Variational balance models for the three-dimensional Euler-Boussinesq equations with full Coriolis force

Gözde Özden[a,b,c] and Marcel Oliver[a]

[a]School of Engineering and Science, Jacobs University, 28759 Bremen, Germany; [b]MARUM – Center for Marine Environmental Sciences, Universität Bremen, 28334 Bremen, Germany; [c]Meteorological Institute, Universität Hamburg, 20144 Hamburg, Germany

ARTICLE HISTORY

Published in Phys. Fluids 33 (2021), 076606.

ABSTRACT

We derive a semi-geostrophic variational balance model for the three-dimensional Euler–Boussinesq equations on the non-traditional f-plane under the rigid lid approximation. The model is obtained by a small Rossby number expansion in the Hamilton principle, with no other approximations made. We allow for a fully non-hydrostatic flow and do not neglect the horizontal components of the Coriolis parameter, that is, we do not make the so-called "traditional approximation". The resulting balance models have the same structure as the "L_1 balance model" for the primitive equations: a kinematic balance relation, the prognostic equation for the three-dimensional tracer field, and an additional prognostic equation for a scalar field over the two-dimensional horizontal domain which is linked to the undetermined constant of integration in the thermal wind relation. The balance relation is elliptic under the assumption of stable stratification and sufficiently small fluctuations in all prognostic fields.

KEYWORDS

Boussinesq equation, full Coriolis force, variational asymptotics, balance models

1. Introduction

We describe the derivation of a semi-geostrophic variational balance model for the three-dimensional Euler–Boussinesq equations on the non-traditional f-plane under the rigid lid approximation. In non-dimensional variables, the Euler–Boussinesq system takes the form

$$\varepsilon \left(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u}\right) + \boldsymbol{\Omega} \times \boldsymbol{u} = -\boldsymbol{\nabla} p - \rho \, \boldsymbol{k} \,, \tag{1a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \,, \tag{1b}$$

$$\partial_t \rho + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho = 0 \,. \tag{1c}$$

where $\varepsilon = U/(fL)$ is the Rossby number (here, U denotes a typical horizontal velocity scale, L a typical horizontal length scale, and f the Coriolis frequency in physical

Corresponding author: Gözde Özden, goezde.oezden@uni-hamburg.de

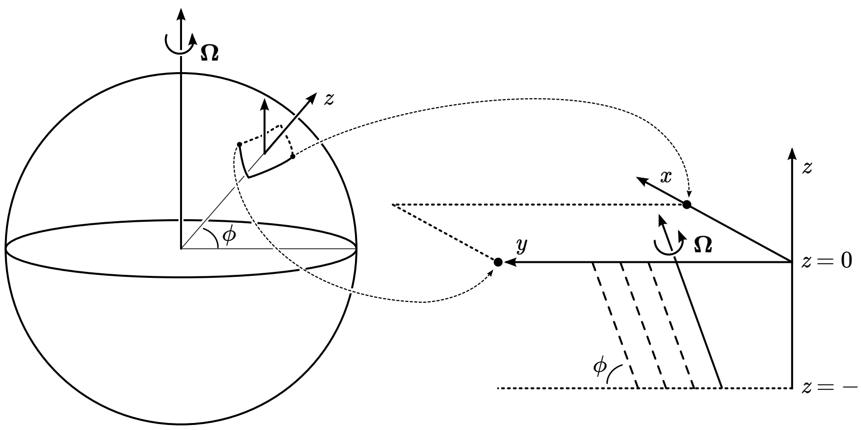

Figure 1. Geometry of the f-plane approximation at latitude ϕ with the full Coriolis vector. Here, x, y, and z are directed toward the east, the north, and upward, respectively. The dashed lines parallel to the axis of rotation in the y-z plane represent the characteristics of the thermal wind relation.

units), assumed small, $\boldsymbol{u}$ is the three-dimensional velocity field,

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ \cos\phi \\ \sin\phi \end{pmatrix} \equiv \begin{pmatrix} 0 \\ c \\ s \end{pmatrix} \tag{2}$$

the full Coriolis vector at constant latitude ϕ which, without loss of generality, is assumed to lie in the y-z plane as shown in Figure 1, p is the pressure, ρ the density, and $\boldsymbol{k}$ the unit vector in the vertical. See, e.g., Majda (2003) or Vallis (2017) for details on the Euler–Boussinesq equations and Franzke, Oliver, Rademacher, and Badin (2019) for an explicit exposition of the semi-geostrophic scaling limit.

For simplicity, we assume periodic boundary conditions in the horizontal and rigid lid boundary conditions in the vertical on a layer of fluid with constant depth,

$$\mathcal{D} = \mathbb{T}^2 \times [-1, 0] \,, \tag{3}$$

so that

$$\boldsymbol{k} \cdot \boldsymbol{u} = 0 \quad \text{at} \quad z = -1, 0 \,. \tag{4}$$

The semi-geostrophic scaling used in (1) corresponds to a regime of strong rotation and weaker stratification, expressed by a Burger number $\mathrm{Bu} = \varepsilon$ (or, equivalently, $\varepsilon = \mathrm{Fr}^2$, where Fr denotes the Froude number). In the more widely studied quasi-geostrophic regime, rotation and stratification are equally important so that the Burger number is $O(1)$. The quasi-geostrophic approximation must be made in conjunction with the assumption that the density is a small perturbation of a constant, stably stratified background density. The quasi-geostrophic limit as it is found in textbooks (e.g. Pedlosky, 1987; Vallis, 2017) is commonly studied under the additional assumption of the hydrostatic and traditional approximations and is widely used for proof-of-concept studies. However, Embid and Majda (1998) have shown that it is possible to derive quasi-geostrophic limit equations without hydrostaticity and with the full Coriolis vector; this limit has a rigorous justification within the framework of averaging over fast scales developed in Embid and Majda (1996). Julien, Knobloch, Milliff, and Werne (2006) systematically explore generalized quasi-geostrophic models by allowing, in contrast to the earlier work of Embid and Majda (1998), different domain aspect ratios and different tilting regimes of the axis of rotation; see Lucas, McWilliams, and Rousseau (2017) for recent rigorous analysis and Nieves, Grooms, Julien, and Weiss (2016) for a numerical study. We also refer the reader to Babin, Mahalov, and Nicolaenko (2002) for a detailed discussion of the different scaling regimes for rotating stratified flow and for a discussion of resonances when averaging over fast waves. A more expository survey can be found in Franzke et al. (2019).

In our setting, there is no need to split off a small perturbation density from a stably stratified background profile *a priori*. However, we shall see that this condition does not disappear altogether but comes back in slightly weaker form as a solvability condition for the balance relation. By keeping the density as a Lagrangian tracer, it is possible to retain the variational formulation of the equation of motion throughout the limit.

As in Embid and Majda (1998), Julien et al. (2006) and Lucas et al. (2017), we use the full Coriolis vector, in contrast to the more common "traditional approximation" where only the vertical component of the Coriolis vector is retained—a consistency requirement if the hydrostatic approximation is made (cf. White, 2002). Vice versa, when the aspect ratio, the ratio of typical horizontal and typical vertical scales is of order one so that the model is fully non-hydrostatic, as is assumed here, both horizontal and vertical components of the Coriolis vector contribute to the leading order on the mid-latitude f-plane. Whitehead and Wingate (2014) perform a more idealized numerical study for the non-hydrostatic Boussinesq equations with only traditional Coriolis forces (a "polar f-plane") in three different limits: the quasi-geostrophic regime, the strongly stratified regime where the Rossby number remains $O(1)$ while the Froude number goes to zero (both limits as considered by Embid & Majda, 1998) as well as the case of strong rotation and weak stratification where the Froude number remains $O(1)$ while the Rossby number goes to zero. Whitehead, Haut, and Wingate (2018) also consider intermediate regimes between quasi-geostrophy and the weak stratification limit, such as the semi-geostrophic limit considered here, within the theoretical framework of Embid and Majda (1996). More recent related theoretical results are due to Ju and Mu (2019). Wetzel, Smith, Stechmann, and Martin (2019) perform a numerical study of balance for a moist atmosphere with phase changes, and Kafiabad and Bartello (2016, 2017, 2018) and Kafiabad, Vanneste, and Young (2021) study the exchange of energy between balanced and unbalanced motion.

Our derivation is based on variational asymptotics. Approximations to Hamilton's variational principle for the equations of rotating fluid flow were pioneered by Salmon (1983, 1985) in the context of the shallow water equations and later, in Salmon (1996), for the primitive equations of the ocean. The basic idea is to consider the "extended Hamilton principle" or "phase space variational principle" and use the leading order balance relation (geostrophic balance for the shallow water equations or the thermal wind relation for the primitive equations) to constrain the momentum variables. The stationary points of the constrained action with respect to variations of the position variables give a balance relation that includes the next-order correction to the leading order constraint. In principle, this method can be iterated to obtain higher-order balance relations. For example, Allen, Holm, and Newberger (2002) derive second order models and show that the so-called L_1 and L_2 models are numerically well behaved.

A second idea, also proposed in Salmon (1985), is the use of a near-identity coordinate transformation to bring the resulting variational principle or the equations of motion into a more convenient form. This transformation may be applied perturbatively, so that the resulting models coincide only up to terms of a certain order. From an analytic perspective, higher-order terms may matter and it turns out that a transformation to a coordinate system in which the Hamilton equations of motion take canonical form, the original motivation behind this approach, leads to ill-posedness of the full system of prognostic equations.

An even more general framework is obtained by reversing the steps "constrain" and "transform". Oliver (2006) noted that is possible to assume an entirely general near-identity change of coordinates and expand the transformed Lagrangian in powers of the small parameter. Whenever the transformation is chosen such that the Lagrangian is degenerate to the desired order, the variational principle *implies* a phase space constraint. This approach gives rise to a greater variety of candidate models that are not accessible via Salmon's approach. In particular, it allows us to retain some important mathematical features such as regularity of potential vorticity inversion. In the context of the shallow water equations on the f-plane, it turns out that the L_1 model first proposed by Salmon is already optimal among a larger family of models (Dritschel, Gottwald, & Oliver, 2017). In other configurations, such as when the Coriolis parameter is spatially varying (Oliver & Vasylkevych, 2013), or in the fully three dimensional situation considered here, the additional flexibility that comes from putting the change of coordinates first is necessary.

While our scaling is consistent with the geostrophic momentum approximation (Eliassen, 1948), the derivation and the resulting equations are different. Hoskins (1975) has shown that the geostrophic momentum approximation can be formulated as potential vorticity advection in transformed coordinates, known as the semigeostrophic equations. His transformation has subsequently been interpreted as a Legendre transformation, providing a sense of generalized solutions via the theory of optimal transport (Benamou & Brenier, 1998; Colombo, 2017; Cullen, 2006; Cullen & Purser, 1984; Roulstone & Sewell, 1997). On the other hand, L_1-type models, at least in the shallow water context, have strong solutions in a classical sense (Çalık, Oliver, & Vasylkevych, 2013). Semi-geostrophic theory can be seen as a particular choice of truncation and transformation in the framework of variational asymptotics (Oliver, 2014); for a different view on generalized semigeostrophic theory, see McIntyre and Roulstone (2002).

In this work, we show that variational asymptotics in the semi-geostrophic regime can be done directly for the three-dimensional Euler–Boussinesq system with full Coriolis force without any preparatory approximations. Our motivation derives, first, from the role of the Boussinesq equation as the common parent model of nearly all of geophysical fluid dynamics. Second, we would like to see how "non-traditional" Coriolis effects, associated with the vertical component of the Coriolis force and horizontal Coriolis forces coming from the vertical velocity, enter the balance dynamics. Non-traditional effects are significant in a variety of circumstances (Brummell, Hurlburt, & Toomre, 1996; de laTorre Juárez, Fisher, & Orton, 2002; Gerkema & Shrira, 2005a, 2005b; Gerkema, Zimmerman, Maas, & van Haren, 2008; Hayashi & Itoh, 2012) and are also studied in the context of other reduced models such as layers of shallow water (Stewart & Dellar, 2010, 2012) and models on the β-plane (Dellar, 2011).

Our approach is similar, in principle, to the primitive equation case (Oliver & Vasylkevych, 2016), and we find that the resulting balance models have the same structure: a kinematic balance relation which is elliptic under suitable assumptions, the prognostic equation for the three-dimensional tracer field, and an additional prognostic equation for a scalar field over the two-dimensional horizontal domain which is linked to the undetermined constant of integration in the thermal wind relation. This field, in the present setting, takes the form of a skewed relative vorticity averaged along the axis of rotation.

On the technical level, the computations are considerably more difficult than for the primitive equation. Some expressions, such as the thermal wind relation, take a natural form in an oblique coordinate system where the axis of rotation takes the role of the "vertical". Other expressions, most notably the top and bottom boundary conditions and the gravitational force, are most easily described in the usual Cartesian coordinates. This incompatibility requires a detailed study of averaging and decomposition of vector fields in oblique coordinates, and of the translations between oblique and Cartesian coordinates.

The remainder of the paper is organized as follows. Section 2 recalls the variational derivation for the Euler–Boussinesq equations in the language of the Euler–Poincaré variational principle. Section 3 explains the general setting for variational asymptotics in this framework. In Section 4, we set up an oblique coordinate system whose vertical direction is aligned with the axis of rotation and introduce the notion of averaging along this axis. Section 5 discusses the leading order balance, the thermal wind relation, on the mid-latitude f-plane with full Coriolis force. In Section 6, we derive the first order balance model Lagrangian. To ensure that the Lagrangian is maximally degenerate, we carefully decompose the expression for kinetic energy into the parts that can be removed by a suitable choice of the transformation vector field and a residual component which cannot be removed. In Section 7, we derive the balance model Euler–Poincaré equations from the balance model Lagrangian. In Section 8, we find that it can be decomposed into a single evolution equation for a scalar field in the two horizontal variables, and a kinematic relationship for all other components. We show that this kinematic relationship is elliptic under the assumption of stable stratification and sufficiently small fluctuations in all prognostic fields. Section 9 discusses the reconstruction of the full velocity field from the balance relation and the prognostic variables. Then, we give a brief discussion and conclusions. Finally, four technical appendixes contain results on averaging along the axis of rotation, the decomposition of vector field in oblique coordinates, associated inner product identities, and some details of the computation of the balance model Lagrangian.

2. Variational principle for the Boussinesq equations

In this section, we recall the derivation of the equations of motion via the Hamilton principle, here in the abstract setting of Euler–Poincaré theory.

We write $\operatorname{Diff}_\mu(\mathcal{D})$ to denote the group of H^s-class volume-preserving diffeomorphisms on $\mathcal{D}$ that leave its boundary invariant. The associated "Lie algebra" of vector fields is the space

$$V_{\text{div}} = \{\boldsymbol{u} \in H^s(\mathcal{D}, \mathbb{R}^3) \colon \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0, \, \boldsymbol{k} \cdot \boldsymbol{u} = 0 \text{ at } z = -1, 0\} \,. \tag{5}$$

Here, H^s is the usual Sobolev space of order s consisting of functions with square integrable weak derivatives up to order s. For $s > 5/2$, $\operatorname{Diff}_\mu(\mathcal{D})$ is a smooth infinite-dimensional manifold and V_{div} is its tangent space at the identity (Ebin & Marsden, 1970; Palais, 1968).

Let $\boldsymbol{\eta} = \boldsymbol{\eta}(\,\cdot\,, t) \in \operatorname{Diff}_\mu(\mathcal{D})$ denote the time-dependent flow generated by a time-dependent vector field $\boldsymbol{u} = \boldsymbol{u}(\,\cdot\,, t) \in V_{\rm div}$, i.e.,

$$\partial_t \boldsymbol{\eta}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{\eta}(\boldsymbol{a}, t), t) \tag{6}$$

or $\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}$ for short; here and in the following we use the symbol "$\circ$" to denote the composition of maps and the dot-symbol to denote the *partial* time derivative. In the Lagrangian description of fluid flow, $\boldsymbol{x}(t) = \boldsymbol{\eta}(\boldsymbol{a}, t)$ is the trajectory of a fluid parcel initially located at $\boldsymbol{a} \in \mathcal{D}$. We retain the letter $\boldsymbol{x}$ for Eulerian positions and $\boldsymbol{a}$ for Lagrangian labels throughout.

To formulate the variational principle, we note that the continuity equation (1c) is equivalent to

$$\rho \circ \boldsymbol{\eta} = \rho_0 \,, \tag{7}$$

where ρ_0 denotes the given initial density field. The Lagrangian for the non-dimensional Euler–Boussinesq system is given by

$$\begin{aligned} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_0) &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \frac{\varepsilon}{2} \, |\dot{\boldsymbol{\eta}}|^2 - \rho \, \boldsymbol{k} \cdot \boldsymbol{\eta} \, d\boldsymbol{a} \\ &= \int_{\mathcal{D}} \boldsymbol{R} \cdot \boldsymbol{u} + \frac{\varepsilon}{2} \, |\boldsymbol{u}|^2 - \rho \, z \, d\boldsymbol{x} \equiv \ell(\boldsymbol{u}, \rho) \,. \end{aligned} \tag{8}$$

The vector $\boldsymbol{R}$ is a vector potential for the Coriolis vector, that is, $\boldsymbol{\nabla} \times \boldsymbol{R} = \boldsymbol{\Omega}$; it arises from the transformation into a rotating frame of reference (e.g. Landau & Lifshitz, 1976). Here, on the f-plane, we can choose

$$\boldsymbol{R} = \frac{1}{2} \mathsf{J} \boldsymbol{x} = \frac{1}{2} \begin{pmatrix} cz - sy \\ sx \\ -cx \end{pmatrix} , \tag{9}$$

where J is the skew-symmetric matrix

$$\mathsf{J} = \begin{pmatrix} 0 & -s & c \\ s & 0 & 0 \\ -c & 0 & 0 \end{pmatrix} . \tag{10}$$

As we see in the second line of (8), the Lagrangian can be written as a functional of $\boldsymbol{u}$ and ρ alone. In this form, it is referred to as the reduced Lagrangian ℓ. Computing variations of the full Lagrangian L with respect to the flow map $\boldsymbol{\eta}$ is equivalent to computing variations of the reduced Lagrangian ℓ with respect to $\boldsymbol{u}$ and ρ subject to so-called Lin constraints. This is summarized in the following theorem.

Theorem 1 (Holm, Marsden, & Ratiu, 1998)**.** *Consider a curve* $\boldsymbol{\eta}$ *in* $\operatorname{Diff}_\mu(\mathcal{D})$ *with Lagrangian velocity* $\dot{\boldsymbol{\eta}}$ *and Eulerian velocity* $\boldsymbol{u} \in V_{\rm div}$. *Then the following are equivalent.*

(i) $\boldsymbol{\eta}$ *satisfies the Hamilton variational principle*

$$\delta \int_{t_1}^{t_2} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_0) \, dt = 0 \tag{11}$$

with respect to variations of the flow map $\delta\boldsymbol{\eta} = \boldsymbol{v} \circ \boldsymbol{\eta}$, *where* $\boldsymbol{v}$ *is a curve in* $V_{\rm div}$ *vanishing at the temporal end points.*

(ii) $\boldsymbol{u}$ *and* ρ *satisfy the reduced Hamilton principle*

$$\delta \int_{t_1}^{t_2} \ell(\boldsymbol{u}, \rho) \, dt = 0 \tag{12}$$

with respect to variations $\delta\boldsymbol{u}$ *and* $\delta\rho$ *that are subject to the Lin constraints* $\delta\boldsymbol{u} = \dot{\boldsymbol{v}} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u}$ *and* $\delta\rho + \boldsymbol{v} \cdot \boldsymbol{\nabla} \rho = 0$, *where* $\boldsymbol{v}$ *is a curve in* $V_{\rm div}$ *vanishing at the temporal end points.*

(iii) $\boldsymbol{m}$ *and* ρ *satisfy the Euler–Poincaré equation*

$$\int_{\mathcal{D}} (\partial_t + \mathcal{L}_{\boldsymbol{u}}) \boldsymbol{m} \cdot \boldsymbol{v} + \frac{\delta \ell}{\delta \rho} \, \mathcal{L}_{\boldsymbol{v}} \rho \, d\boldsymbol{x} = 0 \tag{13}$$

for every $\boldsymbol{v} \in V_{\rm div}$, *where* $\mathcal{L}_{\boldsymbol{u}}$ *is the Lie derivative in the direction of* $\boldsymbol{u}$ *and*

$$\boldsymbol{m} = \frac{\delta \ell}{\delta \boldsymbol{u}} \tag{14}$$

is the momentum one-form.

In the language of vector fields on a region of $\mathbb{R}^3$, (13) reads

$$\int_{\mathcal{D}} \Big(\partial_t \boldsymbol{m} + (\boldsymbol{\nabla} \times \boldsymbol{m}) \times \boldsymbol{u} + \boldsymbol{\nabla} (\boldsymbol{m} \cdot \boldsymbol{u}) + \frac{\delta \ell}{\delta \rho} \, \boldsymbol{\nabla} \rho \Big) \cdot \boldsymbol{v} \, d\boldsymbol{x} = 0 \tag{15}$$

for every $\boldsymbol{v} \in V_{\rm div}$. This implies that the term in parentheses must be zero up to a gradient of a scalar potential ϕ.

For the Euler–Boussinesq Lagrangian (8),

$$\boldsymbol{m} = \frac{\delta \ell}{\delta \boldsymbol{u}} = \boldsymbol{R} + \varepsilon \, \boldsymbol{u} \qquad \text{and} \qquad \frac{\delta \ell}{\delta \rho} = -z \, . \tag{16}$$

Setting $\phi = -p - z\,\rho - \frac{1}{2}\, \varepsilon \, |\boldsymbol{u}|^2$ and using the vector identity $(\boldsymbol{\nabla} \times \boldsymbol{u}) \times \boldsymbol{u} = \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \frac{1}{2} \boldsymbol{\nabla} |\boldsymbol{u}|^2$, we see that (15) implies the Euler–Boussinesq momentum equation (1).

As the Lagrangian L is invariant under time translation, the corresponding Hamiltonian

$$H = \int_{\mathcal{D}} \frac{\delta \ell}{\delta \boldsymbol{u}} \cdot \boldsymbol{u} \, d\boldsymbol{x} - \ell(\boldsymbol{u}, \rho) = \int_{\mathcal{D}} \frac{\varepsilon}{2} \, |\boldsymbol{u}|^2 + z \, \rho \, d\boldsymbol{x} \tag{17}$$

is a constant of the motion. A second conservation law arises via the invariance of the Lagrangian with respect to “particle relabeling”, that is, composition of the flow map with an arbitrary time-independent map in $\mathrm{Diff}_\mu(\mathcal{D})$: Ripa (1981) and Salmon (1982) have shown that this symmetry implies material conservation of the Ertel *potential vorticity*

$$q = (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla} \rho = (\boldsymbol{\Omega} + \varepsilon \, \boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{\nabla} \rho \, , \tag{18}$$

that is, q satisfies the advection equation

$$\partial_t q + \boldsymbol{u} \cdot \boldsymbol{\nabla} q = 0 \, . \tag{19}$$

3. Variational asymptotics

Our variational balance model is based on the following construction. Suppose that the flow of the balance model is related to the flow of the full model via a near-identity change of variables. To be explicit, let $\dot{\boldsymbol{\eta}}_\varepsilon$ denote the Lagrangian velocity and $\boldsymbol{u}_\varepsilon$ the corresponding Eulerian velocity field of the full Euler–Boussinesq flow, that is,

$$\dot{\boldsymbol{\eta}}_\varepsilon = \boldsymbol{u}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \, . \tag{20}$$

The corresponding Lagrangian and Eulerian balance model velocities are denoted $\dot{\boldsymbol{\eta}}$ and $\boldsymbol{u}$, so that

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta} \, . \tag{21}$$

We suppose that the flow map of the full model is related to the flow map of the balance model by a change of coordinates that is the flow of a vector field $\boldsymbol{v}_\varepsilon$ with ε as the flow parameter (thus, formally, ε plays the same role here that t plays above), namely

$$\boldsymbol{\eta}_\varepsilon' = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \, , \qquad \boldsymbol{\eta}_\varepsilon \big|_{\varepsilon=0} = \boldsymbol{\eta} \, . \tag{22}$$

Here and in the following, we use the prime symbol to denote a derivative with respect to ε. Cross-differentiation of (20) and (22) gives

$$\boldsymbol{u}_\varepsilon' = \dot{\boldsymbol{v}}_\varepsilon + \boldsymbol{u}_\varepsilon \cdot \boldsymbol{\nabla} \boldsymbol{v}_\varepsilon - \boldsymbol{v}_\varepsilon \cdot \boldsymbol{\nabla} \boldsymbol{u}_\varepsilon \, . \tag{23}$$

Likewise, differentiation in ε of the Lagrangian density gives

$$\rho_\varepsilon' + \boldsymbol{v}_\varepsilon \cdot \boldsymbol{\nabla} \rho_\varepsilon = 0 \, . \tag{24}$$

These two relations can be seen as the Lin constraints for the ε-flow, cf. the Lin constraints for the t-flow stated in Theorem 1.

At this point, the choice of $\boldsymbol{v}_\varepsilon$ is completely arbitrary and no assumptions have been made. However, we will always restrict ourselves to incompressible transformations which leave the domain invariant, so that $\boldsymbol{v}_\varepsilon \in V_{\text{div}}$.

We now treat ε as a perturbation parameter, expanding (23) and (24) in ε. Likewise, we expand the Lagrangian as a formal power series in ε and use the Lin constraints to eliminate all ε-derivatives of $\boldsymbol{u}$ and ρ. Model reduction is achieved via the following steps:

(i) Truncate the expansion of the Lagrangian at some fixed order. Here, we will only look at the truncation to $O(\varepsilon)$, the first nontrivial case.

(ii) Choose the expansion coefficients of the transformation vector fields such that the resulting Lagrangian is maximally degenerate. When computing the model reduction to first order, the zero-order transformation vector field $\boldsymbol{v} = \boldsymbol{v}_\varepsilon|_{\varepsilon=0}$ is the only choice to be made.

(iii) Apply the Euler–Poincaré variational principle to derive the equations of motion. Since the Lagrangian is degenerate, some degrees of freedom will be kinematic, thus imply a phase-space constraint that can be understood as a so-called Dirac constraint (e.g. Salmon, 1988).

The first two steps are done in reverse order relative to the original method of Salmon (1985) who constrained the system first, using the readily available leading order balance relation, and transformed to different coordinates second. In fact, in his approach, the second step is optional and it turns out that, for f-plane shallow water, the so-called L_1 model, which skips the transformation, is superior to all other models in a more general family (Dritschel et al., 2017). In our viewpoint, the transformation is always necessary. The advantage is that the constraint arises as a consequence of the formalism and does not need to be guessed or derived *a priori*. In simple cases, it is possible to choose the transformation such that it vanishes to the order considered; in this case, we reproduce Salmon's L_1 model. In more complicated cases, in particular in our setting here, it is not possible to cancel all terms in the transformation vector field at the order considered. This means that a direct application of Salmon's method would result in a model with additional spurious prognostic variables. We finally remark that at least in finite dimensions, the procedure is rigorous and the resulting model is correct to the expected order of approximation (Gottwald & Oliver, 2014). For partial differential equations, this question is open; it is clear, though, that additional conditions are necessary.

4. Oblique vertical coordinate and oblique averages

The axis of rotation defines the direction of the characteristics of the thermal wind relation. Here, we introduce notation used throughout the paper to describe the decomposition of vector fields along and perpendicular to this axis, as well as a basic averaging operation along the axis of notation.

Let

$$\mathsf{Q} = \boldsymbol{\Omega}\boldsymbol{\Omega}^{\mathsf{T}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & c^2 & cs \\ 0 & cs & s^2 \end{pmatrix} \tag{25}$$

denote the orthogonal projector onto the direction of the axis of rotation; the projector onto the orthogonal complement is then given by

$$\mathsf{P} = \mathsf{I} - \mathsf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & s^2 & -cs \\ 0 & -cs & c^2 \end{pmatrix} . \tag{26}$$

Note that $\boldsymbol{\Omega}$ spans the kernel of J and is perpendicular to the range of J. Thus, Q is the orthogonal projector onto $\operatorname{Ker} \mathsf{J}$ and P is the orthogonal projector onto $\operatorname{Range} \mathsf{J}$.

We introduce the oblique coordinate system

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \zeta\, \boldsymbol{\Omega} \equiv \mathsf{A}\boldsymbol{\xi} \equiv \boldsymbol{\chi}(\boldsymbol{\xi}) \tag{27}$$

where $\boldsymbol{\xi} \equiv (x, \zeta)^{\mathsf{T}}$, where x and y are the horizontal coordinates of each characteristic line at the surface, ζ is the arclength parameter along the characteristics with $\zeta = 0$ being at the surface and $\zeta = -s^{-1}$ being at the bottom, and

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & s \end{pmatrix} . \tag{28}$$

We note that $\det \mathsf{A} = s$. Then, for any scalar function f,

$$\partial_i (f \circ \boldsymbol{\chi}) = \begin{cases} (\partial_i f) \circ \boldsymbol{\chi} & \text{for } i = x, y \\ (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} f) \circ \boldsymbol{\chi} & \text{for } i = \zeta \end{cases} \tag{29}$$

so that, for arbitrary vector fields $\boldsymbol{u}$, $\boldsymbol{v}$,

$$\boldsymbol{\nabla} \cdot (\boldsymbol{v} \circ \boldsymbol{\chi}) = (\boldsymbol{\nabla} \cdot \mathsf{A} \boldsymbol{v}) \circ \boldsymbol{\chi} \,, \tag{30a}$$
$$(\boldsymbol{\nabla} \cdot \mathsf{P} \boldsymbol{u}) \circ \boldsymbol{\chi} = \boldsymbol{\nabla} \cdot (\mathsf{SP} \boldsymbol{u} \circ \boldsymbol{\chi}) \,, \tag{30b}$$

where (30b) follows from (30a) with $\boldsymbol{v} = \mathsf{A}^{-1} \mathsf{P} \boldsymbol{u}$ since $\mathsf{A}^{-1} \mathsf{P} = \mathsf{SP}$ with

$$\mathsf{S} = \operatorname{diag}(1, s^{-2}, s^{-1}) \,. \tag{31}$$

For any function ϕ, we define $\bar{\phi}$ as its mean along the axis of rotation, that is,

$$\bar{\phi} \circ \boldsymbol{\chi} = s \int_{-s^{-1}}^{0} \phi \circ \boldsymbol{\chi} \, d\zeta \,. \tag{32}$$

Further, we write $\hat{\phi} = \phi - \bar{\phi}$ to denote the deviation from the mean. The definition for vector fields is analogous. Since, for arbitrary functions ϕ and ψ, $\bar{\psi} \circ \boldsymbol{\chi}$ is independent of ζ and $\det \mathsf{A} = s$, we see that

$$\int_{\mathcal{D}} \hat{\phi} \, \bar{\psi} \, d\boldsymbol{x} = s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \hat{\phi} \circ \boldsymbol{\chi} \, d\zeta \, \bar{\psi} \circ \boldsymbol{\chi} \, dx = 0 \,. \tag{33}$$

In other words, mean and fluctuating components in the sense of (32) are L^2-orthogonal.

5. Thermal wind

The formal leading order balance in the Euler–Boussinesq momentum equation (1a) gives an expression for the *thermal* or *geostrophic wind* $\boldsymbol{u}_g$,

$$\boldsymbol{\Omega} \times \boldsymbol{u}_g = -\boldsymbol{\nabla} p - \rho \, \boldsymbol{k} \,. \tag{34}$$

Taking the curl to remove the pressure, we obtain the *thermal wind relation*

$$\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{u}_g = \begin{pmatrix} -\nabla^{\perp} \rho \\ 0 \end{pmatrix} . \tag{35}$$

The characteristics of this first order equation are the lines parallel to the axis of rotation. With the notation set up in Section 4, it is easy to integrate (35) along its characteristic lines. Splitting $\boldsymbol{u}_g = \hat{\boldsymbol{u}}_g + \bar{\boldsymbol{u}}_g$, we first note that (35) determines only the mean-free component $\hat{\boldsymbol{u}}_g$. Indeed,

$$\partial_\zeta(\hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}) = (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \hat{\boldsymbol{u}}_g) \circ \boldsymbol{\chi} = \begin{pmatrix} -\nabla^\perp \rho \circ \boldsymbol{\chi} \\ 0 \end{pmatrix} . \tag{36}$$

The mean-free component $\hat{\boldsymbol{u}}_g$ must further satisfy

$$0 = s \int_{-s^{-1}}^{0} \hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi} \, d\zeta = \hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}(-s^{-1}) - s \int_{-s^{-1}}^{0} \zeta \, \partial_\zeta (\hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}) \, d\zeta \, , \tag{37}$$

where we write $\boldsymbol{\chi}'$ to abbreviate $\boldsymbol{\chi}(x, y, \zeta')$. Then,

$$\begin{aligned} \hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi} &= \hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}(-s^{-1}) + \int_{-s^{-1}}^{\zeta} \partial_\zeta (\hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}') \, d\zeta' \\ &= s \int_{-s^{-1}}^{0} \zeta \, \partial_\zeta (\hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}) \, d\zeta + \int_{-s^{-1}}^{\zeta} \partial_\zeta (\hat{\boldsymbol{u}}_g \circ \boldsymbol{\chi}') \, d\zeta' \, . \end{aligned} \tag{38}$$

Inserting the thermal wind relation (36), we find that the vertical component of the thermal wind is independent of ζ, so $\boldsymbol{k} \cdot \hat{\boldsymbol{u}}_g = 0$, and the horizontal components of the thermal wind are given by

$$\hat{u}_g = \nabla^\perp \theta \tag{39a}$$

with

$$\theta \circ \boldsymbol{\chi} = -s \int_{-s^{-1}}^{0} \zeta \, \rho \circ \boldsymbol{\chi} \, d\zeta - \int_{-s^{-1}}^{\zeta} \rho \circ \boldsymbol{\chi}' \, d\zeta' \, . \tag{39b}$$

Equation (39) shows that $\hat{u}_g$ is horizontally divergence-free. Since $\boldsymbol{k} \cdot \hat{\boldsymbol{u}}_g = 0$, we also have $\boldsymbol{\nabla} \cdot \boldsymbol{u}_g = 0$.

6. Derivation of the first-order balance model

We now implement the procedure outlined in Section 3 at first order in ε. Using the Lin constraints for the ε-flow, (23) and (24), we can write

$$\boldsymbol{u}_\varepsilon = \boldsymbol{u} + \varepsilon \, (\dot{\boldsymbol{v}} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u}) + O(\varepsilon^2) \tag{40}$$

and

$$\rho_\varepsilon = \rho - \varepsilon \, \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) + O(\varepsilon^2) \, . \tag{41}$$

The transformation vector field $\boldsymbol{v}$ will be specified in the following. By construction, we assume that all flows are volume-preserving and leave the domain invariant, so that $\boldsymbol{u}, \boldsymbol{v} \in V_{\text{div}}$. For technical reasons, we also assume that the domain-mean of $\boldsymbol{u}$ is zero.

This is not a restriction since the assumption is only removing a steady solid-body translation from the system, which is a constant of the motion of the full Euler–Boussinesq system so that it remains zero if it vanishes initially. The transformation vector field $\boldsymbol{v}$ shall also be chosen so as not to generate a solid body translation in balance model coordinates.

Inserting these relations into the Euler–Boussinesq Lagrangian (8), expanded in powers of ε and truncated to $O(\varepsilon)$, we obtain after a short computation

$$
\begin{aligned}
L &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon + \tfrac{\varepsilon}{2} \, |\dot{\boldsymbol{\eta}}_\varepsilon|^2 \, d\boldsymbol{a} - \int_{\mathcal{D}} \rho_\varepsilon \, z \, d\boldsymbol{x} \\
&= \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{x} - \rho \, z \, d\boldsymbol{x} + \varepsilon \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v} + \tfrac{1}{2} \, |\boldsymbol{u}|^2 - \rho \, \boldsymbol{v} \cdot \boldsymbol{k} \, d\boldsymbol{x} \\
&= L_0 + \varepsilon \, L_1 \, .
\end{aligned} \tag{42}
$$

Here we have used the divergence theorem to rewrite the potential energy contribution to L_1, where the boundary integral vanishes due to the boundary condition $\boldsymbol{k} \cdot \boldsymbol{v} = 0$.

We must now choose the transformation vector field $\boldsymbol{v}$ such that it removes, to the extent possible, all terms that are quadratic in components of $\boldsymbol{u}$. In preparation, we decompose the kinetic energy part of L_1 as

$$
\int_{\mathcal{D}} |\boldsymbol{u}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} |\hat{\boldsymbol{u}}|^2 + 2 \, \hat{\boldsymbol{u}} \cdot \bar{\boldsymbol{u}} + |\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \, . \tag{43}
$$

The cross term in (43) vanishes due to (33). The square terms are decomposed into terms that can be written as an L^2-pairing with $\mathsf{P}\hat{\boldsymbol{u}}$, and a final remainder term which cannot. Starting with the contribution from $|\hat{\boldsymbol{u}}|^2$, we write

$$
\int_{\mathcal{D}} |\hat{\boldsymbol{u}}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}} \cdot \mathsf{Q}\hat{\boldsymbol{u}} \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}(\hat{\boldsymbol{u}} + \boldsymbol{C}[\hat{\boldsymbol{u}}]) \, d\boldsymbol{x} \, , \tag{44}
$$

using Lemma 7 in the final equality. The contribution from $|\bar{\boldsymbol{u}}|^2$ is split differently. Setting $\mathsf{S}_0 = \operatorname{diag}(1, s^{-2}, 0)$, noting that $s \, (\mathsf{I} - \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}} = \boldsymbol{\Omega} \, \bar{u}_3$, and noting that $(\mathsf{S}_0\mathsf{P})^\mathsf{T} - \mathsf{S}_0\mathsf{P}$ is skew so that $\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} \cdot \bar{\boldsymbol{u}} - \bar{\boldsymbol{u}} \cdot \mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} = 0$, we write

$$
\begin{aligned}
\int_{\mathcal{D}} |\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} &= \int_{\mathcal{D}} (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}} \cdot (\mathsf{I} - \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}} + \mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} \cdot \mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} \, d\boldsymbol{x} \\
&= -s^{-1} \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\boldsymbol{\nabla}(\boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}}) \, z \, d\boldsymbol{x} + \int_{\mathcal{D}} |\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \, .
\end{aligned} \tag{45}
$$

The second equality is based on Lemma 8 with $\bar{\phi} = s^{-1} \, \boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}}$. Note that Lemma 8 could be used in different ways so long as we retain a pairing of some function $\bar{\phi}$ with $\bar{u}_3$. Our chosen splitting is distinguished, due to Lemma 3, by the fact that the remainder integral in (45) is proportional to the kinetic energy of a *divergence-free* two-dimensional vector field. In the following, this remainder term will be the only term that yields an evolution equation in the variational principle. As $\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}$ is divergence-free, this evolution equation can always be written in terms of a scalar stream function, that is, is determined by the evolution of a single scalar field. Any other splitting would yield a two-component evolution equation except for one special tilt of the axis of rotation, which is not desirable.

Collecting terms, we find

$$\int_{\mathcal{D}} |\boldsymbol{u}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\hat{\boldsymbol{V}} \, d\boldsymbol{x} + \int_{\mathcal{D}} |\mathsf{S}_0 \mathsf{P} \bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \tag{46}$$

with

$$\boldsymbol{V} = \hat{\boldsymbol{u}} + \boldsymbol{C}[\hat{\boldsymbol{u}}] - s^{-1} \, \boldsymbol{\nabla}\big(\boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}}\big) \, z \,. \tag{47}$$

Even though $\boldsymbol{V}$ is not necessarily mean-free, only its mean-free component $\hat{\boldsymbol{V}}$ contributes to the Lagrangian due to the pairing with $\hat{\boldsymbol{u}}$.

Since J has a one-dimensional kernel, it is impossible to remove all quadratic terms from the L_1 Lagrangian, but choosing the transformation vector field $\boldsymbol{v}$ as a solution to the equation

$$\mathsf{J}\boldsymbol{v} = -\frac{1}{2}\mathsf{P}\hat{\boldsymbol{V}} \,, \tag{48}$$

up to terms that only depend on ρ, we can make L_1 affine in $\hat{\boldsymbol{u}}$: Noting that J is invertible on Range P with pseudo-inverse J^T, and further that $\mathsf{P}\mathsf{J}^\mathsf{T} = \mathsf{J}^\mathsf{T} = \mathsf{J}^\mathsf{T}\mathsf{P}$, we seek $\boldsymbol{v}$ in the form

$$\boldsymbol{v} = \mathsf{P}\hat{\boldsymbol{v}} + \mathsf{Q}\hat{\boldsymbol{v}} + \bar{\boldsymbol{v}} \tag{49}$$

with

$$\mathsf{P}\hat{\boldsymbol{v}} = -\frac{1}{2}\mathsf{J}^\mathsf{T} \, \hat{\boldsymbol{V}} + \lambda \, \mathsf{J}^\mathsf{T} \hat{\boldsymbol{V}}_g \,. \tag{50}$$

In this expression, λ is a free parameter and

$$\boldsymbol{V}_g = \hat{\boldsymbol{u}}_g + \boldsymbol{C}[\hat{\boldsymbol{u}}_g] \tag{51}$$

which takes the same form as $\boldsymbol{V}$ with $\boldsymbol{u}$ replaced by $\boldsymbol{u}_g$. Since $\boldsymbol{u}_g$ is not constrained by the thermal wind relation, we do not include a term that matches the third term of (47) into the ansatz for $\boldsymbol{V}_g$.

The terms $\mathsf{Q}\hat{\boldsymbol{v}}$ and $\bar{\boldsymbol{v}}$ in (49) are chosen such that $\boldsymbol{v}$ becomes divergence free and tangent to the top and bottom boundaries. Using the construction from Lemma 5, we write $\mathsf{Q}\hat{\boldsymbol{v}} \equiv \boldsymbol{\Omega} \, \hat{g}$, where

$$\hat{g} \circ \boldsymbol{\chi} = \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla} \cdot (\mathsf{S}\mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}') \, d\zeta' \,. \tag{52}$$

By Lemma 6, $\hat{\boldsymbol{v}}$ is divergence-free. By Lemma 9, $\bar{\boldsymbol{v}}$ can be chosen such that the zero flux boundary condition

$$\begin{aligned} \bar{v}_3 \circ \boldsymbol{\chi} &= -\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}(-s^{-1}) - s \, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) \\ &= -\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}(0) - s \, \hat{g} \circ \boldsymbol{\chi}(0) \,. \end{aligned} \tag{53}$$

is satisfied and, moreover, $\bar{\boldsymbol{v}}$ is divergence free. Lemma 10 shows that the choice of the horizontal components $\bar{v}$ will not enter the computation of the equations of motion.

Only the contribution from $\bar{v}_3$ will appear, but it is directly a function of $\hat{v}_3$ by (53), so that the final expression will not contain any references to $\bar{\boldsymbol{v}}$.

Combining the contributions from rotation and kinetic energy, we find

$$\begin{aligned}\int_{\mathcal{D}} \boldsymbol{u}\cdot\mathsf{J}\boldsymbol{v} + \tfrac{1}{2}\,|\boldsymbol{u}|^2\,d\boldsymbol{x} &= \int_{\mathcal{D}} \hat{\boldsymbol{u}}\cdot\mathsf{J}\mathsf{J}^{\mathsf{T}}\left(-\tfrac{1}{2}\,\hat{\boldsymbol{V}} + \lambda\hat{\boldsymbol{V}}_g\right) + \tfrac{1}{2}\,\hat{\boldsymbol{u}}\cdot\mathsf{P}\hat{\boldsymbol{V}} + \tfrac{1}{2}\,|\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \lambda\,\hat{\boldsymbol{u}}\cdot\mathsf{P}\hat{\boldsymbol{V}}_g + \tfrac{1}{2}\,|\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \lambda\,\hat{u}\cdot\hat{u}_g + \tfrac{1}{2}\,|\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2\,d\boldsymbol{x} \end{aligned} \tag{54}$$

where, in the last equality, we have used Lemma 7 and the fact that the geostrophic velocity is exclusively horizontal. The contribution to the L_1-Lagrangian from the potential energy term is

$$-\int_{\mathcal{D}} \rho\,\boldsymbol{v}\cdot\boldsymbol{k}\,d\boldsymbol{x} = \int_{\mathcal{D}} \hat{u}_g\cdot\left(\tfrac{1}{2}\,\hat{u} - \lambda\,\hat{u}_g\right)d\boldsymbol{x}\,; \tag{55}$$

the details of this calculation are given in Appendix D. Altogether, the L_1-Lagrangian then reads

$$\begin{aligned} L_1 &= \int_{\mathcal{D}} \lambda\,\hat{u}\cdot\hat{u}_g + \tfrac{1}{2}\,|\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2 + \hat{u}_g\cdot(\tfrac{1}{2}\,\hat{u} - \lambda\,\hat{u}_g)\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \nu\,\hat{u}\cdot\hat{u}_g + \tfrac{1}{2}\,\bar{\boldsymbol{u}}\cdot\mathsf{M}\bar{\boldsymbol{u}} - \lambda\,|\hat{u}_g|^2\,d\boldsymbol{x}\,, \end{aligned} \tag{56}$$

where $\nu = \lambda + \frac{1}{2}$ and

$$\mathsf{M} = \mathsf{P}\mathsf{S}_0^2\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c/s \\ 0 & -c/s & c^2/s^2 \end{pmatrix}. \tag{57}$$

In the following, we choose λ such that $\nu > 0$.

7. Derivation of the balance model equations of motion

Taking the variation of (56), we obtain

$$\begin{aligned} \delta L_1 &= \int_{\mathcal{D}} \delta\boldsymbol{u}\cdot(\nu\,\hat{\boldsymbol{u}}_g + \mathsf{M}\bar{\boldsymbol{u}}) + \nu\,u\cdot\delta\hat{u}_g - 2\lambda\,\delta\hat{u}_g\cdot\hat{u}_g\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \delta\boldsymbol{u}\cdot\boldsymbol{p} + \delta\hat{u}_g\cdot\hat{b}\,d\boldsymbol{x}\,, \end{aligned} \tag{58}$$

where

$$\boldsymbol{p} = \mathsf{M}\bar{\boldsymbol{u}} + \nu\,\hat{\boldsymbol{u}}_g \qquad \text{and} \qquad \hat{b} = \nu\,\hat{u} - 2\lambda\,\hat{u}_g\,, \tag{59}$$

To rewrite the second term on the right of (58), we insert the expression for the geostrophic velocity (39), change variables, and recall that horizontal derivatives and composition with $\boldsymbol{\chi}$ commute, so that

$$\begin{aligned}\int_{\mathcal{D}} \delta\hat{u}_g \cdot \hat{b}\, d\boldsymbol{x} &= -s \int_{\mathcal{D}} \int_{-s^{-1}}^{\zeta} \nabla^{\perp}\delta\rho \circ \boldsymbol{\chi}'\, d\zeta' \cdot \hat{b} \circ \boldsymbol{\chi}\, d\boldsymbol{\xi} \\ &= s \int_{\mathcal{D}} \nabla^{\perp}\delta\rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \hat{b} \circ \boldsymbol{\chi}'\, d\zeta'\, d\boldsymbol{\xi} \\ &= -s \int_{\mathcal{D}} \delta\rho \circ \boldsymbol{\chi}\, \nabla^{\perp} \cdot \int_{-s^{-1}}^{\zeta} \hat{b} \circ \boldsymbol{\chi}'\, d\zeta'\, d\boldsymbol{\xi} \\ &= - \int_{\mathcal{D}} \delta\rho\, \nabla^{\perp} \cdot B\, d\boldsymbol{x}\,, \end{aligned} \tag{60}$$

where B is the anti-derivative of b along the axis of rotation, i.e.,

$$B = \nu\, U - 2\lambda\, U_g \tag{61}$$

where

$$U \circ \boldsymbol{\chi} = \int_{-s^{-1}}^{\zeta} \hat{u} \circ \boldsymbol{\chi}'\, d\zeta' \tag{62}$$

and U_g is defined likewise. Inserting (60) into (58), we have

$$\delta L_1 = \int_{\mathcal{D}} \delta\boldsymbol{u} \cdot \boldsymbol{p} - \delta\rho\, \nabla^{\perp} \cdot B\, d\boldsymbol{x}\,. \tag{63}$$

We recall the general Euler–Poincaré equation (15), which can be written

$$\partial_t \boldsymbol{m} + (\boldsymbol{\nabla} \times \boldsymbol{m}) \times \boldsymbol{u} + \frac{\delta\ell}{\delta\rho}\, \boldsymbol{\nabla}\rho = \boldsymbol{\nabla}\tilde{\phi} \tag{64}$$

with $\tilde{\phi}$ an arbitrary scalar field. Here,

$$\boldsymbol{m} = \frac{\delta\ell}{\delta\boldsymbol{u}} = \boldsymbol{R} + \varepsilon\, \boldsymbol{p} \qquad \text{and} \qquad \frac{\delta\ell}{\delta\rho} = -z - \varepsilon\, \nabla^{\perp} \cdot B\,, \tag{65}$$

so that, with $\tilde{\phi} = -\phi - z\rho$,

$$\boldsymbol{\Omega} \times \boldsymbol{u} + \rho\boldsymbol{k} + \varepsilon\,(\partial_t\boldsymbol{p} + (\boldsymbol{\nabla} \times \boldsymbol{p}) \times \boldsymbol{u} - \boldsymbol{\nabla}\rho\, \nabla^{\perp} \cdot B) = -\boldsymbol{\nabla}\phi\,. \tag{66}$$

As $c\,\partial_z = -s\,\partial_y$ when applied to averaged quantities, we find that $\boldsymbol{\nabla} \times \mathsf{M}\bar{\boldsymbol{u}} = \boldsymbol{\Omega}\,\omega$ with $\omega = s^{-1}\, \nabla^{\perp} \cdot \mathsf{M}_h\bar{\boldsymbol{u}}$. Hence,

$$\boldsymbol{\xi} \equiv \boldsymbol{\nabla} \times \boldsymbol{p} = \boldsymbol{\Omega}\,\omega + \nu\,\boldsymbol{\gamma} \tag{67}$$

with

$$\boldsymbol{\gamma} = \boldsymbol{\nabla} \times \hat{\boldsymbol{u}}_g = -\boldsymbol{\nabla} \times \operatorname{curl}\theta = \begin{pmatrix} -\nabla\partial_z\theta \\ \Delta\theta \end{pmatrix}, \tag{68}$$

where we recall that $\hat{u}_g = \nabla^\perp \theta$, write $\operatorname{curl} f = \boldsymbol{\nabla} \times (0,0,f)$ to identify the curl of a scalar field with the curl of a vector field oriented in the vertical, and use Δ to denote the *horizontal* Laplacian. With this notation in place, taking the curl of (66) and noting that $\boldsymbol{\nabla} \times (\boldsymbol{\xi} \times \boldsymbol{u}) = \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} \boldsymbol{u}$ as both $\boldsymbol{u}$ and $\boldsymbol{\xi}$ are divergence-free, we find

$$-\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \times (\rho \boldsymbol{k}) + \varepsilon \left(\partial_t \boldsymbol{\xi} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \nabla^\perp \cdot B \right) = 0 \,. \tag{69}$$

The corresponding balance model Hamiltonian, via (17), is given by

$$H = \int_{\mathcal{D}} \rho \, z + \varepsilon \left(\tfrac{1}{2} \, \bar{\boldsymbol{u}} \cdot \mathsf{M} \bar{\boldsymbol{u}} + \lambda \, |\hat{\boldsymbol{u}}_g|^2 \right) d\boldsymbol{x} \,. \tag{70}$$

The potential vorticity for the balance model reads

$$\begin{aligned} q &= (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla} \rho \\ &= (\boldsymbol{\Omega} + \varepsilon \, \boldsymbol{\Omega} \, \omega + \varepsilon \nu \, \gamma) \cdot \boldsymbol{\nabla} \rho \\ &= (s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) \, \partial_z \rho + (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) \cdot \nabla \rho \\ &= -(s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) \, \partial_z (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) - (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) \cdot \nabla (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) \,, \end{aligned} \tag{71}$$

where $\Omega = (0, c)$. In the last equality, we have used that $\rho = -\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta$. Alternatively, we can write

$$q = \begin{vmatrix} -(s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) & \nabla (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) \\ (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) & \partial_z (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) \end{vmatrix} \,. \tag{72}$$

We remark that the relation between q and θ can be seen as a second order nonlinear differential operator of the form

$$q = F(\omega, D^2 \theta) \,, \tag{73}$$

which is nonlinearly elliptic (e.g. Gilbarg & Trudinger, 2001) so long as

$$[F_{ij}] = \frac{\partial F}{\partial (D^2 \theta)} \tag{74}$$

is positive (or negative) definite. Direct computation shows that

$$[F_{ij}] = \begin{pmatrix} \varepsilon \nu \partial_z \rho & -\frac{1}{2} \varepsilon c \, \xi_1 & -\frac{1}{2} \varepsilon \, (\xi_1 + \nu \partial_x \rho) \\ -\frac{1}{2} \varepsilon c \, \xi_1 & \varepsilon \nu \partial_z \rho - c(c + \varepsilon \xi_2) & -\frac{1}{2} \big(2cs + \varepsilon (c \xi_3 + s \xi_2 + \nu \partial_y \rho) \big) \\ -\frac{1}{2} \varepsilon \, (\xi_1 + \nu \partial_x \rho) & -\frac{1}{2} \big(2cs + \varepsilon (c \xi_3 + s \xi_2 + \nu \partial_y \rho) \big) & -s^2 - \varepsilon s \, \xi_3 \end{pmatrix} , \tag{75}$$

so that $-F$ is positive definite provided its principal minors are positive, that is,

$$-\varepsilon \nu \partial_z \rho > 0 \,, \tag{76a}$$

$$-\varepsilon \nu \, \partial_z \rho + c \big[c + \varepsilon \big(\xi_2 + \tfrac{1}{4} \tfrac{c \xi_1^2}{\nu \partial_z \rho} \big) \big] > 0 \,, \tag{76b}$$

$$\det F < 0 \,. \tag{76c}$$

Condition (76c) can be written more explicitly as

$$-(\varepsilon s\nu\partial_z\rho)^2 + \varepsilon^2 f(\boldsymbol{\xi}, \nabla\rho; \boldsymbol{\xi}, \rho) < 0\,, \tag{77}$$

where f is linear in its first two arguments. Hence, these conditions are satisfied if the fluid is stably stratified so that $\partial_z\rho < 0$, the deviations from a steady mean state are sufficiently small, and ε is sufficiently small.

8. Separation of balance relation into dynamic and kinematic components

In the following, we show that the balance relation is mostly a kinematic relationship between density ρ and balanced velocity field. However, there is one horizontal scalar field, ω, which evolves dynamically via the vertical component equation

$$\partial_t\xi_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}\xi_3 = \boldsymbol{\xi}\cdot\boldsymbol{\nabla}u_3 + \nabla\rho\cdot\nabla^\perp\nabla^\perp\cdot B + \varepsilon^{-1}\,\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\hat{u}_3\,. \tag{78}$$

Note that the seemingly unbalanced $O(\varepsilon^{-1})$-term in this equation is actually only an $O(1)$-contribution because $\boldsymbol{u}$ is an $O(\varepsilon)$ perturbation of $\boldsymbol{u}_g$ whose vertical component is zero.

Taking the average of (78) along the axis of rotation and using Lemma 1 and Lemma 2 from the appendix to commute averaging with directional derivatives where possible, we find that the evolution equation for ξ_3 reduces to a prognostic equation for ω,

$$\partial_t\omega + \bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega = \overline{\nabla\rho\cdot\nabla^\perp\nabla^\perp\cdot B} - \nu\,\overline{(\partial_z\nabla\theta\cdot\nabla u_3 - \Delta\theta\,\partial_z u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}\Delta\theta)}\,. \tag{79}$$

All other contributions to (69) are entirely kinematic. To see this, we start with multiplying (79) by $\boldsymbol{\Omega}$, then subtract this expression from (69) to obtain

$$\begin{aligned}0 = &-\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\times(\rho\boldsymbol{k}) + \varepsilon\,\big[\nu\,\partial_t\boldsymbol{\gamma} + \boldsymbol{\Omega}\,\hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega + \nu\,\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{\gamma} - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}\boldsymbol{u} \\ &+ \boldsymbol{\nabla}\rho\times\boldsymbol{\nabla}\nabla^\perp\cdot B + \boldsymbol{\Omega}\,\overline{\nabla\rho\cdot\nabla^\perp\nabla^\perp\cdot B} + \nu\,\boldsymbol{\Omega}\,\overline{(\gamma\cdot\boldsymbol{\nabla}u_3 - \boldsymbol{u}\cdot\boldsymbol{\nabla}\Delta\theta)}\big]\,.\end{aligned} \tag{80}$$

We now assume that the flow is strongly and uniformly stratified. More specifically, we shall assume that

$$\partial_z\rho = -\alpha + \partial_z\tilde{\rho} \tag{81}$$

where $\tilde{\rho}$ is small in a suitable norm, to be specified further below. It is possible to weaken this assumption and allow variations in the stratification profile so long as stratification is uniformly stable across the layer, but the simpler assumption (81) makes the structure of the problem more transparent.

Inserting (81) into the definition of $\boldsymbol{\gamma}$ and noting that $\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\theta = -\rho$, we find

$$\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\dot{\gamma}_3 = \Delta(\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\dot{\theta}) = \Delta(\boldsymbol{u}\cdot\boldsymbol{\nabla}\rho) = -\alpha\,\Delta u_3 + \Delta(\boldsymbol{u}\cdot\boldsymbol{\nabla}\tilde{\rho})\,. \tag{82}$$

Decomposing

$$U = \nabla^\perp\Psi + \nabla\Phi\,, \tag{83}$$

and, setting $\zeta = \Delta\Psi$, we have

$$\boldsymbol{\nabla}\rho \times \boldsymbol{\nabla}\nabla^\perp \cdot B = -\alpha\nu\, \boldsymbol{k} \times \boldsymbol{\nabla}\zeta + \nu\, \boldsymbol{\nabla}\tilde\rho \times \boldsymbol{\nabla}\zeta - 2\lambda\, \boldsymbol{\nabla}\rho \times \boldsymbol{\nabla}\nabla^\perp \cdot U_g \,. \tag{84}$$

Inserting (82) and (84) into (80), taking the horizontal curl of the horizontal component equations, and applying $\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}$ to the vertical component equation, we obtain a pair of kinematic balance relations:

$$(\boldsymbol{\Omega}\cdot\boldsymbol{\nabla})^2\zeta + \varepsilon\alpha\nu\,\Delta\zeta = -\Delta\tilde\rho + \varepsilon\, f\,, \tag{85a}$$

$$(\boldsymbol{\Omega}\cdot\boldsymbol{\nabla})^2 u_3 + \varepsilon\alpha\nu\,\Delta u_3 = \varepsilon\, g \tag{85b}$$

with

$$\begin{aligned} f &= \Omega\cdot\nabla^\perp\,(\hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega) + \nu\,\nabla^\perp\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\gamma) - \nabla^\perp\cdot(\boldsymbol{\xi}\cdot\boldsymbol{\nabla}u) - \nu\,\nabla\cdot(\nabla\tilde\rho\,\partial_z\zeta) \\ &\quad + \nu\,\nabla\cdot(\nabla\zeta\,\partial_z\tilde\rho) + 2\lambda\,\nabla\cdot(\nabla\rho\,\partial_z\nabla^\perp\cdot U_g) - 2\lambda\,\nabla\cdot(\nabla\nabla^\perp\cdot U_g\,\partial_z\rho) \\ &\quad - 2\lambda\,\Omega\cdot\nabla^\perp\,\overline{\nabla\tilde\rho\cdot\nabla^\perp\nabla^\perp\cdot U_g} + \nu\,\Omega\cdot\nabla^\perp\,\overline{(\nabla\tilde\rho\cdot\nabla^\perp\zeta + \boldsymbol{\gamma}\cdot\boldsymbol{\nabla}u_3 - \hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\Delta\theta)}\,, \end{aligned} \tag{85c}$$

$$g = \nu\,\Delta(\boldsymbol{u}\cdot\boldsymbol{\nabla}\tilde\rho) + s\,\hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega + \nu\,\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}(\boldsymbol{u}\cdot\boldsymbol{\nabla}\gamma_3 - \boldsymbol{\xi}\cdot\boldsymbol{\nabla}u_3 - \nabla\tilde\rho\cdot\nabla^\perp\nabla^\perp\cdot B)\,. \tag{85d}$$

This elliptic problem is augmented by homogeneous Dirichlet boundary conditions on both ζ and u_3. These boundary conditions encode, for u_3, the no-flux conditions of the full velocity field, and for ζ that it is the antiderivative along $\boldsymbol{\Omega}$ of a mean-free field with an arbitrary choice of gauge that disappears upon differentiation.

The right hand functions f and g still contain terms that depend on the unknown functions ζ and u_3, so they need to be solved by fixed point iteration. In the next section, we will argue that this can be done under suitable smallness assumptions for $\tilde\rho$ and ω.

9. Closing the balance model

To close the balance relation and the prognostic equations, we need to recover the full velocity field $\boldsymbol{u}$ from ω, ζ, and u_3. We express the horizontal component of the velocity in terms of stream function ψ and velocity potential ϕ,

$$u = \nabla^\perp\psi + \nabla\phi\,. \tag{86}$$

First, note that $\nabla^\perp\cdot\mathsf{M}_h\bar{\boldsymbol{u}} = \nabla^\perp\cdot\bar u - c/s\,\partial_x\bar u_3 = s\,\omega$ by definition, so that

$$\Delta\bar\psi = s\,\omega + \tfrac{c}{s}\,\partial_x\bar u_3\,. \tag{87a}$$

Next, due to (83),

$$\Delta\Psi = \zeta\,. \tag{87b}$$

Finally, by incompressibility,

$$\Delta\phi = -\partial_z u_3 \tag{87c}$$

so that, altogether,

$$u = \nabla^\perp \bar{\psi} + \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \nabla^\perp \Psi + \nabla \phi \,. \tag{87d}$$

This expression for the horizontal components of the velocity field, first, proves that the kinematic balance relation (85) can be solved by iteration. Seeking a solution in the Sobolev space H^{s+2}, H^s denoting the space of square integrable functions with square-integrable derivatives up to order s where we take s large enough so that products of such functions are also contained in H^s, we need to verify that all terms that appear in f and g are contained in H^s. For example, for the term $\nabla^\perp \cdot (\boldsymbol{\xi} \cdot \boldsymbol{\nabla} u)$ which appears in the expression for f, the highest derivatives on the unknown functions appear as

$$\boldsymbol{\xi} \cdot \boldsymbol{\nabla}(\partial_x \bar{u}_3 + \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \zeta) \,, \tag{88}$$

so this term is contained in H^s provided $\zeta, u_3 \in H^{s+2}$. All other terms are either similar or have only lower-order derivatives on the unknowns. Further, the coefficients that appear, here $\boldsymbol{\xi}$, depend only on ω and $\tilde{\rho}$, so they are small if the data is close enough to the stably stratified rest state in a Sobolev norm of sufficiently high order. Thus, the balance relation (85) defines a contraction in H^{s+2} and can be solved by iteration.

Second, the reconstructed velocity field $\boldsymbol{u}$ is used to propagate ω and ρ (or, alternatively, the potential vorticity q), in time. Then the complete set of balance model equations is given by the the kinematic balance relation (85), the reconstruction equations (87), the continuity equation in three dimensions,

$$\partial_t \rho + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho = 0 \,, \tag{89a}$$

and the evolution equation in two spatial dimensions for ω,

$$\partial_t \omega + \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega = \overline{\nabla \rho \cdot \nabla^\perp \nabla^\perp \cdot B} - \nu \, \overline{(\partial_z \nabla \theta \cdot \nabla u_3 - \Delta \theta \, \partial_z u_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} \Delta \theta)} \,. \tag{89b}$$

where B is defined in (61) and θ is the geostrophic stream function which depends on ρ via (39b).

10. Discussion and Conclusion

In this paper, we have achieved a variational model reduction for the full three-dimensional Euler–Boussinesq equation with a full Coriolis force. We have studied a simple setting, namely the f-plane approximation, a layer of fluid of constant depth, and periodic boundary conditions in the horizontal. The picture which emerges is structurally identical to that of variational balance models for the primitive equations as derived by Salmon (1996) and generalized in Oliver and Vasylkevych (2016):

(i) There are two prognostic variables of the first-order balance model, the density ρ (equivalently, a potential temperature) and a scalar generalized vorticity ω.
(ii) The generalized vorticity ω depends only on the two velocity components that are perpendicular to the axis of rotation, and it is averaged along the axis of rotation. Thus, ω is independent of the oblique vertical coordinate ζ.
(iii) All other components of the velocity field, that is, all deviations from the vertical mean as well as the vertical mean of the velocity component pointing along the

axis of rotation, are kinematic. In other words, these components are slaved to ρ and ω via a balance relation.

(iv) The balance relation is elliptic if rotation is sufficiently fast and the prognostic fields are small perturbations of a stably stratified equilibrium state.

(v) The balance model conserves energy and has a materially conserved potential vorticity.

Thus, we have verified that the variational derivation of balance models of semi-geostrophic type extends all the way to one of the most general models of geophysical flows. In particular, the assumption of hydrostaticity and of the "traditional approximation" changes details, but does not change the structural features of the semi-geostrophic limit.

At the same time, a full Coriolis force causes difficulties that are not seen in the simpler cases. Since the axis of rotation is not aligned with the direction of gravitational force, there are two distinguished "vertical" directions. This is an obstacle to using a fully intrinsic geometric formulation of the derivation in the spirit of Arnold and Khesin (1999) or Gilbert and Vanneste (2018), forcing us to resort to detailed coordinate calculations. The resulting equations, therefore, appear to lack the relative simplicity of balance models in more idealized settings; many of the new terms simplify or disappear when the Coriolis force is acting exactly in the horizontal plane.

We emphasize that our derivation requires a nontrivial change of coordinates already at $O(\varepsilon)$. The associated transformation vector field depends on the prognostic part of the mean velocity, cf. the last term in (47). We believe that it is not possible to remove this contribution to the transformation at $O(\varepsilon)$ since there is no leading-order constraint through the thermal wind relation on this component. For the analogous computation for the primitive equations, it suffices to set $\lambda = \frac{1}{2}$ in (50), which formally cancels all terms at $O(\varepsilon)$. For the Euler–Boussinesq system, it is the last term in (47) that cannot be canceled. This additional contribution to the transformation vector field appears even in the case when the axis of rotation is aligned with the geometric vertical and only disappears when the hydrostatic approximation is made. Thus, the more straightforward derivation by Salmon (1996), who inserts the thermal wind relation directly into the extended Lagrangian to constrain the variational principle does not work in this case; the more general setting described in Section 3 must be used.

When the axis of rotation is aligned with the horizontal, i.e., for the equatorial f-plane, all our final expressions remain non-singular, which is surprising given that some of the intermediate expressions do contain diverging terms. What fails, however, is ellipticity of the balance relation. Stratification is providing regularization in the horizontal plane, rotation is providing regularization in the vertical. When the axis of rotation is tilted into the horizontal, all control in the geometric vertical is lost. We do not expect that the model remains well posed in this limit.

The balance relation allows, under the conditions stated, stable reconstruction of the slaved components of the velocity field from sufficiently smooth prognostic variables. A full analysis of well-posedness of the balance model is more difficult, as the right hand side of the balance relation contains high-order derivatives of the prognostic variables, and remains open. In practical terms, the balance relation might be most useful as a diagnostic relation independent of the full system of prognostic relations. Nonetheless, the full balance model can be solved numerically in the formulation given in Section 9. For balance models in two dimensions, potential vorticity based methods provide an alternative, numerically robust setting (Dritschel et al., 2017). As for primitive equation balance models, potential vorticity based numerics require the solution of a nonlinear elliptic equation (cf. Akramov & Oliver, 2020; Oliver & Vasylkevych, 2016). This alternative formulation would require the solution of one more Monge–Ampère-like equation, here with oblique derivatives, but may be more stable because the potential vorticity is materially advected.

Acknowledgment

This paper is a contribution to project M2 (Systematic Multi-Scale Analysis and Modeling for Geophysical Flow) of the Collaborative Research Center TRR 181 "Energy Transfers in Atmosphere and Ocean" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project number 274762653. Additional funding was received through the Ideen- und Risikofund 2020 at Universität Hamburg.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Appendix A. Averaging along the axis of rotation

This appendix states two simple lemmas on the properties of the oblique averaging operation. The first shows that it distributes over products as expected.

Lemma 1. *Let ϕ and ψ be arbitrary functions and $\boldsymbol{v}$ an arbitrary vector field on $\mathcal{D}$. Then*

(i) $\overline{\psi\,\bar\phi} = \bar\psi\,\bar\phi$,
(ii) $\overline{\boldsymbol{v}\cdot\boldsymbol{\nabla}\bar\phi} = \bar{\boldsymbol{v}}\cdot\boldsymbol{\nabla}\bar\phi$.

Proof. For (i), note that $\bar\phi\circ\boldsymbol{\chi}$ is independent of ζ, so that

$$\overline{\psi\,\bar\phi}\circ\boldsymbol{\chi} = s\int_{-s^{-1}}^{0} \psi\circ\boldsymbol{\chi}\,\bar\phi\circ\boldsymbol{\chi}\,d\zeta = s\int_{-s^{-1}}^{0} \psi\circ\boldsymbol{\chi}\,d\zeta\,\bar\phi\circ\boldsymbol{\chi} = \bar\psi\circ\boldsymbol{\chi}\,\bar\phi\circ\boldsymbol{\chi}\,.$$

For (ii), note that $\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\bar\phi = 0$, so that the z-derivative can be replaced by an equivalent y-derivative. Since horizontal derivatives commute with taking the average, part (i) applies and yields the claim. □

Commutation of vertical derivatives of arbitrary functions with averaging is more subtle, as the next lemma shows. Here and in the following, we write $\boldsymbol{\chi}(0)$ and $\boldsymbol{\chi}(-s^{-1})$ to indicate that the expression is evaluated at the top ($\zeta = 0$) or at the bottom boundary ($\zeta = s^{-1}$), as a function of the remaining horizontal variables.

Lemma 2. *Let ϕ be a function with $\phi\circ\boldsymbol{\chi}(0) = \phi\circ\boldsymbol{\chi}(-s^{-1})$ and $\boldsymbol{v}$ an arbitrary vector field on $\mathcal{D}$. Then*

(i) $\partial_z\bar\phi = \overline{\partial_z\phi}$,
(ii) $\overline{\bar{\boldsymbol{v}}\cdot\boldsymbol{\nabla}\phi} = \bar{\boldsymbol{v}}\cdot\boldsymbol{\nabla}\bar\phi$.

Proof. On the one hand, $\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \bar{\phi} = 0$, so that $s \, \partial_z \bar{\phi} = -c \, \partial_y \bar{\phi}$. On the other hand,

$$\begin{aligned} \overline{\partial_z \phi} \circ \boldsymbol{\chi} &= s \int_{-s^{-1}}^{0} (\partial_z \phi) \circ \boldsymbol{\chi} \, d\zeta \\ &= \int_{-s^{-1}}^{0} (\partial_\zeta - c \, \partial_y)(\phi \circ \boldsymbol{\chi}) \, d\zeta \\ &= -\frac{c}{s} \, \partial_y \bar{\phi} \circ \boldsymbol{\chi} + \phi \circ \boldsymbol{\chi}(0) - \phi \circ \boldsymbol{\chi}(-s^{-1}) \, . \end{aligned} \tag{A.1}$$

Under the condition stated, the boundary terms cancel. This implies (i). For (ii), we use Lemma 1 to move $\bar{\boldsymbol{v}}$ out of the average. Horizontal derivatives can be moved out of the average without restrictions; for the z-derivative, part (i) applies. □

Appendix B. Splitting of divergence free vector fields

We proceed to prove a number of identities which describe the splitting of divergence-free vector fields with zero-flux boundary conditions into, on the one hand, mean and mean-free components and, on the other hand, components along and perpendicular to the axis of rotation.

Lemma 3. *Let $\boldsymbol{v} \in V_{\mathrm{div}}$. Then $\nabla \cdot \mathsf{S}_h \mathsf{P} \bar{\boldsymbol{v}} = 0$.*

Proof. For a divergence-free vector field, $\boldsymbol{\nabla} \cdot \mathsf{P} \boldsymbol{v} = -\boldsymbol{\nabla} \cdot \mathsf{Q} \boldsymbol{v} = -\boldsymbol{\Omega} \cdot \boldsymbol{\nabla}(\boldsymbol{\Omega} \cdot \boldsymbol{v})$, so that (30b) turns into

$$\nabla \cdot (\mathsf{S}_h \mathsf{P} \boldsymbol{v} \circ \boldsymbol{\chi}) = -\partial_\zeta (\boldsymbol{\Omega} \cdot \boldsymbol{v} \circ \boldsymbol{\chi}) - \partial_\zeta (\mathsf{S}_3 \mathsf{P} \boldsymbol{v} \circ \boldsymbol{\chi}) = -s^{-1} \, \partial_\zeta (v_3 \circ \boldsymbol{\chi}) \, , \tag{B.1}$$

where the last equality can be verified by direct computation in coordinates. Integrating in ζ and noting that the right hand side is zero due to the boundary conditions, we obtain the statement of the lemma. □

The following is a somewhat weaker converse of Lemma 3.

Lemma 4. *Let $\boldsymbol{v}$ be a vector field with $\nabla \cdot \mathsf{S}_h \mathsf{P} \bar{\boldsymbol{v}} = 0$. Then $\boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$.*

Proof. The assumption implies

$$0 = \nabla \cdot \bar{v} - \frac{c}{s} \, \partial_y \bar{v}_3 \, . \tag{B.2}$$

Since $\bar{v}_3 \circ \boldsymbol{\chi}$ is independent of ζ, this implies $\boldsymbol{\nabla} \cdot (\mathsf{A}^{-1} \bar{\boldsymbol{v}} \circ \boldsymbol{\chi}) = 0$. By (30a), this implies that $\boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$. □

Corollary 1. *If $\boldsymbol{v} \in V_{\mathrm{div}}$, then $\boldsymbol{\nabla} \cdot \hat{\boldsymbol{v}} = \boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$.*

Proof. Lemma 3 followed by Lemma 4 yields $\boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$. Then $\hat{\boldsymbol{u}} = \boldsymbol{u} - \bar{\boldsymbol{u}}$ is also divergence-free. □

Lemma 5. *Let $\boldsymbol{u} \in V_{\mathrm{div}}$. Then $\mathsf{Q}\hat{\boldsymbol{u}}$ is uniquely determined by $\mathsf{P}\hat{\boldsymbol{u}}$ and given by the formula $\mathsf{Q}\hat{\boldsymbol{u}} = \boldsymbol{\Omega}\hat{g}$ with*

$$\hat{g} \circ \boldsymbol{\chi} = -s \int_{-s^{-1}}^{0} \zeta \, \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \, d\zeta - \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}') \, d\zeta' \,. \tag{B.3}$$

Proof. With $\boldsymbol{u} = \mathsf{P}\boldsymbol{u} + \mathsf{Q}\boldsymbol{u}$, the divergence condition reads

$$0 = \boldsymbol{\nabla} \cdot \boldsymbol{u} = \boldsymbol{\nabla} \cdot \mathsf{P}\boldsymbol{u} + \boldsymbol{\nabla} \cdot \mathsf{Q}\boldsymbol{u} \,. \tag{B.4}$$

Recalling that $\mathsf{Q} = \boldsymbol{\Omega}\boldsymbol{\Omega}^{\mathsf{T}}$ and setting $g = \boldsymbol{\Omega} \cdot \boldsymbol{u}$, we can write $\mathsf{Q}\boldsymbol{u} = \boldsymbol{\Omega} g$. Then,

$$\boldsymbol{\Omega} \cdot \boldsymbol{\nabla}\hat{g} = \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} g = -\boldsymbol{\nabla} \cdot \mathsf{P}\boldsymbol{u} \tag{B.5}$$

so that, due to (29),

$$\partial_\zeta(\hat{g} \circ \boldsymbol{\chi}) = (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla}\hat{g}) \circ \boldsymbol{\chi} = -(\boldsymbol{\nabla} \cdot \mathsf{P}\boldsymbol{u}) \circ \boldsymbol{\chi} = -\boldsymbol{\nabla} \cdot (\mathsf{SP}\boldsymbol{u} \circ \boldsymbol{\chi}) = -\boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \,. \tag{B.6}$$

The third equality in (B.6) is due to (30b) and the last equality is due to $\boldsymbol{u} = \hat{\boldsymbol{u}} + \bar{\boldsymbol{u}}$, where $\bar{\boldsymbol{u}} \circ \boldsymbol{\chi}$ is independent of ζ, so that the entire contribution from $\bar{\boldsymbol{u}}$ vanishes by Lemma 3. Equation (B.6) determines g uniquely up to a constant of integration on each of the characteristic lines; we write $g = \hat{g} + \bar{g}$ and choose $\bar{g}$ as this constant of integration. The condition that $\hat{g}$ is mean-free along each characteristic line implies

$$0 = s \int_{-s^{-1}}^{0} \hat{g} \circ \boldsymbol{\chi} \, d\zeta = \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - s \int_{-s^{-1}}^{0} \zeta \, \partial_\zeta(\hat{g} \circ \boldsymbol{\chi}) \, d\zeta \,. \tag{B.7}$$

Then, substituting (B.6) and (B.7) into

$$\hat{g} \circ \boldsymbol{\chi}(\zeta) = \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) + \int_{-s^{-1}}^{\zeta} \partial_\zeta(\hat{g} \circ \boldsymbol{\chi}') \, d\zeta' \,, \tag{B.8}$$

we obtain (B.3), which determines $\mathsf{Q}\hat{\boldsymbol{u}}$ uniquely. □

The next lemma provides a converse statement to Lemma 5.

Lemma 6. *Suppose $\mathsf{P}\hat{\boldsymbol{u}}$ is a given vector field which is mean-free and contained in the range of P. Define $\hat{g}$ as in Lemma 5, that is,*

$$\hat{g} \circ \boldsymbol{\chi} = -s \int_{-s^{-1}}^{0} \zeta \, \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \, d\zeta - \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}') \, d\zeta' \,. \tag{B.9}$$

Then $\hat{\boldsymbol{u}} = \mathsf{P}\hat{\boldsymbol{u}} + \mathsf{Q}\hat{\boldsymbol{u}}$ with $\mathsf{Q}\hat{\boldsymbol{u}} = \boldsymbol{\Omega}\hat{g}$ is mean-free and divergence-free.

Proof. The fact that $\hat{g}$, hence $\hat{\boldsymbol{u}}$, is mean-free is a direct consequence of the choice of constant of integration in the proof of Lemma 5.

To prove that $\hat{\boldsymbol{u}}$ is divergence-free, we take the ζ-derivative of (B.9),

$$\partial_\zeta(\hat{g} \circ \boldsymbol{\chi}) = -\boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \,. \tag{B.10}$$

This implies that

$$
\begin{aligned}
0 &= \partial_\zeta(\hat{g} \circ \boldsymbol{\chi}) + \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \\
&= (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla}\hat{g}) \circ \boldsymbol{\chi} + (\boldsymbol{\nabla} \cdot \mathsf{P}\hat{\boldsymbol{u}}) \circ \boldsymbol{\chi} \\
&= (\boldsymbol{\nabla} \cdot \mathsf{Q}\hat{\boldsymbol{u}}) \circ \boldsymbol{\chi} + (\boldsymbol{\nabla} \cdot \mathsf{P}\hat{\boldsymbol{u}}) \circ \boldsymbol{\chi} \\
&= (\boldsymbol{\nabla} \cdot \hat{\boldsymbol{u}}) \circ \boldsymbol{\chi} \,.
\end{aligned}
\tag{B.11}
$$

Since $\boldsymbol{\chi}$ is invertible, we find that $\boldsymbol{\nabla} \cdot \hat{\boldsymbol{u}} = 0$. □

Appendix C. Inner product identities for decomposed vector fields

Lemma 7. *Let $\boldsymbol{u} \in V_{\text{div}}$; we write $\hat{\boldsymbol{u}}$ to denote its mean-free component as before. Further, let $\hat{\boldsymbol{w}}$ be any mean free vector field. Then*

$$
\int_{\mathcal{D}} \hat{\boldsymbol{w}} \cdot \mathsf{Q}\hat{\boldsymbol{u}} \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\boldsymbol{C}[\hat{\boldsymbol{w}}] \, d\boldsymbol{x}
\tag{C.1}
$$

with $\boldsymbol{C}[\hat{\boldsymbol{w}}]$ defined by

$$
\boldsymbol{C}[\hat{\boldsymbol{w}}] \circ \boldsymbol{\chi} = -\mathsf{S}\boldsymbol{\nabla} \int_{-s^{-1}}^{\zeta} \boldsymbol{\Omega} \cdot \hat{\boldsymbol{w}} \circ \boldsymbol{\chi}' \, d\zeta' \,.
\tag{C.2}
$$

Proof. By Lemma 5,

$$
\int_{\mathcal{D}} \hat{\boldsymbol{w}} \cdot \mathsf{Q}\hat{\boldsymbol{u}} \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{w}} \cdot \boldsymbol{\Omega} \, \hat{g} \, d\boldsymbol{x} = s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \boldsymbol{\Omega} \cdot \hat{\boldsymbol{w}} \circ \boldsymbol{\chi} \, \hat{g} \circ \boldsymbol{\chi} \, d\boldsymbol{\xi} \,,
\tag{C.3}
$$

where $\hat{g} \circ \boldsymbol{\chi}$ is given by (B.3). As it is integrated against a mean-free vector field, the first term on the right of (B.3) does not contribute to the integral (C.3), so that

$$
\begin{aligned}
\int_{\mathcal{D}} \hat{\boldsymbol{w}} \cdot \mathsf{Q}\hat{\boldsymbol{u}} \, d\boldsymbol{x} &= -s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \boldsymbol{\Omega} \cdot \hat{\boldsymbol{w}} \circ \boldsymbol{\chi} \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}') \, d\zeta' \, d\boldsymbol{\xi} \\
&= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \int_{-s^{-1}}^{\zeta} \boldsymbol{\Omega} \cdot \hat{\boldsymbol{w}} \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi} \\
&= -s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \hat{\boldsymbol{u}} \circ \boldsymbol{\chi} \cdot \mathsf{PS}\boldsymbol{\nabla} \int_{-s^{-1}}^{\zeta} \boldsymbol{\Omega} \cdot \hat{\boldsymbol{w}} \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi} \\
&= \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\boldsymbol{C}[\hat{\boldsymbol{w}}] \, d\boldsymbol{x}
\end{aligned}
\tag{C.4}
$$

where $\boldsymbol{C}[\hat{\boldsymbol{w}}]$ is given by (C.2). We remark that the second inequality is based on integration by parts in ζ, the third equality is due to the divergence theorem. In both cases, the boundary terms vanish due to the mean-free condition on $\hat{\boldsymbol{w}}$. We have further used the symmetry of the matrices S and P. □

Lemma 8. *Let $\boldsymbol{u} = \hat{\boldsymbol{u}} + \bar{\boldsymbol{u}} \in V_{\text{div}}$ and let $\bar{\phi}$ be the vertical mean of an arbitrary scalar field. Then*

$$\int_{\mathcal{D}} \bar{\phi}\, \bar{u}_3 \, d\boldsymbol{x} = - \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P} \boldsymbol{\nabla} \bar{\phi}\, z \, d\boldsymbol{x} \, . \tag{C.5}$$

Proof. By Lemma 5, $\mathsf{Q}\hat{\boldsymbol{u}} = \boldsymbol{\Omega} \hat{g}$ with

$$\hat{g} \circ \boldsymbol{\chi}(-s^{-1}) = -s \int_{-s^{-1}}^{0} \zeta \, \boldsymbol{\nabla} \cdot (\mathsf{SP} \hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \, d\zeta \, . \tag{C.6}$$

At the bottom boundary, $\boldsymbol{k} \cdot \bar{\boldsymbol{u}} = -\boldsymbol{k} \cdot \hat{\boldsymbol{u}} = -\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} - \boldsymbol{k} \cdot \mathsf{Q}\hat{\boldsymbol{u}}$, so that

$$\begin{aligned}
\int_{\mathcal{D}} \bar{\phi}\, \bar{u}_3 \, d\boldsymbol{x} &= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \bar{\phi} \circ \boldsymbol{\chi} \left[-\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) - s \, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) \right] d\boldsymbol{\xi} \\
&= \int_{\mathbb{T}^2} \bar{\phi} \circ \boldsymbol{\chi} \left[-\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) + s^2 \int_{-s^{-1}}^{0} \zeta \, \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \, d\zeta \right] dx \\
&= - \int_{\mathbb{T}^2} \bar{\phi} \circ \boldsymbol{\chi} \, \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) \, dx \\
&\quad + s^2 \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \bar{\phi} \circ \boldsymbol{\chi} \left(\boldsymbol{\nabla} \cdot (\zeta \, \mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) - \boldsymbol{\nabla}\zeta \cdot \mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi} \right) d\boldsymbol{\xi} \, .
\end{aligned} \tag{C.7}$$

Since $\boldsymbol{\nabla}\zeta = \boldsymbol{k}$, the second term in the last integral vanishes as the vertical integration is over a mean free quantity. For the first term in the last integral, we integrate by parts. The boundary term from the upper boundary is zero. The boundary term from the lower boundary exactly cancels the integral on the second last line, so that

$$\begin{aligned}
\int_{\mathcal{D}} \bar{\phi}\, \bar{u}_3 \, d\boldsymbol{x} &= -s^2 \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \hat{\boldsymbol{u}} \circ \boldsymbol{\chi} \cdot \mathsf{PS} \boldsymbol{\nabla} (\bar{\phi} \circ \boldsymbol{\chi}) \, \zeta \, d\boldsymbol{\xi} \\
&= -s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \hat{\boldsymbol{u}} \circ \boldsymbol{\chi} \cdot \mathsf{PS} (\mathsf{A}^{\mathsf{T}} \boldsymbol{\nabla} \bar{\phi} \, z) \circ \boldsymbol{\chi} \, d\boldsymbol{\xi} \, ,
\end{aligned} \tag{C.8}$$

where the last equality is due to $\boldsymbol{\nabla}(f \circ \boldsymbol{\chi}) = (\mathsf{A}^{\mathsf{T}} \boldsymbol{\nabla} f) \circ \boldsymbol{\chi}$ and $z = s\zeta$. Noting that $\mathsf{PSA}^{\mathsf{T}} = \mathsf{P}$ and changing back to Cartesian coordinates, we obtain (C.5). □

Lemma 9. *Under the conditions of Lemma 6, there exists $\bar{\boldsymbol{u}}$, also divergence free, such that $\boldsymbol{u} = \hat{\boldsymbol{u}} + \bar{\boldsymbol{u}}$ satisfies the zero-flux boundary condition $\boldsymbol{k} \cdot \boldsymbol{u} = 0$ at $z = 0, -1$.*

Proof. By Lemma 6, $\hat{\boldsymbol{u}}$ is divergence free. We now choose $\bar{u}_3$ such that the vector field $\boldsymbol{u}$ is tangent to the top and bottom boundaries. At $z = -1, 0$, we require

$$\boldsymbol{k} \cdot \bar{\boldsymbol{u}} = -\boldsymbol{k} \cdot \hat{\boldsymbol{u}} \, . \tag{C.9}$$

Observe that

$$\begin{aligned}
\hat{u}_3 \circ \boldsymbol{\chi}(0) &= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s \, \hat{g} \circ \boldsymbol{\chi}(0) \\
&= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s \, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - s \int_{-s^{-1}}^{0} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}) \, d\zeta
\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s\,\hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) \\
&= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) + s\,\hat{g} \circ \boldsymbol{\chi}(-s^{-1})\,. \qquad \text{(C.10)}
\end{aligned}
$$

This implies $\hat{u}_3$ takes the same value at the bottom and top boundaries along any line in the direction of the axis of rotation. Consequently, we can use (C.9) to *define* $\boldsymbol{k} \cdot \bar{\boldsymbol{u}}$ at the bottom, that is, we set

$$
\bar{u}_3 \circ \boldsymbol{\chi} = -\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) - s\,\hat{g} \circ \boldsymbol{\chi}(-s^{-1})\,. \tag{C.11}
$$

Then the boundary condition (C.9) is satisfied at $z = 0$ as well.

We now choose $\bar{u}$ such that $\boldsymbol{u}$ is divergence-free. Indeed, due to Lemma 4, it suffices to ensure that $\nabla \cdot \mathsf{S}_h \mathsf{P}\bar{\boldsymbol{u}} = 0$, cf. (B.2) for an explicit expression. Setting $\bar{u} = \nabla\phi$, we see that this implies

$$
\Delta\phi = \frac{c}{s}\,\partial_y \bar{u}_3\,, \tag{C.12}
$$

which can be solved as a Poisson equation on $\mathbb{T}^2$. □

Lemma 10. *Let $\boldsymbol{u}, \boldsymbol{v} \in V_{\text{div}}$ be such that their domain-mean is zero. Then*

$$
\int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v}\,d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{J}\hat{\boldsymbol{v}}\,d\boldsymbol{x}\,. \tag{C.13}
$$

Proof. The mean of $\boldsymbol{u}$ over the domain $\mathcal{D}$ is zero if and only if the horizontal mean of $\bar{\boldsymbol{u}}$ is zero, and likewise for $\boldsymbol{v}$. Moreover, by Lemma 3, $\nabla\cdot(\mathsf{S}_h\mathsf{P}\bar{\boldsymbol{u}}) = 0$ and $\nabla\cdot(\mathsf{S}_h\mathsf{P}\bar{\boldsymbol{v}}) = 0$. Thus, there exist scalar fields ψ and θ such that $\mathsf{S}_h\mathsf{P}\bar{\boldsymbol{u}} = \nabla^\perp\psi$ and $\mathsf{S}_h\mathsf{P}\bar{\boldsymbol{v}} = \nabla^\perp\theta$. We write

$$
\mathsf{S}_h^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & s^2 \\ 0 & -cs \end{pmatrix} \tag{C.14}
$$

to denote the pseudo-inverse of S_h on $\operatorname{Range}\mathsf{P}$. It satisfies

$$
\mathsf{S}_h\,\mathsf{S}_h^{-1} = \mathsf{I}_2\,, \tag{C.15a}
$$

where I_2 denotes the 2×2 identity matrix, and

$$
\mathsf{S}_h^{-1}\,\mathsf{S}_h\,\mathsf{P} = \mathsf{P}\,. \tag{C.15b}
$$

Then, $\mathsf{P}\bar{\boldsymbol{u}} = \mathsf{S}_h^{-1}\,\nabla^\perp\psi$ and $\mathsf{P}\bar{\boldsymbol{v}} = \mathsf{S}_h^{-1}\,\nabla^\perp\theta$. Further, observing that

$$
\mathsf{S}_h^{-\mathsf{T}}\,\mathsf{J}\,\mathsf{S}_h^{-1} = s\,\mathsf{J}_2\,, \tag{C.16}
$$

where J_2 denotes the canonical 2×2 symplectic matrix, and recalling that $\mathsf{PJ} = \mathsf{J} = \mathsf{JP}$, we compute

$$
\int_{\mathcal{D}} \bar{\boldsymbol{u}} \cdot \mathsf{J}\bar{\boldsymbol{v}}\,d\boldsymbol{x} = \int_{\mathcal{D}} \mathsf{P}\bar{\boldsymbol{u}} \cdot \mathsf{J}\mathsf{P}\bar{\boldsymbol{v}}\,d\boldsymbol{x}
$$

$$
\begin{aligned}
&= \int_{\mathcal{D}} \mathsf{S}_h^{-1} \nabla^\perp \psi \cdot \mathsf{J} \mathsf{S}_h^{-1} \nabla^\perp \theta \, d\boldsymbol{x} \\
&= \int_{\mathcal{D}} \nabla^\perp \psi \cdot \mathsf{S}_h^{-\mathsf{T}} \mathsf{J} \mathsf{S}_h^{-1} \nabla^\perp \theta \, d\boldsymbol{x} \\
&= -s \int_{\mathcal{D}} \nabla^\perp \psi \cdot \nabla \theta \, d\boldsymbol{x} \,.
\end{aligned} \tag{C.17}
$$

By orthogonality of gradients and curls, the last integral is zero, which implies (C.13). □

Appendix D. Derivation of potential energy contribution to L_1

In the following, we give a detailed derivation of the potential energy contribution to the L_1-Lagrangian. Inserting the boundary condition for the transformation vector field (53) and the representation of $\mathsf{Q}\bar{\boldsymbol{v}}$ via (52), we compute:

$$
\begin{aligned}
-\int_{\mathcal{D}} \rho \, \boldsymbol{v} \cdot \boldsymbol{k} \, d\boldsymbol{x} &= -s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \rho \circ \boldsymbol{\chi} \left(\boldsymbol{k} \cdot \bar{\boldsymbol{v}} \circ \boldsymbol{\chi} + \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi} + \boldsymbol{k} \cdot \mathsf{Q}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi} \right) d\boldsymbol{\xi} \\
&= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \rho \circ \boldsymbol{\chi} \Big[\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}(-s^{-1}) + s \, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi} \\
&\qquad - s \, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) + s \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla} \cdot (\mathsf{S}\mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}') \, d\zeta' \Big] \, d\boldsymbol{\xi} \\
&= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \rho \circ \boldsymbol{\chi} \Big[\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}(-s^{-1}) - \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi} \\
&\qquad + \int_{-s^{-1}}^{\zeta} \partial_\zeta (\boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}') \, d\zeta' + s \int_{-s^{-1}}^{\zeta} \nabla \cdot (\mathsf{S}_h \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}') \, d\zeta' \Big] \, d\boldsymbol{\xi} \\
&= -s^2 \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \nabla \rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \mathsf{S}_h \mathsf{P}\hat{\boldsymbol{v}} \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi}
\end{aligned} \tag{D.1}
$$

Further, inserting (50), using the identity

$$
\mathsf{S}_h \, \mathsf{J}^\mathsf{T} = -s^{-1} \, \mathsf{J}_2 \, \mathsf{P}_h \,, \tag{D.2}
$$

and noting that horizontal gradients and composition with $\boldsymbol{\chi}$ commute, we obtain

$$
\begin{aligned}
-\int_{\mathcal{D}} \rho \, \boldsymbol{v} \cdot \boldsymbol{k} \, d\boldsymbol{x} &= -s^2 \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \nabla \rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \mathsf{S}_h \mathsf{J}^\mathsf{T} \left(-\tfrac{1}{2} \, \hat{\boldsymbol{V}} + \lambda \, \hat{\boldsymbol{V}}_g \right) \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi} \\
&= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \nabla \rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \mathsf{J}_2 \, \mathsf{P}_h \left(-\tfrac{1}{2} \, \hat{\boldsymbol{V}} + \lambda \, \hat{\boldsymbol{V}}_g \right) \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi} \\
&= s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \nabla^\perp \rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \mathsf{P}_h \left(\tfrac{1}{2} \, \hat{\boldsymbol{V}} - \lambda \, \hat{\boldsymbol{V}}_g \right) \circ \boldsymbol{\chi}' \, d\zeta' \, d\boldsymbol{\xi} \,.
\end{aligned} \tag{D.3}
$$

Inserting the thermal wind relation (36), integrating by parts with respect to ζ, and changing variables, we continue the computation:

$$
\begin{aligned}
-\int_{\mathcal{D}} \rho\, \boldsymbol{v} \cdot \boldsymbol{k}\, d\boldsymbol{x} &= -s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \partial_\zeta (\hat{u}_g \circ \boldsymbol{\chi}) \cdot \int_{-s^{-1}}^{\zeta} \mathsf{P}_h \left(\tfrac{1}{2} \hat{\boldsymbol{V}} - \lambda \hat{\boldsymbol{V}}_g \right) \circ \boldsymbol{\chi}'\, d\zeta'\, d\boldsymbol{\xi} \\
&= \int_{\mathcal{D}} \hat{u}_g \cdot \mathsf{P}_h \big(\tfrac{1}{2} \hat{\boldsymbol{V}} - \lambda \hat{\boldsymbol{V}}_g \big)\, d\boldsymbol{x} \\
&= \int_{\mathcal{D}} \hat{u}_g \cdot \big(\tfrac{1}{2} \hat{u} - \lambda \hat{u}_g \big)\, d\boldsymbol{x} \qquad \text{(D.4)}
\end{aligned}
$$

where, in the last step, we have made use of Lemma 7, Lemma 8, and the fact that $\boldsymbol{k} \cdot \boldsymbol{u}_g = 0$.

References

Akramov, I., & Oliver, M. (2020). On the existence of solutions to a bi-planar Monge–Ampère equation. *Acta Math. Sci.*, *40*(2), 379–388.

Allen, J. S., Holm, D. D., & Newberger, P. A. (2002). Toward an extended-geostrophic Euler–Poincaré model for mesoscale oceanographic flow. In J. Norbury & I. Roulstone (Eds.), *Large-scale atmosphere–ocean dynamics* (Vol. 1, pp. 101–125). Cambridge University Press.

Arnold, V. I., & Khesin, B. A. (1999). *Topological methods in hydrodynamics.* Springer.

Babin, A., Mahalov, A., & Nicolaenko, B. (2002). Fast singular oscillating limits of stably-stratified 3d Euler and Navier–Stokes equations and ageostrophic wave fronts. In J. Norbury & I. Roulstone (Eds.), *Large-scale atmosphere–ocean dynamics* (Vol. 1, pp. 126–201). Cambridge University Press.

Benamou, J. D., & Brenier, Y. (1998). Weak existence for the semigeostrophic equations formulated as a coupled Monge–Ampère/transport problem. *SIAM J. Appl. Math.*, *58*(5), 1450–1461.

Brummell, N. H., Hurlburt, N. E., & Toomre, J. (1996). Turbulent compressible convection with rotation. I. Flow structure and evolution. *Astrophys. J.*, *473*, 494–513.

Çalık, M., Oliver, M., & Vasylkevych, S. (2013). Global well-posedness for the generalized large-scale semigeostrophic equations. *Arch. Ration. Mech. Anal.*, *207*(3), 969–990.

Colombo, M. (2017). *Flows of non-smooth vector fields and degenerate elliptic equations. with applications to the vlasov-poisson and semigeostrophic systems* (Vol. 22). Edizioni della Normale, Pisa.

Cullen, M. J. P. (2006). *A mathematical theory of large-scale atmosphere/ocean flow.* Imperial College Press.

Cullen, M. J. P., & Purser, R. J. (1984). An extended Lagrangian theory of semi-geostrophic frontogenesis. *J. Atmos. Sci.*, *41*(9), 1477–1497.

de laTorre Juárez, M., Fisher, B., & Orton, G. (2002). Large scale geostrophic winds with a full representation of the Coriolis force: Application to IR observations of the upper Jovian troposphere. *Geophys. Astrophys. Fluid Dyn.*, *96*, 87–114.

Dellar, P. J. (2011). Variations on a beta-plane: derivation of non-traditional beta-plane equations from Hamilton's principle on a sphere. *J. Fluid Mech.*, *674*, 174–195.

Dritschel, D. G., Gottwald, G. A., & Oliver, M. (2017). Comparison of variational balance models for the rotating shallow water equations. *J. Fluid Mech.*, *822*, 689–716.

Ebin, D. G., & Marsden, J. (1970). Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)*, *92*, 102–163.

Eliassen, A. (1948). The quasi-static equations of motion with pressure as independent variable. *Geofys. Publ.*, *17*, 1–44.

Embid, P. F., & Majda, A. J. (1996). Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity. *Comm. Partial Differential Equations*, *21*(3-4), 619–658.

Embid, P. F., & Majda, A. J. (1998). Low Froude number limiting dynamics for stably stratified flow with small or finite Rossby numbers. *Geophys. Astrophys. Fluid Dyn.*, *87*, 1–50.

Franzke, C. L. E., Oliver, M., Rademacher, J. D. M., & Badin, G. (2019). Multi-scale methods for geophysical flows. In C. Eden & A. Iske (Eds.), *Energy Transfers in Atmosphere and Ocean* (pp. 1–51). Cham: Springer.

Gerkema, T., & Shrira, V. I. (2005a). Near-inertial waves in the ocean: beyond the 'traditional approximation'. *J. Fluid Mech.*, *529*, 195–219.

Gerkema, T., & Shrira, V. I. (2005b). Near-inertial waves on the "nontraditional" β-plane. *J. Geophys. Res. Oceans*, *110*(C1).

Gerkema, T., Zimmerman, J. T. F., Maas, L. R. M., & van Haren, H. (2008). Geophysical and astrophysical fluid dynamics beyond the traditional approximation. *Rev. Geophys.*, *46*(2), RG2004.

Gilbarg, D., & Trudinger, N. S. (2001). *Elliptic partial differential equations of second order.* Springer.

Gilbert, A. D., & Vanneste, J. (2018). Geometric generalised Lagrangian-mean theories. *J. Fluid Mech.*, *839*, 95–134.

Gottwald, G. A., & Oliver, M. (2014). Slow dynamics via degenerate variational asymptotics. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, *470*(2170), 20140460.

Hayashi, M., & Itoh, H. (2012). The importance of the nontraditional Coriolis terms in large-scale motions in the tropics forced by prescribed cumulus heating. *J. Atmos. Sci.*, *69*, 2699–2716.

Holm, D. D., Marsden, J. E., & Ratiu, T. S. (1998). The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, *137*(1), 1–81.

Hoskins, B. J. (1975). The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, *32*(2), 233–242.

Ju, Q., & Mu, P. (2019). Low Froude and Rossby number three-scale singular limits of the rotating stratified Boussinesq equations. *Z. Angew. Math. Phys.*, *70*(6), Paper No. 161, 20.

Julien, K., Knobloch, E., Milliff, R., & Werne, J. (2006). Generalized quasi-geostrophy for spatially anisotropic rotationally constrained flows. *J. Fluid Mech.*, *555*, 233–274.

Kafiabad, H. A., & Bartello, P. (2016). Balance dynamics in rotating stratified turbulence. *J. Fluid Mech.*, *795*, 914–949.

Kafiabad, H. A., & Bartello, P. (2017). Rotating stratified turbulence and the slow manifold. *Comput. & Fluids*, *151*, 23–34.

Kafiabad, H. A., & Bartello, P. (2018). Spontaneous imbalance in the non-hydrostatic Boussinesq equations. *J. Fluid Mech.*, *847*, 614–643.

Kafiabad, H. A., Vanneste, J., & Young, W. R. (2021). Wave-averaged balance: a simple example. *J. Fluid Mech.*, *911*, Paper No. R1.

Landau, L. D., & Lifshitz, E. M. (1976). *Mechanics* (3rd ed.). Butterworth-Heinemann.

Lucas, C., McWilliams, J. C., & Rousseau, A. (2017). On nontraditional quasi-geostrophic equations. *ESAIM Math. Model. Numer. Anal.*, *51*(2), 427–442.

Majda, A. (2003). *Introduction to pdes and waves for the atmosphere and ocean.* American Mathematical Society.

McIntyre, M. E., & Roulstone, I. (2002). Are there higher-accuracy analogues of semi-geostrophic theory? In J. Norbury & I. Roulstone (Eds.), *Large-scale atmosphere–ocean dynamics* (Vol. 2, pp. 301–364). Cambridge University Press.

Nieves, D., Grooms, I., Julien, K., & Weiss, J. B. (2016). Investigations of non-hydrostatic, stably stratified and rapidly rotating flows. *J. Fluid Mech.*, *801*, 430–458.

Oliver, M. (2006). Variational asymptotics for rotating shallow water near geostrophy: a transformational approach. *J. Fluid Mech.*, *551*, 197–234.

Oliver, M. (2014). A variational derivation of the geostrophic momentum approximation. *J. Fluid Mech.*, *751*, R2 (10 pages).

Oliver, M., & Vasylkevych, S. (2013). Generalized LSG models with spatially varying Coriolis parameter. *Geophys. Astrophys. Fluid Dyn.*, *107*, 259–276.

Oliver, M., & Vasylkevych, S. (2016). Generalized large-scale semigeostrophic approximations for the f-plane primitive equations. *J. Phys. A: Math. Theor.*, *49*, 184001.

Palais, R. S. (1968). *Foundations of Global Non-Linear Analysis.* W. A. Benjamin, Inc., New York-Amsterdam.

Pedlosky, J. (1987). *Geophysical fluid dynamics* (Second ed.). Springer.

Ripa, P. (1981). Symmetries and conservation laws for internal gravity waves. *AIP Conference Proceedings*, *76*(1), 281–306.

Roulstone, I., & Sewell, M. J. (1997). The mathematical structure of theories of semi-geostrophic type. *Philos. Trans. Roy. Soc. London Ser. A*, *355*(1734), 2489–2517.

Salmon, R. (1982). Hamilton's principle and ertel's theorem. *AIP Conference Proceedings*, *88*(1), 127–135.

Salmon, R. (1983). Practical use of Hamilton's principle. *J. Fluid Mech.*, *132*, 431–444.

Salmon, R. (1985). New equations for nearly geostrophic flow. *J. Fluid Mech.*, *153*, 461–477.

Salmon, R. (1988). Semigeostrophic theory as a Dirac-bracket projection. *J. Fluid Mech.*, *196*, 345–358.

Salmon, R. (1996). Large-scale semigeostrophic equations for use in ocean circulation models. *J. Fluid Mech.*, *318*, 85–105.

Stewart, A. L., & Dellar, P. J. (2010). Multilayer shallow water equations with complete Coriolis force. Part 1. Derivation on a non-traditional beta-plane. *J. Fluid Mech.*, *651*, 387.

Stewart, A. L., & Dellar, P. J. (2012). Multilayer shallow water equations with complete Coriolis force. Part 2. Linear plane waves. *J. Fluid Mech.*, *690*, 16–50.

Vallis, G. K. (2017). *Atmospheric and oceanic fluid dynamics: Fundamentals and large-scale circulation* (Second ed.). Cambridge University Press.

Wetzel, A. N., Smith, L. M., Stechmann, S. N., & Martin, J. E. (2019). Balanced and unbalanced components of moist atmospheric flows with phase changes. *Chin. Ann. Math. Ser. B*, *40*(6), 1005–1038.

White, A. A. (2002). A view of the equations of meteorological dynamics and various approximations. In J. Norbury & I. Roulstone (Eds.), *Large-scale atmosphere–ocean dynamics* (Vol. 1, pp. 1–100). Cambridge University Press.

Whitehead, J. P., Haut, T., & Wingate, B. A. (2018). The effect of two distinct fast time scales in the rotating, stratified Boussinesq equations: variations from quasi-geostrophy. *Theor. Comput. Fluid Dyn.*, *32*(6), 713–732.

Whitehead, J. P., & Wingate, B. A. (2014, 10). The influence of fast waves and fluctuations on the evolution of the dynamics on the slow manifold. *J. Fluid Mech.*, *757*, 155–178.