water dynamics with continuity equation

$$\hat{h}_t + \nabla \cdot \mathbf{u} = 0. \tag{A1}$$

However, the same strategy applies to the exact continuity equation (2.2). If we rewrite (2.2) as a statement of vanishing spacetime divergence,

$$(\partial_t, \partial_x, \partial_y) \cdot \left(\hat{h}, \hat{h}u, \hat{h}v\right) = 0, \tag{A2}$$

then it follows from (A2) that

$$\left(\hat{h}, \hat{h}u, \hat{h}v\right) = (\partial_t, \partial_x, \partial_y) \times (-\psi, A, B) \tag{A3}$$

for some vector $(-\psi, A, B)$. That is,

$$\hat{h} = B_x - A_y, \tag{A4}$$

$$\hat{h}u = -\psi_y - B_t, \tag{A5}$$

$$\hat{h}v = \psi_x + A_t. \tag{A6}$$

Following steps similar to those in S14, it is straightforward to show that the non-rotating shallow water equations, with vorticity confined to delta functions located at $\mathbf{x}_i(t)$, are equivalent to the requirement that

$$L = \iiint d\mathbf{x}dt \ \left(\tfrac{1}{2}\frac{(A_t+\psi_x)^2}{(B_x-A_y)} + \tfrac{1}{2}\frac{(B_t+\psi_y)^2}{(B_x-A_y)} - \tfrac{1}{2}c^2(B_x-A_y)^2 \right) + \sum_i \Gamma_i \int dt \, (\psi(\mathbf{x}_i(t),t) - A(\mathbf{x}_i(t),t)\dot{x}_i(t) - B(\mathbf{x}_i(t),t)\dot{y}_i(t)) \tag{A7}$$

be stationary with respect to variations in ψ, A, B, and $\mathbf{x}_i$.

If we set $B_x - A_y = 1$ (i.e. $h = h_0$) in the denominators of (A7), then the resulting

$$L = \iiint d\mathbf{x}dt \ \left(\tfrac{1}{2}(A_t+\psi_x)^2 + \tfrac{1}{2}(B_t+\psi_y)^2 - \tfrac{1}{2}c^2(B_x-A_y)^2 \right) + \sum_i \Gamma_i \int dt \, (\psi(\mathbf{x}_i(t),t) - A(\mathbf{x}_i(t),t)\dot{x}_i(t) - B(\mathbf{x}_i(t),t)\dot{y}_i(t)) \tag{A8}$$

is equivalent to the Lagrangian given in S14. The dynamics corresponding to (A8) were called 'wave-vortex dynamics' in S14. Because the first line in (A7) is non-quadratic in the potentials ψ, A, and B, gravity waves interact nonlinearly in exact shallow-water dynamics, even in the absence of vorticity. However, because the first line in (A8) is quadratic, the gravity waves in wave-vortex dynamics are linear waves that never interact. In that respect they resemble electrodynamic waves, and, as noted in S14, the Lagrangian (A8) is in fact closely analogous to the Lagrangian for classical electrodynamics. See, for example, Landau & Lifshitz (1975, pp. 67-69). In this analogy, Γ_i corresponds to the charge on the particle located at $\mathbf{x}_i(t)$. In fact, the only difference between (A8) and the Lagrangian for classical electrodynamics in two space dimensions is that (A8) contains no term analogous to

$$\int mc \, ds, \tag{A9}$$

where m is the mass of the electron, c is the speed of light, and ds is the differential of proper time.

The vector $(-\psi, A, B)$ is not unique; to it we may add a vector $(\partial_t, \partial_x, \partial_y)\lambda(\mathbf{x},t)$ without affecting (A3). Thus we are free to add a gauge condition. S14 adopted the Lorentz gauge. In the present paper, we choose the Coulomb gauge,

$$A_x + B_y = 0. \tag{A10}$$

Then

$$(A, B) = (\gamma_y, x - \gamma_x) \tag{A11}$$

for some γ. The x-term in (A11) just separates out the mean depth. Substituting (A11) into (A4)-(A6), we obtain (2.8)-(2.10). Substituting (A11) into (A7) we obtain (2.18) with $f_0 = 0$. The choice of the Coulomb gauge shifts the emphasis from the momentum equations to the vorticity and divergence equations.

The introduction of nonvanishing coordinate system rotation is easy at any stage of the development. However, $f_0 \neq 0$ precipitates a choice between physically equivalent formulations of the dynamics. S14 introduced charges as lumps of potential vorticity. In the present paper the charges represent lumps of APV. These two approaches are equally valid, but if the former is chosen, then coordinate system rotation enters as a background medium with a uniform electric charge. The inertia-gravity waves become gravity waves

that continually interact with the background charge. This proves to be a clumsy way of looking at inertia-gravity waves.

Because no term like (A9) appears in (A7), (2.16), which is analogous to the Lorenz force law for the i-th charged particle,

$$m_i \ddot{\boldsymbol{x}}_i = q_i \left(\mathbf{E} + \dot{\boldsymbol{x}}_i \times \mathbf{B} \right), \tag{A12}$$

reduces to the analogue of

$$0 = \mathbf{E} + \dot{\boldsymbol{x}}_i \times \mathbf{B}, \tag{A13}$$

in which the electric force cancels the magnetic force. That is, (2.16) lacks the acceleration term $\ddot{\mathbf{x}}_i$ because point vortices, unlike electrons, are massless.

In overall summary, fluid dynamics is more complicated than electrodynamics in that the fluid waves interact with each other, even in the absence of charge (i.e. vorticity). However, electrodynamics is more complicated than fluid dynamics in that electrons, unlike point vortices, have mass.

The Lagrangians given in the present paper are believed to be new. They differ from the more conventional Lagrangians for a perfect fluid, which typically depend upon labeled fluid particles or upon a set of Clebsch potentials. See, for example, Salmon (1998, Chapter 7). Given the analogy between fluid dynamics and classical electrodynamics developed in the present paper, one could ask whether electrodynamics has a variational principle like the conventional variational principle of fluid dynamics, that is, solely in terms of the labeled locations of charged particles, and not involving the electromagnetic potentials at all. The answer, of course, is yes. It is the Fokker action principle discussed in the famous paper by Wheeler & Feynman (1949).

Appendix B. Agreement with the result of Wagner and Young

In this appendix we show that our (7.35) agrees with WY's (7.36). WY work in Cartesian coordinates. In Cartesian coordinates, the linear wave dynamics are

$$\xi_{tt} - f_0 \eta_t = -p_x, \tag{B1}$$

$$\eta_{tt} + f_0 \xi_t = -p_y, \tag{B2}$$

$$p_z = b, \tag{A3}$$

$$b_t + N_0^2 \zeta_t = 0, \tag{B4}$$

$$\xi_x + \eta_y + \zeta_z = 0, \tag{B5}$$

where p is the pressure and b is the buoyancy. If we define

$$p = B', \tag{B6}$$

$$\zeta = -N_0^{-2} B'_z, \tag{B7}$$

$$b = B'_z, \tag{B8}$$

then (7.4)-(7.6) are formally identical to (B1)-(B5). (We can ignore the difference between the derivatives in the two coordinate systems; this difference is negligible at the order of q^w.) Directly from (B1)-(B5), WY demonstrate that

$$\langle \boldsymbol{\xi}_3 \cdot \nabla_3 p \rangle_z = -2N_0^2 \langle \boldsymbol{\xi}_3 \cdot \nabla_3 \zeta + \frac{1}{2} \zeta^2 \partial_z \ln N_0^2 \rangle, \tag{B9}$$

where $\boldsymbol{\xi}_3 = (\xi, \eta, \zeta)$ and $\nabla_3 = (\partial_x, \partial_y, \partial_z)$. (This is their equation (A10).) It follows from (B7) and (B9) that

$$\langle \boldsymbol{\xi} \cdot \nabla B' \rangle_z = -2N_0^2 \langle \boldsymbol{\xi} \cdot \nabla \zeta \rangle. \tag{B10}$$

The linear dynamics (7.4)-(7.6) or (B1)-(B5) imply

$$(\eta_x - \xi_y)_t = -f_0(\xi_x + \eta_y). \tag{B11}$$

It follows from (B11) that

$$\nabla \cdot \langle (\eta_x - \xi_y)\boldsymbol{\xi}_t \rangle = f_0 \nabla \cdot \langle (\xi_x + \eta_y)\boldsymbol{\xi} \rangle. \tag{B12}$$

Substituting (B10) and (B12) into (7.35) we obtain

$$q^w = \langle J(\xi_t, \xi) \rangle + \langle J(\eta_t, \eta) \rangle + f_0 \nabla \cdot \langle (\xi_x + \eta_y)\boldsymbol{\xi} \rangle + f_0 \langle \boldsymbol{\xi} \cdot \nabla \zeta \rangle_z. \tag{B13}$$

Using (B5) it is straightforward to show that (B13) is equivalent to WY's result (7.36).

REFERENCES

Bretherton, F. P. 1969 On the mean motion induced by internal gravity waves. *J. Fluid Mech.* **36**, 785-803.

Bühler, O. & McIntyre, M. E. 1998 On non-dissipative wave-mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**, 301-343.

Bühler, O. & McIntyre, M. E. 2003 Remote recoil: a new wave-mean interaction effect. *J. Fluid Mech.* **492**, 207-230.

Bühler, O. & McIntyre, M. E. 2005 Wave capture and wave-vortex duality. *J. Fluid Mech.* **534**, 67-95.

Landau, L. D. & Lifshitz, E. M. 1975 The Classical Theory of Fields. Pergamon, 397 pp.

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

Salmon, R. 1998 Lectures on Geophysical Fluid Dynamics. Oxford, 372 pp.

Salmon, R. 2014 Analogous formulation of electrodynamics and two-dimensional fluid dynamics. *J. Fluid Mech.* **761**, R2, 1-12.

Wagner, G. L. & Young, W. R. 2015 Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**, 401-424.

Wheeler, J. A. & Feynman, R. P. 1949 Classical electrodynamics in terms of direct interparticle action. *Rev. Mod. Phys.* **21**, 3, 425-433.

Whitham, G. B., 1965 A general approach to linear and nonlinear waves using a Lagrangian. *J. Fluid Mech.* **22**, 273-283.

Whitham, G. B., 1974 Linear and Nonlinear Waves. John Wiley, 628 pp.

Xie, J.-H. & Vanneste, J. 2015 A generalized-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143-169.

Young, W. R. 2012 An exact thickness-weighted average formulation of the Boussinesq equations. *J. Phys. Ocean.* **42**, 692-707.

Young, W. R. & Ben Jelloul, M. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55**, 735-766.