Variational Model Reduction for Non-hydrostatic Stratified Flows in the Mid-latitude and the Equator

by

Gözde Özden

A thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
in Mathematics

Approved Dissertation Committee

Prof. Dr. Marcel OLIVER, Jacobs University Bremen

Prof. Dr. Sergey DANILOV, Alfred Wegener Institut

Dr. Sergiy VASYLKEVYCH, Universität Hamburg

Date of the defense: 23 June 2022

Mathematics and Logistics

Statutory Decleration

Family Name, Given/First Name	Gözde Özden
Matriculationnumber	20331969
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	PhD-Thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.
The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.
Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

. . . .

Date, Signature

To my family…

Summary

This thesis studies balance models for a rotating stratified three-dimensional fluid on a tangent plane with full Coriolis force. Derivations are done for two different regions, namely mid-latitude and equator, which we considered separately.

Each model is studied via a variational approach which is based on Lagrangian dynamics assuming smallness of the Rossby number and allowing for anisotropy in the horizontal length scales. We assume semigeostrophic scaling, akin to the derivation of the L_1 model by Salmon (1985) for the rotating shallow water equations. Contrary to Salmon's derivation, we start with an arbitrary change of coordinates and then choose the transformation to fix the degeneracy on the first order of the Lagrangian, L_1, as suggested by Oliver (2006). In our setting, the full projection of the rotation vector of the Earth is considered, so that the horizontal component of the Coriolis vector is taken into account. For each model, conservation laws for the energy and the potential vorticity are valid because of the Hamiltonian structure.

Our first model is a balance model for the mid-latitude on the f-plane in semigeostrophic scaling. It is an extended version of the primitive equation model derived by Oliver and Vasylkevych (2016). We achieve to obtain a valid balance model keeping the L_1 structure. It includes two prognostic equations for density and potential vorticity, and one kinematic equation which provides a relationship between some component of the velocity field and prognostic variables. It is an elliptic equation when the fluid is stably stratified and all prognostic variables have sufficiently small fluctuation. It is the most general model obtained so far in semi-geostrophic scaling.

The other model concerns balance model on the equatorial β-plane. The situation differs totally from the f-plane model because the vertical component of the Coriolis vector is not constant on the β-plane. This brings additional complexity. Other than this, obtaining a workable scaling is a major challenge in this region. We discuss three different scaling assumptions corresponding to different dynamics. We then focus on one of them for which the rotation is dominant. Under the additional assumption of construction of zero-meridional velocity as suggested by the leading order dynamics, an equatorial balance model is obtained.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Dr. Marcel Oliver. He give me excellent feedback and priceless scientific guidance during this process.

I thank Dr. Sergiy Vasylkevych for his encouragement and countless long discussions. I must thank Prof. Dr. Sergey Danilov, being the member of my thesis committee. I am grateful for his valuable suggestions.

I am sincerely thankful to Prof. Dr. Nedjeljka Žagar for the splendid opportunity to accept me work within her group which help to keep my eyes open on my research topic.

This research was supported by subproject M2 of the Collaborative Research Centre TRR181 (Energy Transfers in Atmosphere and Ocean), funded through grant 274762653 by the German Research Foundation (DFG). I am happy to be a part of this project as it gave me a chance to get in touch with colleagues from different disciplines. I thank the graduate school GLOMAR for giving the opportunity to improve my skills.

I am thankful Applied Analysis group at Jacobs University Bremen, especially Khadeeja Afzal with whom I always enjoyed discussions in the office. I am deeply grateful to all the members of The Atmospheric Dynamics and Predictability group at Hamburg University for their support. I enjoyed and learned a lot from them during lunch breaks.

Last but not the least, I would like to express my gratitude to my family and my friends who have never withheld their support and always encouraged me on this journey.

Contents

List of Figures

Introduction

Balance is a well-known concept in the geophysical fluid dynamics. It is very useful in climate sciences due to the link between the balance motion and large-scale atmospheric or oceanic circulation processes. Basically, a balance model is obtained by filtering out inertia-gravity waves which have a higher frequency set comparing to the other waves, especially Rossby waves. Ideally, the resulting model does not contain any inertia-gravity wave, as the model includes only the slow dynamics. One can think that inertia-gravity waves evolve on a small time scale and they do not interact with other waves according to linear wave theory. Then, the scale separation is easier in linear dynamics. However, it becomes a challenge in non-linear models because of the interactions. At this point, some characteristic ratios like the Rossby number and Froude number help to make this elimination via scaling.

One of the most famous scaling was described by Charney (1948). It is quasi-geostrophic scaling. However, it is not proper for large-scale dynamics. Around that time, semi-geostrophic scaling was introduced by Eliassen (1948), but it gained the popularity after Hoskins and Bretherton (1972) and Hoskins (1975). Mostly, Rossby number or Froude number is used to obtain the dynamics in these specific scalings. On the other hand, Leith (1980) obtained the balance model in quasi-geostrophic scaling, but without using the Rossby number. Their method might be useful in regions like the equator where the Rossby number vanishes.

There are mainly two types of equations which have dynamically distinct character in a balance model. One of them is the diagnostic eqution for the fast variable. It is known as the balance relation, too. In other words, it is an invariant object through a phase-space constraint (Franzke et al., 2019). The entire motion is controlled by it because it is acting as a constraint. The other one is the prognostic equation or the evolution equation for the slow variable. It is mostly for a conserved quantity like potential vorticity. It gives the dynamics of the model (McIntyre and Norton, 2000) (Hoskins, McIntyre, and Robertson, 1985). The importance of having this conserved quantity in the model directly takes the attention to models based on the Hamilton's principle because it provides the conserved quantities automatically. Salmon (1982; 1985; 1988) is the pioneer in this field. Roulstone and Sewell (1996), Allen, Holm, and Newberger (2002) followed the Hamilton's principle on their balance models.

In our derivations, the balance model compromises of the set of equations of motion derived via variational principle which is applied to the Lagrangian of the fluid. One of the most important advantages of this approach is the relationship with conservation laws because this approach provides the symmetry property. More precisely, the resulting model brings conserved quantities by Noether's theorem. For instance, the fluid has a symmetry called as particle labeling symmetry. According to it, the Lagrangian of the fluid is invariant under composition of the flow map (Holm, Marsden, and Ratiu, 1998; Franzke et al., 2019). The related conserved quantity is the potential vorticity. This property makes it convenient to solve the system even if it changes its behavior but remains conserved. Then, conservation laws help to study the flow behavior.

We use the full Coriolis force in our models. It is obtained by the non-zero horizontal component of the Coriolis vector, and so it includes two more terms implied with "full". One of them appears in the vertical component of the Coriolis force and the other one is on the horizontal term with the vertical velocity. The axis of rotation does not project to the local vertical with this assumption. Then, the model is more realistic than the one with the traditional approximation, which retains the projection of the axis of rotation to the local vertical. Thus, it is expected that these extra terms cause some complexity on the resulting model. Effects of them firstly studied on giant planets by Busse (1994) and Yano (1998). More general phenomena related to it is review by Gerkema et al. (2008a).

The usage of the full Coriolis force is more pronounced to explain some missing parts such as dynamics on internal wave (Gerkema and Shrira, 2005b), equatorial circulations like Madden-Julian oscillation (Hayashi and Itoh, 2012) and Intertropical Convergence Zone (ITCZ) (Ong and Roundy, 2019). Then, new components which are ignored in the traditional approach might bring new perspectives on the theory. In addition to these phenomena, Kohma and Sato (2013a) show how the horizontal component of the Coriolis vector promotes the evolution of edge waves like boundary Kelvin wave and topographic Rossby waves using anelastic equations set on f-plane.

Our derivations are done for two different regions, namely mid-latitude and equator. We start with the mid-latitude to show that a non-hydrostatic balance model is viable. Then, we apply our approach to the equatorial model in the long-wave scaling to examine the feasibility of the method around the equator and we aim to obtain a valid balance model for this specific region.

Coriolis vector aligns parallel to the local horizontal plane around the equator. Therefore, the traditional approximation is not proper for rotating flows in this region because it ignores the rotation by having zero components there. In other words, the fluid does not rotate in the same sense as in the mid-latitude. Thus, the horizontal component of this vector gains importance in order to keep the rotation effect.

There is a clear time scale separation on generated waves at least in the mid-latitude, as they have different characters. For example, inertia-gravity waves are moving very fast compared to Rossby waves because these waves are high frequency waves. By choosing a proper scaling, it is possible to eliminate them from the model. However, the picture changes dramatically around the equator since two more waves, which have both fast and slow parts, appear in this region.

Matsuno (1966) made the first description of the equatorial waves as an eigenmode solution using shallow water equations on the β-plane. He linearized inviscid shallow water equations about the stationary state. He showed that Kelvin waves and mixed Rossby-gravity waves which are generated only around the equator in addition to Rossby and inertia-gravity waves. These new types have different characters from Rossby waves and inertia gravity waves which move in certain speed, fast or slow. For instance, Kelvin waves with low wavenumbers behave slowly, but higher wavenumbers are fast. Similarly, mixed Rossby-gravity waves, which are sometimes called Yanaii waves, westward mixed Rossby-gravity waves move slowly like Rossby waves. However, the eastward part of them is fast like inertia-gravity waves. Thus, it is hard to apply known scale separation in this region. The existence of these new waves were shown by observations (Wallace, 1968; Kiladis et al., 2009) and so their slow parts must be included into the balance model.

Around the equator, the surface temperature is quite warm. This causes the wave-guide effect so that some waves like Kelvin wave are trapped in this region. This feature has a significant influence on the weather prediction for the mid-latitude in the long term (Dutrifoy and Majda, 2006). Then, it is worth to have a balance model near the equator. However, present equatorial models do not satisfy the notion of balance, which implies the dynamics without only inertia-gravity waves. Depending on the approach, some slow waves do not appear on the model (Verkley and Velde, 2010; Mohebalhojeh and Theiss, 2011; Chan and Shepherd, 2013; McIntyre, 2015). Then, the approach with proper scaling is very important for this critical region.

For the balance model, the full Coriolis force is suggested to handle the singularity in the vicinity of the equator by Verdière and Schopp (1994). They show that the hydrostatic approximation can be applied to the upper part of the ocean (above the characteristics) and it is necessary to use the non-traditional term for the rest. This is taken up in Section 1.2.2. Then, we keep the horizontal component of it because we work on the characteristics for both regions and we assume infinitely deep layer only around the equator.

We choose rotating Boussinesq equations on the local tangent plane on which the potential term covers the pressure and the geopotential. This system of equations is based on the perturbation on the density and the pressure, which are considered more important for incompressible and compressible fluids, respectively. Here, we consider the stratified and incompressible fluid. Our derivations are done for non-hydrostatic dynamics. Specifically, we work on the f-plane for the mid-latitude and we use β-plane for the equatorial case. Then, this work may be used as a reference for other giant planets such as Jupiter and Saturn since shallow layer assumption fails (Juárez, Fisher, and Orton, 2002).

Last, we note that non-hydrostatic approximation needs to be integrated to present models as the horizontal resolution is getting higher by improvements in the performance on computing technology. For instance, the European Centre for Medium Range Weather Forecasts (ECMWF) nearly doubles the horizontal resolution every 8 years (Wedi, 2014). Then, the higher resolution on the horizontal scale makes the vertical scale comparable with the horizontal one. As a result, hydrostatic models become invalid near feature. Hence, we apply this approximation on our models too and so we keep the vertical momentum equation.

The aim of the thesis

The aim of this work is to study variational balance models in the mid-latitude and in the vicinity of the equator. We derive a model for each case with the full Coriolis force. Although we have accurate models with the traditional approximation, recently it was shown that neglected components have significant effects under specific circumstances by Gerkema and Shrira (2005a). Considering the full Coriolis force, this dissertation firstly addresses the following question:

- Is it possible to extend the balance model which is derived by primitive equations (Oliver and Vasylkevych, 2016) to the three-dimensional case with the full Coriolis force on f-plane?

We observe that the resulting model allows us to work on a three-dimensional model with extra terms coming from the horizontal component of the Coriolis vector. Following it, we question the similar structure of the model for another region, the equator. It is known that a balance model is a challenge in this region because the traditional Coriolis component vanishes. However, the non-traditional component is claimed to handle this complexity (Verdière and Schopp, 1994). Therefore, this dissertation focuses on the next question:

- How can a balance model be constructed in the vicinity of the equator via variational approach on the β-plane?

While working on the second question, we use the long-wave scaling specifically because we use the anisotropy on the horizontal length scales as a small number. The main aim while answering this question is to obtain a valid balance model covering all slow waves in the long-wave scaling.

Plan of the thesis

This dissertation includes one peer-reviewed paper and one manuscript addressing the aforementioned questions in the study of balance models for the mid-latitude and the equator one by one.

The outline of the thesis is described as follows. In this part, the introduction of the research topic and the motivation are given. The rest of the thesis is divided to five main chapters. In the first chapter, we elaborate on some basic concepts such as the Boussinesq equations, the full Coriolis force and the scale analysis. Moreover, the model equation with the full Coriolis force is shown on the f-plane and β-plane separately because the geometry and so the scaling differs on them. In Chapter 2, the idea of the Hamilton's principle and how to apply it to the geostrophic flows are given. Chapter 3 represents the full text of the peer-previewed article about the balance model on the f-plane. Details of some derivations are deferred to the Appendix at the end of this chapter. Chapter 4 provides the manuscript about balance model on the β-plane in the vicinity of the equator. Chapter 5 draws the summary of the research findings and the outlook. Finally, Appendix A gives a brief derivation on one of the scaling option for the equator and Appendix B shows how to derive the horizontal mean field on the f-plane.

Chapter 1

Model

1.1 Boussinesq equations

The idea of Boussinesq approximation comes from the density decomposition. It is used more common in the ocean dynamics as it ignores the density fluctuation and so the fluid is stratified (LeBlond and Mysak, 1978; Franzke et al., 2019). For compressible fluids, some more assumptions are necessary. For instance, Boussinesq equations do not support acoustic waves and so the velocity must be less than the speed of the sound. On our models, we assume that the fluid is strongly stratified so that the total density is decomposed as the perturbation density ρ' and the vertical profile $\bar{\rho}$

$$\rho(\boldsymbol{x}, t) = \bar{\rho}(z) + \rho'(\boldsymbol{x}, t)\,, \tag{1.1}$$

with the condition $\bar{\rho} \gg \rho'$. Besides, the flow is assumed to be inviscid and incompressible.

To economize the notation, we use boldface for three-dimensional vector fields and non-boldface without suffix for horizontal vector fields unless otherwise stated. For instance, u_i for $i = 1, 2, 3$ represents a component of the velocity field, $\boldsymbol{u} = (u_1, u_2, u_3)$ is a three-dimensional velocity field, and $u = (u_1, u_2)$ is the horizontal velocity field. Similarly, we introduce the gradient operator as $\boldsymbol{\nabla} = (\partial_x, \partial_y, \partial_z)$ acting on three-dimensional fields and $\nabla = (\partial_x, \partial_y)$ acting on two-dimensional fields. Then, the system is

$$D_t\boldsymbol{u} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\frac{1}{\rho_0}\boldsymbol{\nabla} p - \frac{g}{\rho_0}\rho\boldsymbol{k}\,, \tag{1.2a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0\,, \tag{1.2b}$$

$$D_t\rho = 0\,. \tag{1.2c}$$

where p is the pressure and, g is the gravity constant and the Coriolis vector with components is

$$2\mathbf{\Omega} = \begin{pmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{pmatrix} . \tag{1.3}$$

It is known as angular velocity too. Components of it depend on the constant latitude ϕ. We define the material derivative as

$$D_t = \partial_t + \boldsymbol{u} \cdot \boldsymbol{\nabla} . \tag{1.4}$$

For simplicity, the diffusion is neglected because we analyze the large-scale circulation where the dissipation range is small. Besides, we want to keep the Hamiltonian structure on the model, so that the energy is conserved with the lack of dissipation. The equation system in (1.2) provides the general model and some further assumptions are done via scaling in Section 1.3 and 1.4.

Rigid lid approximation is used for the vertical boundaries. According to this approximation, the fluid domain is restricted to a flat surface, but free surface displacements are recovered via the pressure equation (Gill, 1982). For lateral boundaries, the domain is periodic at least on the mid-latitude model. We have a periodic boundary condition only in the x-direction for the equatorial model.

The fluid is restricted vertically in a non-dimensional unity length for the mid-latitude model. It has certain boundaries at the top and the bottom. However, we have no strict bottom boundary around the equator. For simplicity, it is assumed that the vertical layer is infinitely deep. Similarly the meridional plane is taken infinitely long as the meridional velocity vanishes towards the boundary. In the vertical direction, there is zero-flux boundary condition i.e,

$$\boldsymbol{k} \cdot \boldsymbol{u} = 0 \quad \begin{cases} \text{at } z = -1, 0 & \text{for the mid-latitude} , \\ \text{at } z = 0 & \text{for the equator} , \end{cases} \tag{1.5}$$

where vertical boundaries appear only at the surface in the vicinity of the equator. Therefore, we define the corresponding domain for each region as

$$\mathcal{D} = \begin{cases} \mathbb{T}^2 \times [-1, 0] & \text{for the mid-latitude}\,, \\ \mathbb{T} \times [-\infty, \infty] \times [-\infty, 0] & \text{for the equator}\,. \end{cases} \tag{1.6}$$

1.1.1 Dispersion relationship for stratified flows

In this section, we show the normal mode solution of the linear Boussinesq equations with the full Coriolis force. The linear Boussinesq equations are

$$\partial_t u_1 - \Omega_z\, u_2 + \Omega_y\, u_3 = -\frac{1}{\rho_0}\,\partial_x p\,, \tag{1.7a}$$

$$\partial_t u_2 + \Omega_z\, u_1 = -\frac{1}{\rho_0}\,\partial_y p\,, \tag{1.7b}$$

$$\partial_t u_3 - \Omega_y\, u_1 = -\frac{1}{\rho_0}\,\partial_z p - \frac{g}{\rho_0}\,\rho'\,, \tag{1.7c}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0\,, \tag{1.7d}$$

$$\partial_t \rho' + \partial_z \bar{\rho}\, u_3 = 0\,. \tag{1.7e}$$

We sum up the x derivative of (1.7a), the y derivative of (1.7b) and the z derivative of (1.7c). Then, we have

$$-\,\partial_x(\Omega_z\, u_2) + \partial_x(\Omega_y\, u_3) + \partial_y(\Omega_z\, u_1) - \partial_z(\Omega_y\, u_1) = -\frac{1}{\rho_0}\boldsymbol{\Delta} p - \frac{g}{\rho_0}\partial_z \rho'\,, \tag{1.8}$$

because $\partial_t \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$ by (1.7d). Next, we apply the three-dimensional Laplacian $\boldsymbol{\Delta}$ to (1.7c) and we obtain

$$\partial_t \boldsymbol{\Delta} u_3 - \boldsymbol{\Delta}(\Omega_y\, u_1) = -\frac{1}{\rho_0}\,\partial_z \boldsymbol{\Delta} p - \frac{g}{\rho_0}\,\boldsymbol{\Delta}\rho'\,. \tag{1.9}$$

To eliminate the pressure term, we take the partial derivative of (1.8) with respect to z which is

$$-\,\partial_{xz}(\Omega_z\, u_2) + \partial_{xz}(\Omega_y u_3) + \partial_{yz}(\Omega_z\, u_1) - \partial_{zz}(\Omega_y\, u_1) = -\frac{1}{\rho_0}\partial_z\,\boldsymbol{\Delta} p - \frac{g}{\rho_0}\partial_{zz}\rho'\,. \tag{1.10}$$

Then, we combine (1.9) and (1.10) into one equation which is

$$\partial_t \mathbf{\Delta} u_3 - \mathbf{\Delta}(\Omega_y u_1) = -\,\partial_{xz}(\Omega_z u_2) + \partial_{xz}(\Omega_y u_3) + \partial_{yz}(\Omega_z u_1) - \partial_{zz}(\Omega_y u_1) + \frac{g}{\rho_0}\partial_{zz}\rho' - \frac{g}{\rho_0}\mathbf{\Delta}\rho' \,. \tag{1.11}$$

At the same time, we follow Majda (2003) to see the effect of the gravity. We rewrite the continuity equation (1.7e)

$$\frac{g}{\rho_0}\partial_t \rho' - N^2 u_3 = 0\,, \tag{1.12}$$

where

$$N^2 = -\frac{g}{\rho_0}\frac{\partial\bar{\rho}}{\partial z} \tag{1.13}$$

is the buoyancy or Brunt-Väisälä frequency. Afterwards, we combine the time derivative of (1.11) and the second vertical derivative of (1.12) into one

$$\partial_{tt}\mathbf{\Delta} u_3 - \partial_t \mathbf{\Delta}(\Omega_y u_1) = -\,\partial_{xzt}(\Omega_z\, u_2) + \partial_{xzt}(\Omega_y\, u_3) + \partial_{yzt}(\Omega_z\, u_1) - \partial_{zzt}(\Omega_y\, u_1) + \partial_{zz}(N^2\, u_3) - \mathbf{\Delta}(N^2\, u_3)\,. \tag{1.14}$$

Next, we choose the ansatz as $(u_1, u_2, u_3, \rho) = (\check{u}_1, \check{u}_2, \check{u}_3, \check{\rho})\, e^{i(kx+ly+mz-\omega t)}$ for the eigenmode system and insert it into (1.14), it yields

$$-i\,\omega\big((k^2+l^2)\,\Omega_y + lm\,\Omega_z\big)\check{u}_1 + ikm\,\Omega_z\,\omega\,\check{u}_2 + \big(\omega^2(k^2+l^2+m^2) - ikm\,\Omega_y\,\omega - N^2(k^2+l^2)\big)\check{u}_3 = 0\,, \tag{1.15}$$

where we assume Ω_y, Ω_z and N^2 are constant.

We apply the same procedure that we did for the vertical momentum equation to the horizontal momentum equations. We obtain

$$\big(\omega^2(k^2+l^2+m^2) - ikl\,\Omega_z\,\omega + ikm\,\Omega_y\,\omega\big)\check{u}_1 - i(l^2+m^2)\,\Omega_z\,\omega\,\check{u}_2 + \big(i(l^2+m^2)\,\Omega_y\,\omega + kmN^2\big)\check{u}_3 = 0\,, \tag{1.16}$$

and

$$i\,\omega\big((k^2+m^2)\,\Omega_z + lm\,\Omega_y\big)\check{u}_1 + \big(\omega^2\,(k^2+l^2+m^2) + ikl\,\Omega_z\,\omega\big)\check{u}_2 \\ - l\,(ik\,\Omega_y\,\omega - mN^2)\check{u}_3 = 0\,. \quad (1.17)$$

At the end, we have 3 linear homogeneous equations which are (1.16), (1.17) and (1.15). This system has non-trivial solutions only if its determinant is zero. This leads to a non-linear relationship between ω and (k,l,m), which is called dispersion relation. Here, we only give a simple special case. The full dispersion relation is more complex. Firstly, in the absence of rotation, the dispersion relation is

$$\omega^4\,(k^2+l^2+m^2)^2\,\big(\omega^2\,(k^2+l^2+m^2) - N^2\,(k^2+l^2)\big) = 0\,. \quad (1.18)$$

The non-zero solution of (1.18) is

$$\omega_\pm = \pm\frac{N\sqrt{k^2+l^2}}{\sqrt{k^2+l^2+m^2}}\,. \quad (1.19)$$

Next, we consider that the Coriolis vector has a non-zero component only vertically, i.e. $\mathbf{\Omega} = (0,0,\Omega_z)$, the dispersion relation then reads

$$\omega^4\,(k^2+l^2+m^2)^2\big(\omega^2\,(k^2+l^2+m^2) - \Omega_z^2\,m^2 - (k^2+l^2)\,N^2\big) = 0\,. \quad (1.20)$$

It results

$$\omega_\pm = \pm\left(\frac{\Omega_z^2\,m^2 + (k^2+l^2)N^2}{k^2+l^2+m^2}\right)^{1/2}\,. \quad (1.21)$$

Waves on the β-plane

We assume linear Boussinesq equations on the β-plane with the full Coriolis force, see Section 1.2 for a detailed discussion. It is based on the small angle approximation

$$\sin\phi \approx \phi\,, \quad \text{and } \cos\phi \approx 1 \quad (1.22)$$

where $\phi = y/a$ with the radius of the Earth a. Then, the vertical component of the Coriolis vector is given as $\Omega_z = \beta y$ and β is given by

$$\beta = \frac{2|\Omega|}{a} \approx 2.3 \times 10^{-11} m^{-1} s^{-1} . \tag{1.23}$$

Then, the Coriolis vector is $\mathbf{\Omega} = (0, \Omega_y, \beta y)$. We follow the same procedure as above, but we choose $(u_1, u_2, u_3, \rho) = \big(\breve{u}_1(y), \breve{u}_2(y), \breve{u}_3(y)\big) e^{i(kx+mz-\omega t)}$. Then, we have

$$\begin{aligned} &\omega\Big[\big(\omega(k^2+m^2) - k\beta + ikm\,\Omega_y\big)\breve{u}_1 - k\,\beta\,y\,\partial_y\breve{u}_1 - \omega^2\,\partial_y^2\breve{u}_1\Big] \\ &+i\omega\Big[\big(k^2(\beta\,y-1) - m^2\big)\breve{u}_2 + \beta\,\partial_y\breve{u}_2 + \beta\,y\,\partial_y^2\breve{u}_2\Big] + \Big[\big(i\omega m^2\Omega_y + kmN^2\big) - i\omega\,\Omega_y\,\partial_y^2\breve{u}_3\Big] = 0\,, \end{aligned} \tag{1.24a}$$

$$\begin{aligned} &\omega\Big[-\beta\,y(k^2+m^2)\breve{u}_1 + (-\beta + im\,\Omega_y)\partial_y\breve{u}_1\Big] \\ &+\omega\Big[\big(-k^2 - m^2 + i\beta k\big)\breve{u}_2 + ik\,\beta\,y\,\partial_y\breve{u}_2 + \partial_y^2\breve{u}_2\Big] + \Big[-i\omega k\,\Omega_y + mN^2\Big]\partial_y\breve{u}_3 = 0\,, \end{aligned} \tag{1.24b}$$

$$\begin{aligned} &-i\omega\Big[\big(\Omega_y(-k^2-m^2) - im\beta - m^2\,\Omega_y\big)\breve{u}_1 - im\,\beta\,y\,\partial_y\breve{u}_1 - \Omega_y\,\partial_y^2\breve{u}_1\Big] \\ &+i\omega km\,\beta\,y\,\breve{u}_2 + \Big[\big(\omega^2(k^2+m^2) - i\omega km\,\Omega_y + \omega^2\,N^2\big)\breve{u}_3 - \omega^2\,\partial_y^2\breve{u}_3\Big] = 0\,, \end{aligned} \tag{1.24c}$$

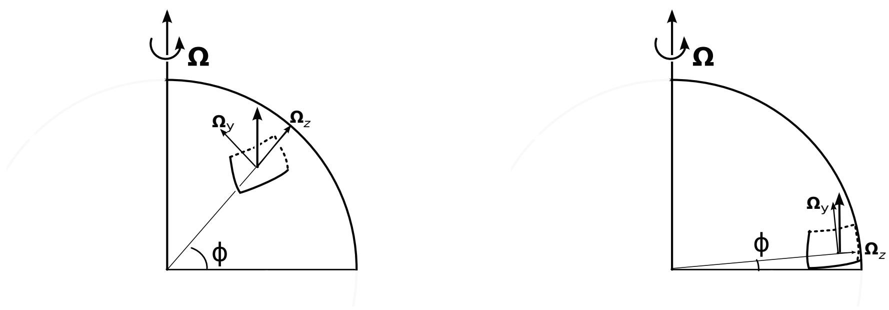
where $\breve{u}_1, \breve{u}_2$ and $\breve{u}_3$ are y-dependent. As a result, we have three equations and three variables. We expect that the reduced equation for meridional velocity matches with the result of Gerkema and Shrira (2005a). Then, the result contains Rossby waves, internal gravity waves and related Kelvin waves and mixed Rossby-gravity waves. They did not go further on their derivations as we did here because of the complexity. The full derivation of the dispersion relation for the internal waves using three-dimensional Boussinesq equation is still an open question.

1.2 Full Coriolis force

The simplest assumption with the Coriolis vector is called traditional approximation, which was defined by Eckart (1960). It is obtained by projections of the Earth's rotation vector to the local vertical. Then, the rotation of the Earth has non-zero vertical component and the Coriolis force is constructed with it. This idea was first used by Laplace (Gerkema et al., 2008a).

When the axis of rotation is retained as it is in the actual form (without any projection to the local vertical), the rotation is expressed more realistic. For this case, the corresponding force is called "full Coriolis force".

Figure 1.1 illustrates the components of the Coriolis vector on a tangent plane at the latitude ϕ. On the mid-latitude, every component has non-zero values. However, horizontal component is nearly parallel to the horizontal plane around the equator and so Ω_z is close to zero there. Then, traditional approximation, which keeps only the vertical component of the Coriolis vector, looses the validity for rotating flows near the equator where Ω_z vanishes.



(A) at the mid-latitude (B) near the equator

FIGURE 1.1: Schematic representation of components of the Coriolis vector on a tangent plane at the latitude ϕ.

Before showing components of the full Coriolis force, we define a vector $\boldsymbol{R}$ as the vector potential of the Coriolis vector to be used on further derivations and so

$$\nabla \times \boldsymbol{R} = \begin{pmatrix} -\partial_z R_y + \partial_y R_z \\ \partial_z R_x - \partial_x R_z \\ -\partial_y R_x + \partial_x R_y \end{pmatrix} = 2\boldsymbol{\Omega}\,. \tag{1.25}$$

$\boldsymbol{R}$ can be chosen as

$$\boldsymbol{R} = \frac{1}{2}\mathsf{J}\boldsymbol{x} = \frac{1}{2}\begin{pmatrix} 0 & -\Omega_z & \Omega_y \\ \Omega_z & 0 & -\Omega_x \\ -\Omega_y & \Omega_x & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\Omega_z y + \Omega_y z \\ \Omega_z x - \Omega_x z \\ -\Omega_y x + \Omega_x y \end{pmatrix}. \tag{1.26}$$

The matrix J is skew-symmetric, i.e, $\mathsf{J}^{\mathsf{T}} = -\mathsf{J}$. Moreover,

$$\mathsf{J} = \boldsymbol{\nabla} \boldsymbol{R} - \boldsymbol{\nabla} \boldsymbol{R}^{\mathsf{T}} \tag{1.27}$$

and, for every vector $\boldsymbol{v}$,

$$\mathsf{J}\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{v}\,. \tag{1.28}$$

In the following, we choose coordinates in which Ω_x is zero. Then, the unit vector in the direction of the axis of rotation is

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ \cos\phi \\ \sin\phi \end{pmatrix}, \tag{1.29}$$

for a constant latitude ϕ and

$$\mathsf{J} = \begin{pmatrix} 0 & -\Omega_z & \Omega_y \\ \Omega_z & 0 & 0 \\ -\Omega_y & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\sin\phi & \cos\phi \\ \sin\phi & 0 & 0 \\ -\cos\phi & 0 & 0 \end{pmatrix}. \tag{1.30}$$

Then, the full Coriolis force is

$$\boldsymbol{\Omega} \times \boldsymbol{u} = \mathsf{J}\boldsymbol{u} = \begin{pmatrix} -\Omega_z u_2 + \Omega_y u_3 \\ \Omega_z u_1 \\ -\Omega_y u_1 \end{pmatrix}. \tag{1.31}$$

Terms with Ω_y in (1.31) are not seen on the model derived by the traditional approximation. Following parts give some details about effects of the Coriolis vector with non-zero horizontal component.

1.2.1 Geostrophy

The simplest balance between Coriolis force and the pressure gradient is called as geostrophic balance,

$$\boldsymbol{\Omega} \times \boldsymbol{u} = -\boldsymbol{\nabla} p - \rho \boldsymbol{k}\,. \tag{1.32}$$

Obtained from the momentum equations in the leading order of a small parameter, it provides a relationship which is called *thermal wind relation*. The solution of this equation gives the geostrophic velocity. To obtain, the pressure term is eliminated by taking the curl of (1.32). Then, the left hand side of it reads by the vector identity

$$\boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \boldsymbol{u}) = \boldsymbol{\Omega}(\boldsymbol{\nabla} \cdot \boldsymbol{u}) - (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\boldsymbol{u} + (\boldsymbol{u} \cdot \boldsymbol{\nabla})\, \boldsymbol{\Omega} - \boldsymbol{u}(\boldsymbol{\nabla} \cdot \boldsymbol{\Omega})\,. \tag{1.33}$$

The first term is zero by the incompressibility feature of the fluid. The last term yields zero for each region even for the equatorial model because the vertical component of the Coriolis vector depends on y by the β-plane approximation. The result of the remaining terms in (1.33) depends on the region.

For the mid-latitude, the problem is simpler because we take the Coriolis vector as constant. Then, the reduced equation for the mid-latitude yields

$$(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\boldsymbol{u} = \begin{pmatrix} -\nabla^{\perp}\rho \\ 0 \end{pmatrix}, \tag{1.34}$$

where $\nabla^{\perp} = (-\partial_y, \partial_x)$.

The vertical component of the Coriolis vector varies with y as given in (1.68) near by the equator. Therefore, we have

$$-\,(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\boldsymbol{u} + (\boldsymbol{u} \cdot \boldsymbol{\nabla})\, \boldsymbol{\Omega} = -\boldsymbol{\nabla} \times (\rho \boldsymbol{k})\,. \tag{1.35}$$

Although we have the same reduced equation horizontally, the situation is interesting in this region. On the vertical component we have

$$-\,(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})u_3 + u_2 = 0\,. \tag{1.36}$$

This was shown by Sverdrup (1947). The integration of the vertical velocity with zero flux boundary conditions along the z-axis provides the Ekman pumping.

We refer the equation (1.35) as a general thermal wind relation for incompressible fluids. It suggests that the velocity field can be predicted as a constant of the integration of an observed quantity like density along specific lines. These lines are called characteristics and the idea on how to work with them is given in the next part.

1.2.2 Characteristics

In this section, we analyze (1.35) using the method of characteristics. Characteristics are a bunch of lines on which we have constant fields like velocity. In our models, they have a distinct character for each region.

We return our reduced problem which is a first order PDE (Partial Differential Equation) given in (1.35). For the solution, we convert it to an appropriate ODE (Ordinary Differential Equation) using method of characteristics (Evans, 1998). Then, we describe

$$\boldsymbol{x}(s) = \big(x(s), y(s), z(s)\big) \tag{1.37}$$

with the help of parameter s. We assume that $\boldsymbol{u}$ has a solution on $\mathcal{D}$ and so

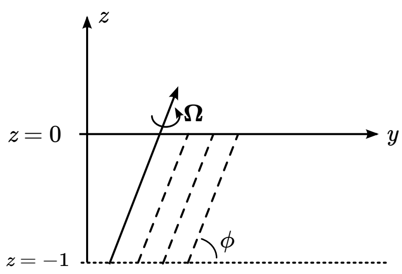
$$\boldsymbol{p}(s) := \boldsymbol{u}(\boldsymbol{x}(s)) \tag{1.38}$$

is the value of $\boldsymbol{u}$ on the curve. Then, we write characteristic equations for (1.35)

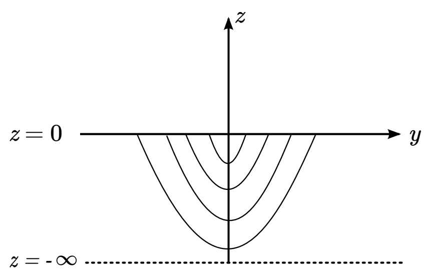
$$\dot{\boldsymbol{x}}(s) = \boldsymbol{\Omega}\,, \tag{1.39a}$$

$$\dot{\boldsymbol{p}}(s) = \boldsymbol{p}(s) \cdot \boldsymbol{\nabla}\boldsymbol{\Omega} - \boldsymbol{\nabla} \times (\rho \boldsymbol{k})\,, \tag{1.39b}$$

where the derivative is taken with respect to s, $\big(\dot{} = \frac{d}{ds}\big)$, only for the system in (1.39). Therefore, (1.35) is reduced to an ODE system via (1.39) and so characteristics provide a surface on which the solution is supposed to be. For the mid-latitude, these curves are straight lines because the first term of (1.39b) is zero. On the other hand, the situation is different around the equator as the vertical component of $\boldsymbol{\Omega}$ depends on y. The Figure 1.2 provides the schematic representation of these lines for each region on the tangent plane.



(A) at the mid-latitude



(B) near the equator

FIGURE 1.2: Characteristics for each region on the tangent plane. They are shown by dashed lines for the mid-latitude parallel to the axis of rotation. Near the equator, they are curve and the axis of rotation is tangent to them, inspired from Verdière and Schopp (1994).

Firstly, these lines intercept the vertical boundaries on the mid-latitude because there are certain points at the top and the bottom. They are straight and parallel to the axis of rotation on the local tangent plane. We use the arclenght parameter along the characteristics on the derivations for the mid-latitude as a convenient choice.

In the equator, characteristics do not intercept the bottom boundary. This is unfortunate because it is necessary to have certain integral boundaries on the derivations. To deal with it, we use the both hemisphere of the Earth. This brings that we have two boundary conditions again, but both of them are at the surface. Different than the mid-latitude, these are curves, not straight lines. It means that the value of the components are different for any two points on different latitudes. They vary with the value of y because of the $\beta-$plane approximation. Equations (1.35) and (1.34) points out this difference. For simplicity, we prefer to use $y-$coordinate for the parameterization in this region. Otherwise, we have the determinant varying with y in the mapping and this causes some complex terms in the integration.

1.3 Scale analysis for the mid-latitude

In this section, we show how to obtain the equations of motion for the mid-latitude on the f-plane. The model is expected to show the large-scale dynamics. Then, non-dimensionalization is done considering the semi-geostrophic scaling to reduce the complex physical system to a simpler one. This scaling is defined with some certain parameters like the Rossby number and the Froude number. On the following parts, these details are given for the mid-latitude model.

1.3.1 Non-dimensionalization

The main aim of the non-dimensionalization is to obtain the simplest equations of motion with minimal set of parameters that characterize the system. For this, dimensional quantities are introduced by characteristic scales. For instance, we non-dimensionalize with U, V, and W which are the characteristic zonal, meridional, and vertical velocity scales, X, L and H which are the characteristic zonal, meridional, and vertical length scales. T is the characteristic time scale, P is the characteristic pressure scale, and Γ is the typical density perturbation scale so that

$$\begin{aligned} x = X\hat{x}\,, \quad y = L\hat{y}\,, \quad z = H\hat{z}\,, \quad t = T\hat{t}\,, \quad \rho = \Gamma\hat{\rho}\,, \\ u_1 = U\hat{u}_1\,, \quad u_2 = V\hat{u}_2\,, \quad u_3 = W\hat{u}_3\,, \quad p = P\hat{p}\,, \end{aligned} \tag{1.40}$$

where dimensionless quantities are shown by hat and capital letters represent the characteristic scales for each variable.

The aspect ratio is defined as the ratio between horizontal and vertical length scales

$$\alpha = \frac{H}{L}\,, \tag{1.41}$$

and the anisotropy parameter is

$$\delta = \frac{L}{X}\,. \tag{1.42}$$

In general, we write the Coriolis vector given in (1.3) with components $2\boldsymbol{\Omega} = (0, \Omega_y, \Omega_z)$. Its non-dimensional form will be given separately for the mid-latitude and the equator later because it differs on f- and β-plane.

Afterwards, we ignore the $(\hat{\ })$ sign for dimensionless variables and Boussinesq equations are

$$\frac{U}{T}\,\partial_t u_1 + \frac{UV}{L}(u\cdot\nabla)u_1 + \frac{UW}{H}u_3\,\partial_z u_1 - \Omega_z V u_2 + \Omega_y W u_3 = -\frac{P}{X\rho_0}\,\partial_x p\,, \tag{1.43a}$$

$$\frac{V}{T}\,\partial_t u_2 + \frac{V^2}{L}(u\cdot\nabla)u_2 + \frac{VW}{H}u_3\,\partial_z u_2 + \Omega_z U u_1 = -\frac{P}{L\rho_0}\,\partial_y p\,, \tag{1.43b}$$

$$\frac{W}{T}\,\partial_t u_3 + \frac{VW}{L}(u\cdot\nabla)u_3 + \frac{W^2}{H}u_3\,\partial_z u_3 - \Omega_y U u_1 = -\frac{P}{\alpha L\rho_0}\,\partial_z p - \frac{g\Gamma}{\rho_0}\,\rho\,, \tag{1.43c}$$

$$\frac{V}{L}(\nabla\cdot u) + \frac{W}{H}\partial_z u_3 = 0\,, \tag{1.43d}$$

$$\frac{\Gamma}{T}\,\partial_t\rho + \frac{V\Gamma}{L}(u\cdot\nabla)\rho + \frac{WN_0^2\,\rho_0}{g}\,u_3\,\partial_z\rho = 0\,, \tag{1.43e}$$

where the typical buoyancy frequency is defined

$$N_0^2 = -\frac{g}{\rho_0}\left[\frac{\partial\rho}{\partial z}\right]\,, \tag{1.44}$$

where $[\partial\rho/\partial z]$ denotes the typical vertical density gradient. We note that the difference to (1.13) where the constant $\frac{\partial\bar{\rho}}{\partial z}$ is defined into N^2. Here, we keep the perturbation part of the density on the general buoyancy definition into N_0 for scaling. Then, we prefer to make an assumption on the strength of the stratification later.

We obtain the scaling bound for the vertical velocity using density equation as (Franzke et al., 2019),

$$W \le \frac{g\Gamma}{\rho_0}\,\frac{V}{N_0^2\,L} = \frac{g\Gamma}{\rho_0}\,\frac{L}{V}\,\alpha^2\,\mathrm{Fr}^2\,, \tag{1.45}$$

on which the density is taken as three-dimensional advected quantity depending on the stratification profile by the Froude number, Fr, defined in (1.48). The other important relation is obtained by the horizontal momentum equation and hydrostatic approximation via

Assumption F1. The vertical momentum equation is assumed nearly hydrostatic, i.e.,

$$\Omega_z L U\rho_0 \sim P \sim g\Gamma H\,. \tag{1.46}$$

1.3.2 Characteristic parameters

The fluid motion is characterized with two important features, rotation and stratification. The smallness assumption on some dimensionless parameters related with them gives the time separation between two different motions. Firstly, the rotation is characterized by the Rossby number. It is the ratio between the frequency of the flow to the non-zero component of the Coriolis vector,

$$\mathrm{Ro} = \frac{f_{flow}}{f_{Coriolis}} = \frac{U/L}{f} = \frac{U}{fL}\,. \tag{1.47}$$

Ro $\ll 1$ states the velocity of the fluid is smaller than the velocity due to the rotation of the Earth, and that the rotation is strong. Therefore, the flow has a slow motion in such a regime. Then, the smallness assumption on the Rossby number is more suitable for large-scale motions on which the rotation is dominant. However, this case is underestimated for low latitudes because the Coriolis vector is parallel to the surface of the Earth there. This special case will be taken up later.

When the Rossby number is large, i.e, Ro ≈ 1 or Ro $\gg 1$, the scales are not easily separated. Then, fast and slow motions are coupled and so the emission of inertia-gravity wave is allowed (Zeitlin, 2008).

The other important parameter is the Froude number which is the ratio between the fluid velocity and the characteristic wave speed. It is mathematically given as:

$$\mathrm{Fr} = \frac{U}{c} \tag{1.48}$$

where $c = c_1 = \frac{N_0 H}{\pi}$ is the first internal gravity wave speed for Boussinesq flows. It identifies the stratification of the flow with buoyancy frequency given in (1.44). When it is small, i.e, Fr $\ll 1$ with positive buoyancy frequency, $N_0 > 0$, the stratification is strong. If $\partial_z \rho < 0$, then the flow is stably stratified. It implies that the density decreases with the height. Otherwise, $\partial_z \rho > 0$, the density increases with the height and the flow is unstably stratified (Majda, 2003). If Fr $\gg 1$, the stratification is weak.

With these two dimensionless parameters, one can nearly eliminate the fast motion from the slow motion which gives the dynamical core of the flow. Besides, the model becomes simpler.

The regime of the fluid can be defined considering the relationship between Fr and Ro. To understand it, we get help from another parameter called Burger number,

$$\mathrm{Bu} = \left(\frac{L_d}{L}\right)^2 \tag{1.49}$$

where $L_d = c/f$ is the Rossby deformation radius. This horizontal length scale is frequently used in atmosphere and oceanic dynamics to indicate that the efecct of the rotation is important as much as the buoyancy (Gill, 1982). This ratio can be written using the Froude number and the Rossby number (Franzke et al., 2019),

$$\mathrm{Bu} = \frac{\mathrm{Ro}^2}{\mathrm{Fr}^2}\,. \tag{1.50}$$

For example, when we assume $\mathrm{Fr} \sim \mathrm{Ro}$, then, the resulting Burger number becomes nearly unity $\mathrm{Bu} \sim 1$. That is known as quasi-geostrophic regime. In this regime, the length scale is nearly equal to the deformation radius, $L \sim L_d$.

Another important regime appears when $\mathrm{Fr}^2 = \mathrm{Ro}$, and so $\mathrm{Bu} \sim \mathrm{Ro}$. It is called semi-geostrophic regime. Generally, it is used for specific regions like mid-latitude where the non-zero vertical component of the Coriolis vector provides non-zero Ro. This approximation works for larger scale regimes than quasi-geostrophic one because the length scale is larger than the deformation radius, $L > L_d$.

1.3.3 The equations of motions on the f-plane

We assume that components of the Coriolis vector are constant in the mid-latitude. It is non-dimensionalized as

$$2\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ c \\ s \end{pmatrix}, \tag{1.51}$$

where $c = \cos\phi$ and $s = \sin\phi$ for a specified latitude ϕ. Then, we make following assumptions:

Assumption F2. The horizontal flow is assumed isotropic,

$$U \sim V\,. \tag{1.52}$$

Assumption F3. The fluid parcel is advected on the horizontal time scale,

$$\frac{1}{T} \sim \frac{U}{L}\,. \tag{1.53}$$

We apply the scaling parameters to the Boussinesq equations (1.43). Then, non-dimensional equations of motion read

$$\mathrm{Ro}\left(\partial_t u_1 + (u\cdot\nabla)u_1\right) + \mathrm{Fr}^2 u_3\,\partial_z u_1 - u_2 + \alpha c\,\frac{\mathrm{Fr}^2}{\mathrm{Ro}}u_3 = -\partial_x p\,, \tag{1.54a}$$

$$\mathrm{Ro}\left(\partial_t u_2 + (u\cdot\nabla)u_2\right) + \mathrm{Fr}^2\, u_3\, \partial_z u_2 + u_1 = -\partial_y p\,, \tag{1.54b}$$

$$\alpha^2\,\mathrm{Fr}^2\left(\partial_t u_3 + (u\cdot\nabla)u_3\right) + \alpha^2\frac{\mathrm{Fr}^4}{\mathrm{Ro}} u_3\,\partial_z u_3 - \alpha\frac{c}{s}u_1 = -\partial_z p - \rho\,, \tag{1.54c}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0\,, \tag{1.54d}$$

$$D_t\rho = 0\,. \tag{1.54e}$$

Assumption F4. Under the assumption of semi-geostrophic scaling on the f-plane, the relationship between characteristic parameters is set as

$$\mathrm{Ro} \sim \mathrm{Fr}^2\,. \tag{1.55}$$

This scaling provides strong non-linearity, i.e, three-dimensional advected fluid. The non-linearity is weaker on the quasi-geostrophic models since the fluid is horizontally advected.

Assumption F5. The fluid is assumed non-hydrostatic,

$$\alpha = 1\,. \tag{1.56}$$

Assumption F6. Rossby number is the small parameter in the model to separate the fast and slow motions,

$$\varepsilon = Ro = \frac{U}{sL}\,. \tag{1.57}$$

Altogether, Boussinesq equations in the semi-geostrophic scaling on the f-plane reads

$$\begin{aligned} \varepsilon\, D_t\boldsymbol{u} + \boldsymbol{\Omega}\times\boldsymbol{u} &= -\boldsymbol{\nabla} p - \rho\boldsymbol{k}\,,\\ \boldsymbol{\nabla}\cdot\boldsymbol{u} &= 0\,,\\ D_t\rho &= 0\,. \end{aligned} \tag{1.58}$$

1.4 Scaling near the equator

The vertical component of the Coriolis vector becomes zero at the equator, and so this does not allow to have a clear scale separation between the fast and the slow motions as we have in the mid-latitude. This degeneracy indicates a new mathematical phenomena around this region.

Yoshida (1959) introduced that equatorial region acts as a wave-guide. Later, Matsuno (1966) showed that some of equatorial waves, which have low frequency like Kelvin waves, are trapped in this region. Thus, the standard isotropic scaling is not proper on this region. The domain needs to cover the circumference of the Earth in the zonal direction as the dynamics suggests. Then, the anisotropy between horizontal length scales on which the length scale in the zonal direction is taken very long as required there.

Long-wave approximation was firstly defined by Pedlosky (1965). In that work, this approximation was used for the mid-latitude. It was introduced for an equatorial oceanic model by Cane and Sarachik (1977) using anisotropy. By this approximation, the geostrophic balance is obtained only in the meridional direction. In addition, waves on this dynamics have low frequencies. Thus, it is known as low frequency and meridional geostrophy too (Boyd, 2018). A model with this approximation provides only Rossby wave and low frequency Kelvin waves around the equator. It nearly does not contain any inertia-gravity wave and mixed-Rossby gravity wave, which requires non-zero meridional velocity. This approximation with the linear theory is often preferred to model El-Nino phenomena for the equatorial ocean dynamics (Gill, 1982; Majda, 2003).

Long-wave scaling was used by Majda (2003) and Chan and Shepherd (2013) in the concept of the equatorial balance. We state options on the long-wave scaling considering differences on these reference works in following parts.

1.4.1 The equations of motion on the β-plane

In this part, we show the scaling for the equator in the long-wave scaling to filter out the fast waves in the asymptotic limits. Whole derivations are done for the equatorial $\beta-$plane.

We assume *equatorial long-wave scaling* where

$$\delta = \frac{L}{X} \ll 1 \,. \tag{1.59}$$

Moreover, the vertical aspect ratio is also small, i.e.,

$$\alpha = \frac{H}{L} \ll 1\,. \tag{1.60}$$

Then, the rescaled divergence condition reads

$$0 = \frac{U}{X}\,\partial_x u_1 + \frac{V}{L}\,\partial_y u_2 + \frac{W}{H}\,\partial_z u_3 = \frac{V}{L}\left(\frac{\delta U}{V}\,\partial_x u_1 + \partial_y u_2 + \frac{W}{\alpha V}\,\partial_z u_3\right). \tag{1.61}$$

Thus, to have a sensible leading-order balance, the meridional and vertical terms must be in balance, i.e., $W = \alpha V$. In principle, the leading-order balance could also be between the zonal and the vertical contribution, with the meridional component coming in at the lower order. However, this would require an extreme smallness assumption on the meridional velocity. Alternatively, it might be between the first and the second terms on (1.61). Then, the parameter

$$\gamma = \frac{\delta U}{V} \tag{1.62}$$

may either be 1, which corresponds to the long-wave scaling of Majda (2003), or small, e.g. $\gamma = \delta$, which corresponds to the long-wave scaling of Chan and Shepherd (2013). To cover both cases, we define the rescaled gradient operator

$$\boldsymbol{\nabla}_\gamma = (\gamma\,\partial_x, \partial_y, \partial_z)\,. \tag{1.63}$$

The Coriolis vector on the β-plane is

$$2\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ f_0 \\ \beta y \end{pmatrix}, \tag{1.64}$$

where $f_0 = 2|\boldsymbol{\Omega}|$ by (1.22). With these provisions, inserting $u = U\hat{u}$, $x = X\,\hat{x}$, etc., into the Euler–Boussinesq equations (1.2) and dropping the "hats", we obtain

$$\frac{U}{T}\,\partial_t u_1 + \frac{UV}{L}\,\boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1 - \beta LV\,y\,u_2 + \alpha\,f_0\,V\,u_3 = -\delta\,\frac{P}{L\rho_0}\,\partial_x p\,, \tag{1.65a}$$

$$\frac{V}{T}\,\partial_t u_2 + \frac{V^2}{L}\,\boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2 + \beta L U\, y\, u_1 = -\frac{P}{L\rho_0}\,\partial_y p\,, \tag{1.65b}$$

$$\frac{\alpha V}{T}\,\partial_t u_3 + \frac{\alpha V^2}{L}\,\boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_3 - f_0\, U\, u_1 = -\frac{P}{\alpha L\rho_0}\,\partial_z p - \frac{g\Gamma}{\rho_0}\,\rho\,, \tag{1.65c}$$

$$\frac{V}{L}\,\boldsymbol{\nabla}_\gamma\cdot\boldsymbol{u} = 0\,, \tag{1.65d}$$

$$\frac{\Gamma}{T}\,\partial_t\rho + \frac{V\Gamma}{L}\,(\gamma\, u_1\,\partial_x\rho + u_2\,\partial_y\rho) + \frac{\alpha V N_0^2\,\rho_0}{g}\,u_3\,\partial_z\rho = 0\,. \tag{1.65e}$$

Here, N_0 is the buoyancy frequency given in (1.44). We set

$$\kappa = \frac{\alpha f_0}{\beta L}\,. \tag{1.66}$$

Then, we note that the rescaled Coriolis vector potential reads

$$\boldsymbol{R} = \beta L^2\,(\kappa\, z - \tfrac{1}{2}\, y^2)\,\boldsymbol{i}\,, \tag{1.67}$$

and correspondingly,

$$2\boldsymbol{\Omega} = \beta L \begin{pmatrix} 0 \\ \kappa/\alpha \\ y \end{pmatrix}\,. \tag{1.68}$$

We now make a few basic assumptions.

Assumption E1. Pressure gradient and "traditional" Coriolis forces balance to the leading order in the zonal momentum equation (1.65a), i.e.,

$$\beta L V = \delta\,\frac{P}{L\,\rho_0}\,. \tag{1.69}$$

Assumption E2. We assume a "meridionally advective scaling" where

$$\frac{1}{T} = \frac{V}{L}\,. \tag{1.70}$$

Assumption E3. We assume that the vertical momentum equation is hydrostatic and so

$$\beta L^2 U\,\rho_0 \sim P \sim g\Gamma H\,. \tag{1.71}$$

Inserting Assumptions E1-E3 into the momentum equations, we obtain

$$\frac{\gamma}{\delta\beta LT}\left(\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1\right) - y\,u_2 + \kappa\,u_3 = -\partial_x p\,, \tag{1.72a}$$

$$\frac{\delta}{\beta LT}\left(\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2\right) + \gamma\,y\,u_1 = -\partial_y p\,, \tag{1.72b}$$

$$\frac{\delta\,\alpha^2}{\beta LT}\left(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma w\right) - \gamma\,\kappa\,u_1 = -\partial_z p - \rho\,. \tag{1.72c}$$

The rescaled mass conservation equation reads

$$\partial_t\rho + (\gamma\,u_1\,\partial_x + u_2\,\partial_y)\rho + \frac{\delta}{\beta LT}\frac{1}{\mathrm{Fr}^2}\,u_3\,\partial_z\rho = 0\,, \tag{1.72d}$$

where the Foude number is defined using the meridional velocity

$$\mathrm{Fr} = \frac{V}{N_0 H}\,. \tag{1.73}$$

Assumption E4. We assume

$$\mathrm{Fr}^2 = \frac{\delta}{\beta LT} \tag{1.74}$$

to have three-dimensional advective density in (1.72d). Basicaly, we use the circumference of the Earth for the zonal length scale and we know the value of β by (1.23). The zonal wind is taken $5m/s$ as observations suggest (Gill, 1982, (Fig 9.3)) for the value of T. Then, the ratio roughly yields Fr $\approx$ 0.05. This implies that the fluid is strongly stratified.

Alternatively, the factor in front of the vertical density advection might be assumed small, and so

$$\partial_t\rho + u\cdot\nabla_\gamma\rho = 0\,. \tag{1.75}$$

With this horizontally advected fluid, we have again small Fr, and so the fluid is still strongly stratified. In addition to this, when γ is taken small, and we have simpler case

$$\partial_t\rho + u_2\,\partial_y\rho = 0\,, \tag{1.76}$$

which gives one-dimensional transport equation.

1.4.2 Cases on the equatorial scaling

Now to get a sensible leading-order balance for the meridional momentum equation (1.72b), we have the following options.

Case 1. The meridional pressure gradient balances the Coriolis term, i.e. $\gamma = 1$, and we allow propagation of zonal linear waves at leading order, i.e.,

$$\frac{1}{\delta \beta L T} = 1 . \tag{1.77}$$

Then, the momentum equations read

$$\partial_t u_1 + \boldsymbol{u} \cdot \boldsymbol{\nabla} u_1 - y\, u_2 + \kappa\, u_3 = -\partial_x p\,, \tag{1.78a}$$

$$\delta^2 \left(\partial_t u_2 + \boldsymbol{u} \cdot \boldsymbol{\nabla} u_2 \right) + y\, u_1 = -\partial_y p\,, \tag{1.78b}$$

$$\alpha^2\, \delta^2 \left(\partial_t u_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} u_3 \right) - \kappa\, u_1 = -\partial_z p - \rho\,, \tag{1.78c}$$

and

$$\partial_t \rho + u \cdot \nabla \rho + \frac{\delta^2}{\mathrm{Fr}^2}\, u_3\, \partial_z \rho = 0 . \tag{1.79}$$

Case 1 corresponds to the non-linear equatorial long-wave scaling used by Majda (2003, Section 9.3), there for the shallow water equations. We refer a short analysis for his model to Appendix A which shows the derivation of the kinematic balance relation. Apparently, there is a problem on the solvability. Then, we expect to similar issues for our Boussinesq model with the same scaling. Therefore, we skip this case and leave it as an open question in the context of existence and the well-posedness of the model with this scaling as Majda's did.

Case 2. The meridional pressure gradient balances the Coriolis term, i.e. $\gamma = 1$, and we allow the propagation of zonal linear waves at higher order, i.e.,

$$\frac{1}{\delta \beta L T} = \hat{\varepsilon} \ll 1 . \tag{1.80}$$

In this case, we obtain

$$\hat{\varepsilon} \left(\partial_t u_1 + \boldsymbol{u} \cdot \boldsymbol{\nabla} u_1 \right) - y\, u_2 + \kappa\, u_3 = -\partial_x p\,, \tag{1.81a}$$

$$\hat{\varepsilon}\, \delta^2 \left(\partial_t u_2 + \boldsymbol{u} \cdot \boldsymbol{\nabla} u_2 \right) + y\, u_1 = -\partial_y p\,, \tag{1.81b}$$

$$\hat{\varepsilon}\,\delta^2\,\alpha^2\,(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla} u_3) - \kappa\,u_1 = -\partial_z p - \rho\,, \tag{1.81c}$$

and

$$\partial_t\rho + u\cdot\nabla\rho + \hat{\varepsilon}\,\frac{\delta^2}{\mathrm{Fr}^2}\,u_3\,\partial_z\rho = 0\,. \tag{1.82}$$

So this is more extreme scaling than Case 1 and should lead to simpler dynamics. Therefore, we follow this option on the derivation of the equatorial balance model.

Alternatively, we have one more option which might be interesting for the theoretical perspective due to the necessity of Lagrange multiplier for the incompressibility, (Majda, 2003).

Case 3. The meridional pressure gradient balances the inertial forces, i.e.

$$\frac{\delta}{\beta LT} = 1\,. \tag{1.83}$$

Then, to get a sensible leading-order balance in the zonal momentum equation (1.72a), we require that $\gamma = \delta^2$ or even smaller. In this case, we obtain

$$\frac{\gamma}{\delta^2}\,\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1 - y\,u_2 + \kappa\,u_3 = -\partial_x p\,, \tag{1.84a}$$

$$\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2 + \gamma\,y\,u_1 = -\partial_y p\,, \tag{1.84b}$$

$$\alpha^2\,(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_3) - \gamma\,\kappa\,u_1 = -\partial_z p - \rho\,, \tag{1.84c}$$

and

$$\partial_t\rho + u\cdot\nabla_\gamma\rho + \frac{1}{\mathrm{Fr}^2}\,u_3\,\partial_z\rho = 0\,. \tag{1.85}$$

In this regime, the rotation is not the only dominant factor. Then, we do not prefer to use it, but it might be an interesting application because the incompressibility condition changes. It works on to the meridional-vertical velocity.

1.4.3 Discussion of alternatives

In this section, we collect comments on alternatives and on the rationales for not choosing them.

The alternative to the meridionally advective scaling, Assumption E2, would be a "weakly nonlinear scaling" where

$$\frac{\varepsilon}{T} = \frac{V}{L}\,, \tag{1.86}$$

for a small number ε. This corresponds to the scaling used by Chan and Shepherd (2013) with $\gamma = \delta$. We observe that this scaling breaks the geometry, but it is possible to do formal variational asymptotics with weakly non-linearly scaled Lin constraints. Similarly, Case 3 can be tailored in this way.

Another alternative is that the zonal inertia term balances the Coriolis force, i.e.,

$$\frac{U}{T} = \beta L V\,. \tag{1.87}$$

In that case, the only reasonable leading order balance in the meridional momentum equation is that between pressure gradient and Coriolis term; if it wasn't, the leading order horizontal dynamics would be fully pressure-less, which is unphysical. I.e., we also impose

$$\beta L U = \frac{P}{L\rho_0}\,. \tag{1.88}$$

Keeping, once again, the "meridionally advective scaling" (1.70) and near-hydrostaticity (1.46), the resulting momentum equations read

$$\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1 - y\,u_2 + \kappa\,u_3 = -\gamma\,\partial_x p\,, \tag{1.89a}$$

$$\frac{\varepsilon^2}{\gamma^2}\left(\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2\right) + y\,u_1 = -\partial_y p\,, \tag{1.89b}$$

$$\frac{\alpha^2\,\varepsilon^2}{\gamma^2}\left(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_3\right) - \kappa\,u_1 = -\partial_z p - \rho\,, \tag{1.89c}$$

and

$$\partial_t\rho + u\cdot\nabla_\gamma\rho + \frac{\varepsilon}{\gamma}\frac{c^2}{V^2}\frac{1}{\beta L T}\,u_3\,\partial_z\rho = 0\,. \tag{1.90}$$

When $\gamma = 1$, we are back to Case 1, but here, γ may also be smaller, so it looks like a proper generalization of Case 1. However, for $\gamma \ll 1$, the leading-order dynamics is Burgers with rotation in the zonal direction, which is likely unphysical. Also, a variational "weak pressure" approximation breaks the geometry.

Finally, for the sake of discussion, let us assume that the zonal leading-order balance is between the traditional and the non-traditional Coriolis force, i.e.,

$$\beta L = \alpha f_0 \,. \tag{1.91}$$

In the meriodional momentum equation, we still need to assume that (1.88) holds true, for the same reason as in the previous case. Keeping, once again, the "meridionally advective scaling" (1.70) and near-hydrostaticity (1.46), the resulting momentum equations read

$$\frac{\gamma}{\varepsilon}\,\frac{1}{\beta LT}\,(\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1) - y\,u_2 + u_3 = -\gamma\,\partial_x p\,, \tag{1.92a}$$

$$\frac{\varepsilon}{\gamma}\,\frac{1}{\beta LT}\,(\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2) + y\,u_1 = -\partial_y p\,, \tag{1.92b}$$

$$\frac{\varepsilon}{\gamma}\,\frac{\alpha^2}{\beta LT}\,(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_3) - u_1 = -\partial_z p - \rho\,, \tag{1.92c}$$

and

$$\partial_t\rho + u\cdot\nabla_\gamma\rho + \frac{\varepsilon}{\gamma}\,\frac{c^2}{V^2}\,\frac{1}{\beta LT}\,u_3\,\partial_z\rho = 0\,. \tag{1.93}$$

Here, we have quite a bit freedom to play. For example, we might set $\gamma = \varepsilon$ and

$$\frac{1}{\beta LT} = \varepsilon\,, \tag{1.94}$$

to obtain

$$\varepsilon\,(\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_1) - y\,u_2 + u_3 = -\varepsilon\,\partial_x p\,, \tag{1.95a}$$

$$\varepsilon\,(\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_2) + y\,u_1 = -\partial_y p\,, \tag{1.95b}$$

$$\varepsilon\,\alpha^2\,(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}_\gamma u_3) - u_1 = -\partial_z p - \rho\,, \tag{1.95c}$$

and

$$\partial_t\rho + u\cdot\nabla_\gamma\rho + \varepsilon\frac{c^2}{V^2}\,u_3\,\partial_z\rho = 0\,. \tag{1.96}$$

Chapter 2

Balance models via the Hamilton principle

In the fluid dynamics context, Hamiltonian balance models have the advantage of the absence of the dissipation and the conservation of circulation (or equivalently conservation of potential vorticity). Then, we show how a balance model is constructed via the Hamilton's principle in this chapter. This construction not only brings the equations of motion, but also maintains analogs of the conservation laws.

2.1 Variational principle

In this part, it is shown how to obtain the equations of fluid motion by a variational principle. To do this, we explain Lagrangian coordinates first because the variational principle is most naturally formulated in this coordinate system. The main difference from the Eulerian coordinates is that Lagrangian coordinates refer to a fixed fluid parcel, while Eulerian coordinates refer to a fixed observer. In the following, we write $\boldsymbol{x}$ to denote the Eulerian or fixed observer position and $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t)$ to denote the corresponding velocity. Similarly, we write $\boldsymbol{a}$ to denote the Lagrangian label of the parcel initially at location $\boldsymbol{a}$. The time dependent map from the label space to positions is called the flow map $\boldsymbol{\eta}$. Then, $\boldsymbol{x} = \boldsymbol{\eta}(\boldsymbol{a}, t)$ is the position of the parcel originating at $\boldsymbol{a}$ after time t. At time $t = 0$, $\boldsymbol{\eta}$ is the identity map, i.e., $\boldsymbol{a} = \boldsymbol{\eta}(\boldsymbol{a}, 0)$. Figure 2.1 provides the mapping from the label space to the inertial space.

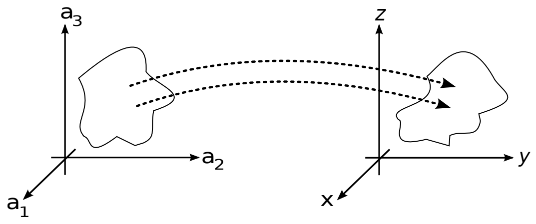


FIGURE 2.1: The mapping from the label space to the inertial space, inspired from Pozrikidis (2011).

We assume that $\boldsymbol{\eta}$ is differentiable with respect to label $\boldsymbol{a}$ and time t. Then, the Lagrangian velocity, the velocity of the parcel labeled $\boldsymbol{a}$, is related to the Eulerian velocity field as

$$\partial_t \boldsymbol{\eta}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{\eta}(\boldsymbol{a}, t), t)\,. \tag{2.1}$$

It can be abbreviated using the symbol "$\circ$" which denotes the composition of maps,

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}\,. \tag{2.2}$$

The variations for the flow map, which are in the Lagrangian setting, can be associated with an Eulerian vector field via

$$\delta\boldsymbol{\eta} = \boldsymbol{v} \circ \boldsymbol{\eta}\,, \tag{2.3}$$

where the variation vector field $\boldsymbol{v}$ associated with $\boldsymbol{\eta}$ is divergence free because $\boldsymbol{\eta}$ must be volume preserving.

2.1.1 Lin Constraints

Generally speaking, it is easier to work on Eulerian coordinates. Therefore, we need to find an expression for the variations of $\boldsymbol{u}$ in terms of variations of $\boldsymbol{\eta}$. It allows to switch between Eulerian and Lagrangian framework in the variational principle. It is obtained by cross differentiation of (2.2) and (2.3). Firstly, the variation is applied to the abstract velocity given in (2.2),

$$\begin{aligned}
\delta\dot{\boldsymbol{\eta}} &= \delta(\boldsymbol{u}\circ\boldsymbol{\eta}) \\
&= \delta\boldsymbol{u}\circ\boldsymbol{\eta} + (\nabla\boldsymbol{u})\circ\boldsymbol{\eta}\,\delta\boldsymbol{\eta} \\
&= \delta\boldsymbol{u}\circ\boldsymbol{\eta} + (\nabla\boldsymbol{u})\circ\boldsymbol{\eta}\,\boldsymbol{v}\circ\boldsymbol{\eta} \\
&= \big(\delta\boldsymbol{u} + \boldsymbol{v}\cdot\nabla\boldsymbol{u}\big)\circ\boldsymbol{\eta}\,.
\end{aligned} \tag{2.4}$$

Then, the time derivative of (2.3) is taken

$$\begin{aligned}
\delta\dot{\boldsymbol{\eta}} &= \partial_t(\boldsymbol{v}\circ\boldsymbol{\eta}) \\
&= \partial_t\boldsymbol{v}\circ\boldsymbol{\eta} + (\nabla\boldsymbol{v})\circ\boldsymbol{\eta}\,\dot{\boldsymbol{\eta}} \\
&= \dot{\boldsymbol{v}}\circ\boldsymbol{\eta} + (\nabla\boldsymbol{v})\circ\boldsymbol{\eta}\,\boldsymbol{u}\circ\boldsymbol{\eta} \\
&= \big(\dot{\boldsymbol{v}} + \boldsymbol{u}\cdot\nabla\boldsymbol{v}\big)\circ\boldsymbol{\eta}\,.
\end{aligned} \tag{2.5}$$

We combine these two equations into one, and we have

$$\delta\boldsymbol{u} = \dot{\boldsymbol{v}} + \boldsymbol{u}\cdot\nabla\boldsymbol{v} - \boldsymbol{v}\cdot\nabla\boldsymbol{u}\,, \tag{2.6}$$

where $\boldsymbol{v}$ is the variation vector field. The advection of the density equation (1.2c) can be written in an equivalent form

$$\rho\circ\boldsymbol{\eta} = \rho_{in}\,, \tag{2.7}$$

where ρ_{in} is the initial density distribution . Then, we have

$$\begin{aligned}
\delta(\rho\circ\boldsymbol{\eta}) &= \delta\rho\circ\boldsymbol{\eta} + (\nabla\rho)\circ\boldsymbol{\eta}\,\delta\rho \\
&= \delta\rho\circ\boldsymbol{\eta} + (\nabla\rho)\circ\boldsymbol{\eta}\,\boldsymbol{v}\circ\boldsymbol{\eta} \\
&= (\delta\rho + \boldsymbol{v}\cdot\nabla\rho)\circ\boldsymbol{\eta}\,,
\end{aligned} \tag{2.8}$$

and so

$$\delta\rho + \boldsymbol{v}\cdot\nabla\rho = 0\,, \tag{2.9}$$

a conservation law for the density. (2.6) and (2.9) are known as "Lin constraints" which were firstly shown by Bretherton (1970).

2.1.2 The equations of motion

The Lagrangian for the rotating Boussinesq equations is given by Franzke et al. (2019)

$$
\begin{aligned}
L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_{in}) &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \frac{1}{2}\, |\dot{\boldsymbol{\eta}}|^2 - \frac{g}{\rho_0}\, \rho_{in} \eta_3 \, d\boldsymbol{a} \\
&= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot (\boldsymbol{u} \circ \boldsymbol{\eta}) + \frac{1}{2}\, |\boldsymbol{u} \circ \boldsymbol{\eta}||\boldsymbol{u} \circ \boldsymbol{\eta}| - \frac{g}{\rho_0}\, (\rho \circ \boldsymbol{\eta})(z \circ \boldsymbol{\eta}) \, d\boldsymbol{a} \\
&= \int_{\mathcal{D}} \big(\boldsymbol{R} \cdot \boldsymbol{u} + \frac{1}{2}\, |\boldsymbol{u}| \cdot |\boldsymbol{u}| - \frac{g}{\rho_0}\, \rho z\big) \circ \boldsymbol{\eta} \, d\boldsymbol{a} \\
&= \int_{\mathcal{D}} \big(\boldsymbol{R} \cdot \boldsymbol{u} + \frac{1}{2}\, |\boldsymbol{u}|^2 - \frac{g}{\rho_0}\, \rho z\big) \, d\boldsymbol{x} \equiv l(\boldsymbol{u}, \rho)\,,
\end{aligned}
\tag{2.10}
$$

where the vector potential of the Coriolis vector $\boldsymbol{R}$ is defined in (1.26). The Lagrangian is written here in the dimensional form. It is given in the Lagrangian perspective first. Then, the "reduced Lagrangian" $l(\boldsymbol{u}, \rho)$, the equivalent form from the Eulerian perspective, is obtained via (2.2), the change of the variables formula, and (2.7). We note that the Lagrangian includes the rotational contribution here, and so it is not only the difference between kinetic and potential energies.

The action S is defined as the integral with fixed t_1 and t_2 of the Lagrangian,

$$
S = \int_{t_1}^{t_2} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}, t) \, dt\,. \tag{2.11}
$$

Critical values of (2.11) are given by

$$
\delta S = 0\,. \tag{2.12}
$$

It means that the action is stationary. This is known as Hamiltonian's principle. We insert the Lagrangian defined in (2.10) into the action S. Then, we take the variation as follows

$$
\delta S = \int_{t_1}^{t_2} \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \delta \dot{\boldsymbol{\eta}} + \boldsymbol{\nabla} \boldsymbol{R} \circ \boldsymbol{\eta}\, \delta \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \dot{\boldsymbol{\eta}} \cdot \delta \dot{\boldsymbol{\eta}} - \frac{g}{\rho_0}\, \delta(\rho_{in} \eta_3) \, d\boldsymbol{a} \, dt
$$

$$
\begin{aligned}
&= -\int_{t_1}^{t_2}\int_{\mathcal{D}}\bigg(-\nabla \boldsymbol{R}\circ\boldsymbol{\eta}\,\dot{\boldsymbol{\eta}}\cdot\delta\boldsymbol{\eta} + \delta\boldsymbol{\eta}\cdot\nabla\boldsymbol{R}^{\mathsf{T}}\circ\boldsymbol{\eta}\cdot\dot{\boldsymbol{\eta}} + \ddot{\boldsymbol{\eta}}\cdot\delta\boldsymbol{\eta} + \nabla\boldsymbol{u}\circ\boldsymbol{\eta}\,\dot{\boldsymbol{\eta}}\cdot\delta\boldsymbol{\eta} \\
&\qquad\qquad + \frac{g}{\rho_0}\,\rho\circ\boldsymbol{\eta}\,\delta\eta_3\bigg)\,d\boldsymbol{a}\,dt \\
&= -\int_{t_1}^{t_2}\int_{\mathcal{D}}\bigg((\nabla\boldsymbol{R}^{\mathsf{T}} - \nabla\boldsymbol{R})\boldsymbol{u}\cdot\boldsymbol{v} + \partial_t\boldsymbol{u}\cdot\boldsymbol{v} + \nabla\boldsymbol{u}\boldsymbol{u}\cdot\boldsymbol{v} + \frac{g}{\rho_0}\,\rho\,\boldsymbol{k}\cdot\boldsymbol{v}\bigg)\circ\boldsymbol{\eta}\,d\boldsymbol{a}\,dt \\
&= -\int_{t_1}^{t_2}\int_{\mathcal{D}}\bigg(\partial_t\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \big(\nabla\times\boldsymbol{R}\big)\times\boldsymbol{u} + \frac{g}{\rho_0}\,\rho\,\boldsymbol{k}\bigg)\cdot\boldsymbol{v}\,d\boldsymbol{x}\,dt\,.
\end{aligned}
\tag{2.13}
$$

On this derivation, integration by parts is applied on the second and third lines. Terms in the parentheses on the last line equal to the gradient due to Hodge decomposition, i.e,

$$
\partial_t\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \big(\nabla\times\boldsymbol{R}\big)\times\boldsymbol{u} + \frac{g}{\rho_0}\,\rho\,\boldsymbol{k} = \nabla\phi \tag{2.14}
$$

with the potential

$$
\phi = -\frac{p}{\rho_0} - \frac{1}{2}|\boldsymbol{u}|^2\,, \tag{2.15}
$$

where p denotes the pressure and the vector identity

$$
(\nabla\times\boldsymbol{u})\times\boldsymbol{u} = \boldsymbol{u}\cdot\nabla\boldsymbol{u} - \frac{1}{2}\nabla|\boldsymbol{u}|^2 \tag{2.16}
$$

is used. Therefore, momentum equations are recovered and we have

$$
\partial_t\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \big(\nabla\times\boldsymbol{R}\big)\times\boldsymbol{u} = -\frac{1}{\rho_0}\nabla p - \frac{g}{\rho_0}\,\rho\boldsymbol{k}\,. \tag{2.17}
$$

They are now called Euler-Poincaré equations via Lagrangian l (Holm, Marsden, and Ratiu, 1998). The corresponding theorem is given in Theorem 3.2.1 using the variational principle on the action S, and it states the equivalence between the full and the reduced variational principle.

2.2 Conservation Laws

Working with the variational principle brings the conserved quantites by the symmetry property. Noether has stated a general theorem that every (Lie) symmetry of the Lagrangian in the action gives rise to a conservation law (Noether, 1918; Kosmann-Schwarzbach, 2011). In this section, we have a brief look on the conserved quantites in our particular setting.

First of all, the model obtained via Lagrangian is invariant under time translation. The corresponding Hamiltonian for Boussinesq equations is

$$\begin{aligned} H &= \int_{\mathcal{D}} \frac{\delta l}{\delta \boldsymbol{u}} \cdot \boldsymbol{u}\, d\boldsymbol{x} - l(\boldsymbol{u}, \rho) \\ &= \int_{\mathcal{D}} \frac{1}{2}\, |\boldsymbol{u}|^2 + \frac{g}{\rho_0}\, \rho\, z\, d\boldsymbol{x}\,, \end{aligned} \tag{2.18}$$

where l is given on the last line of (2.10). By chain rule, we define $\delta f(t) = \dot{f}(t)\, \delta t$ for an arbitrary time dependent function f. We take the variation of the Lagrangian using (2.18)

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \frac{d}{dt} \left(\int_{\mathcal{D}} \frac{\delta l}{\delta \boldsymbol{u}} \cdot \boldsymbol{u}\, d\boldsymbol{x} - H \right) \delta t\, dt \\ &= \int_{t_1}^{t_2} \int_{\mathcal{D}} \left(\frac{d(\boldsymbol{R} + \boldsymbol{u})}{dt} \cdot \boldsymbol{u} + (\boldsymbol{R} + \boldsymbol{u}) \cdot \frac{d\boldsymbol{u}}{dt} \right) d\boldsymbol{x}\, \delta t\, dt - \int_{t_1}^{t_2} \frac{dH}{dt}\, \delta t\, dt \\ &= - \int_{t_1}^{t_2} \frac{dH}{dt}\, \delta t\, dt\,. \end{aligned} \tag{2.19}$$

We apply the integration by parts in time to the first integral in the second line and it vanishes. Therefore, we have only variation of the Hamiltonian. It implies

$$\frac{dH}{dt} = 0\,, \tag{2.20}$$

and so the energy is conserved as a result of the Hamilton's principle in the case $\delta S = 0$.

Next, we work on the particle relabeling symmetry because it provides the potential vorticity as a conserved quantity. In this case, the Lagrangian is invariant to translation in compositions of the flow map with arbitrary volume and domain preserving maps. Geometrically, Euler-Poincaré equation reads

$$(\partial_t + \mathcal{L}_u)\, \boldsymbol{m} + \frac{\delta l}{\delta \rho}\, \mathrm{d}\rho = \mathrm{d}\phi\,, \tag{2.21}$$

where d represents the exterior derivative and $\mathcal{L}_u$ is the Lie derivative in the direction of $\boldsymbol{u}$. First, we write the exterior derivative of (2.21)

$$(\partial_t + \mathcal{L}_u)\,\mathrm{d}\boldsymbol{m} = \mathrm{d}\big(-\tfrac{\delta l}{\delta \rho}\,\mathrm{d}\rho + \mathrm{d}\phi\big) = -\mathrm{d}\rho \wedge \mathrm{d}z\,. \tag{2.22}$$

Similarly, the density satisfies

$$(\partial_t + \mathcal{L}_u)\,\rho = 0\,, \tag{2.23}$$

and so the exterior derivative of it

$$(\partial_t + \mathcal{L}_u)\,\mathrm{d}\rho = 0\,. \tag{2.24}$$

Then, the exterior product of (2.22) and (2.24) is by product rule for Lie derivatives

$$(\partial_t + \mathcal{L}_u)\,\mathrm{d}\boldsymbol{m} \wedge \mathrm{d}\rho = 0\,, \tag{2.25}$$

or in the vector calculus notation

$$\partial_t q + \boldsymbol{u} \cdot \boldsymbol{\nabla} q = 0\,. \tag{2.26}$$

Here q is the conserved quantity on fluid particles and it is called potential vorticity i.e,

$$q = (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla}\rho\,. \tag{2.27}$$

Invariance of the Lagrangian under particle relabeling provides the potential vorticity as a conserved quantity. It is known as Ertel's theorem, too (Ertel, 1942). Here, we take the advantage of the abstract notation as having brief and concise derivation. Alternatively, these derivations can be performed directly via the action principle as we do it for the energy conservation.

For details of the conservation laws, we refer to Ripa (1981), Salmon (1982; 1983), Holm (2011).

2.3 Hamilton's principle for nearly geostrophic fluids

Hamilton's principle is used to derive a balance model for many years and Salmon (1983) is the pioneer in this field. He showed a new approach for nearly geostrophic flow in a layer of shallow water. He applied to all approximations to the Lagrangian of the model, and then he obtained the equations of motion using Hamilton's principle. His major reason for choosing this method is to have the symmetry property of the Hamiltonian systems. With this property, conservation laws like energy and potential vorticity are preserved automatically.

Salmon (1985) applied his method to a semi-geostrophic scaled model. In this model, variation in the vertical component of the Coriolis vector is allowed. He constructed equations in two steps. First, he extended the Lagrangian which includes the rotational energy of the fluid in the leading of small Rossby number. In the first order, there is no kinetic energy contribution as a result of the extension. Then, he replaced the kinetic energy contribution by the one from the geostrophic velocity in the next order. In this model, the geostrophic velocity is a part of the total velocity and the rest is called ageostrophic velocity. Thus, one can think that the geostrophic velocity provides the slow motion because it is eliminated as its local rate of change is zero. Then, inertia-gravity waves are directly eliminated from the dynamics. The new Lagrangian provides new equations set via the Hamilton's principle. This new model is called L_1 model. By doing this, he showed that the model for nearly geostrophic flows on which the horizontal length scale is assumed larger than the deformation of the Rossby radius. This makes the model is appropriate for simple numerical models of ocean thermocline. It includes the relative vorticity effect on the large-scale flow.

Second, he imposed the transformation to L_1 to make the Hamiltonian structure canonical. However, it makes the model dynamically ill-posed, even transformation makes the model simpler. L_1 model still obeys conservation laws because the scale analysis is directly applied to Lagrangian. Then, the transformation coordinate is determined by the direct substitution to L_1. In the model, the geostrophic velocity is replaced by the real velocity.

He applied the Hamilton's method to the stratified fluid using f-plane primitive equations in semi-geostrophic scale in the assumption of small Rossby number. He named it as LSGE model (Salmon, 1996). On the derivations, the velocity field is decomposed into mean and mean-free parts along the vertical axis and the elliptic equation is written for the mean-free part. Again, he applied the transformation to make the model constrained by the geostrophic velocity simpler. However, higher-order terms are dropped by doing this even he has the desired L_1 term. He claimed that the resulting model does not include any fast motion. In addition to this argument, he mentioned that 2D elliptic equation is enough in contrast to quasi- and semi-geostrophic approximations on which it is necessary 3D elliptic equation.

Oliver (2006) revisited and generalized Salmon's method. In contrast to Salmon, his derivations start with an arbitrary change of coordinate for variables. Then, he expanded the Lagrangian in powers of the small number. At the end, the Lagrangian truncates at the desired order, and so the system is constrained to a sub-manifold as a Dirac constraint. Then, the balance in the leading order in a small number approximation like small Rossby number is obtained. The degenerate Lagrangian brings the balance relation for semi-geostrophic models.

To sum up, Salmon puts the constraint to the model in advance on which the balance model is obtained after the transformation. On the other hand, Oliver (2006) made the transformation before constraining in contrast to Salmon. Then, he suggests to look for a degeneracy on the Lagrangian which brings a constraint automatically. By doing this, higher order models which are not available Salmon's method become possible to obtain. Besides, his approach allows us to keep some important mathematical features such as regularity of the potential vorticity inversion. It is also possible to obtain a balance model for different scales, i.e. semi-geostrophic, quasi-geostrophic and LSGE with his generalization. In the next part, we explain how to set the degeneracy condition with Oliver (2006)'s approach.

2.3.1 Transformation on L_1 model

In this part, we explain how to construct a transformation that creates a degeneracy on the truncated Lagrangian to have a balance relation. First, we define a near-identity change of coordinates for the flow map and the density as

$$\boldsymbol{\eta}_\varepsilon = \boldsymbol{\eta} + \varepsilon\,\boldsymbol{\eta}' + \mathcal{O}(\varepsilon^2)\,, \tag{2.28a}$$

$$\rho_\varepsilon = \rho + \varepsilon\,\rho' + \mathcal{O}(\varepsilon^2)\,, \tag{2.28b}$$

for some small number ε. Next, we insert them to the Lagrangian of Boussinesq equations in the non-dimensional form. Here, it is written for a general model. Then, variables might be different for the scaled region. For instance, the velocity field is taken as a full velocity for the mid-latitude scaling, but it keeps only the zonal velocity contribution as $\boldsymbol{u} = (u_1, 0, 0)$ for the derivations around the equator. Thus, the following derivations are taken up in Chapter 3 and 4 separately in detail. The expansion of the non-dimensionalized Lagrangian reads

$$\begin{aligned} L_\varepsilon &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon + \frac{\varepsilon}{2}\,|\dot{\boldsymbol{\eta}}_\varepsilon|^2\,d\boldsymbol{a} - \int_{\mathcal{D}} \rho_\varepsilon\,z\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} (\boldsymbol{R} \circ \boldsymbol{\eta} + \varepsilon\,(\boldsymbol{\nabla}\boldsymbol{R}) \circ \boldsymbol{\eta}\,\boldsymbol{\eta}') \cdot (\dot{\boldsymbol{\eta}} + \varepsilon\,\dot{\boldsymbol{\eta}}') + \frac{\varepsilon}{2}\,|\dot{\boldsymbol{\eta}} + \varepsilon\,\dot{\boldsymbol{\eta}}'|^2\,d\boldsymbol{a} \\ &\quad - \int_{\mathcal{D}} (\rho + \varepsilon\,\rho')\,z\,d\boldsymbol{x} + \mathcal{O}(\varepsilon^2)\,. \end{aligned} \tag{2.29}$$

Here, we are interested in (2.29) up to the first order. Then, we insert Lin constraint given in (2.6) to obtain the equivalent Eulerian form,

$$\begin{aligned} L_\varepsilon &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}}\,d\boldsymbol{a} - \int_{\mathcal{D}} \rho\,z\,d\boldsymbol{x} \\ &\quad + \varepsilon \int_{\mathcal{D}} (\boldsymbol{\nabla}\boldsymbol{R}) \circ \boldsymbol{\eta}\,\boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}} + \boldsymbol{R} \circ \boldsymbol{\eta}\dot{\boldsymbol{\eta}}' + \frac{1}{2}\,|\dot{\boldsymbol{\eta}}|^2\,d\boldsymbol{a} + \varepsilon \int_{\mathcal{D}} \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v})\,z\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}}\,d\boldsymbol{a} - \int_{\mathcal{D}} \rho\,z\,d\boldsymbol{x} \\ &\quad + \varepsilon \int_{\mathcal{D}} (\boldsymbol{\nabla}\boldsymbol{R} - (\boldsymbol{\nabla}\boldsymbol{R})^T) \circ \boldsymbol{\eta}\,\boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}} + \frac{1}{2}\,|\dot{\boldsymbol{\eta}}|^2\,d\boldsymbol{a} - \varepsilon \int_{\mathcal{D}} \rho\,\boldsymbol{v} \cdot \boldsymbol{\nabla} z\,d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{x} - \rho\,z\,d\boldsymbol{x} + \varepsilon \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v} + \frac{1}{2}\,|\boldsymbol{u}|^2 - \rho\,\boldsymbol{k} \cdot \boldsymbol{v}\,d\boldsymbol{x} \\ &= L_0 + \varepsilon\,L_1\,, \end{aligned} \tag{2.30}$$

where

$$L_0 = \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{x} - \rho\, z\, d\boldsymbol{x}\,, \tag{2.31a}$$

$$L_1 = \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v} + \frac{1}{2}\,|\boldsymbol{u}|^2 - \rho\, \boldsymbol{k} \cdot \boldsymbol{v}\, d\boldsymbol{x}\,. \tag{2.31b}$$

Before going further, we split the velocity field

$$\boldsymbol{u} = \bar{\boldsymbol{u}} + \hat{\boldsymbol{u}}\,, \tag{2.32}$$

where $\bar{\boldsymbol{u}}$ is the mean velocity along the characteristics. The reminder, the perturbation from the mean velocity, is denoted $\hat{\boldsymbol{u}}$.

The degeneracy condition is imposed via the vector $\boldsymbol{v}$ which corresponds to the change of coordinates. Thus, it must be chosen in the form which makes the Lagrangian of order ε affine in the mean-free velocity field $\hat{\boldsymbol{u}}$. Here, the matrix J is not invertible. Then, we cannot write directly the approximate form of the vector $\boldsymbol{v}$. Some more assumptions need to be done. For the f-plane model, we use the projection matrix P and the corresponding kinematic relation is derived in Chapter 3. On the β-plane model, we have simpler case because of the scaling but nonconstant Coriolis vector causes complexity on the derivation of the model. Details are given in Chapter 4.

To illustrate the procedure, we consider a hydrostatic fluid and we take $\boldsymbol{\Omega} = (0, 0, 1)$. This corresponds to the primitive equations setting Oliver and Vasylkevych (2016). Then, the Lagrangian is

$$L = \int_{\mathcal{D}} u \cdot R - \rho\, z\, dx + \varepsilon \int_{\mathcal{D}} u \cdot v^{\perp} + \frac{1}{2}\,|u|^2 - \rho\, \boldsymbol{k} \cdot \boldsymbol{v}\, d\boldsymbol{x}\,. \tag{2.33}$$

We set the transformation vector as following

$$\boldsymbol{v} = \frac{1}{2} \begin{pmatrix} \hat{u}^{\perp} \\ \nabla^{\perp} \cdot \hat{U} \end{pmatrix} - \lambda \begin{pmatrix} u_g^{\perp} \\ \nabla^{\perp} \cdot U_g \end{pmatrix}\,, \tag{2.34}$$

where the vertical components are chosen such that both terms on the right-hand side of (2.34) we namely divergence free,

$$\hat{U} = \int_{-1}^{z} \hat{u}\, dz' , \quad \text{and} \quad U_g = \int_{-1}^{z} u_g \, dz' . \tag{2.35}$$

After inserting (2.34) into (2.33) of order ε, we have

$$\begin{aligned} L_1 &= \int_{\mathcal{D}} \lambda\, \hat{u} \cdot u_g + \frac{1}{2}|\bar{u}|^2 - \rho\, \nabla^{\perp} \cdot \left(\frac{1}{2}\hat{U} - \lambda U_g \right) d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \nu\, \hat{u} \cdot u_g + \frac{1}{2}|\bar{u}|^2 - \lambda\, |u_g|^2 \, d\boldsymbol{x} , \end{aligned} \tag{2.36}$$

where $\nu = \frac{1}{2} + \lambda$. On the first line, integration by parts is applied to the last term and the thermal wind relation is taken as $\partial_z u_g = -\nabla^{\perp}\rho$. We take the variation of (2.36) to obtain the equations of motion via Theorem 3.2.1. We have

$$\begin{pmatrix} u^{\perp} \\ \rho \end{pmatrix} + \varepsilon \begin{pmatrix} \partial_t p + u^{\perp}\zeta + u_3\, \partial_z p \\ -u \cdot \partial_z p \end{pmatrix} - \varepsilon \nabla\rho\, \nabla^{\perp} \cdot B = \boldsymbol{\nabla}\phi , \tag{2.37}$$

where $B = \nu\hat{U} - 2\lambda U_g$, the potential ϕ is

$$\phi = -p - \rho\, z - \frac{1}{2}|u|^2 , \tag{2.38}$$

and $\zeta = \nabla^{\perp}\bar{u} + \nu\, \Delta\theta$ with

$$\theta = -\int_{-1}^{0} z\, \rho\, dz' - \int_{-1}^{z} \rho\, dz' . \tag{2.39}$$

We take the curl of (2.37) to eliminate the potential term. Its horizontal component is

$$\begin{aligned} &-(1 + \varepsilon\, \zeta)\, \partial_z^2 B + \varepsilon\nu\, \partial_z\rho\, \Delta B - 2\varepsilon\nu\, \nabla\rho\, \nabla\partial_z B \\ &\qquad = (\nu - 2\lambda - 2\varepsilon\lambda\zeta - 2\varepsilon\nu\lambda\Delta\theta)\nabla^{\perp}\rho + \varepsilon\nu^2 \big(\nabla\bar{u} + (\nabla\bar{u})^{\perp}\big)\, \nabla\rho + 4\varepsilon\nu\lambda\nabla u_g \nabla\rho . \end{aligned} \tag{2.40}$$

Equation (2.40) provides a relationship between the velocity field and the density. The average of the vertical component of the curl of (2.37) is

$$\partial_t \omega + \bar{u} \cdot \nabla \omega = \overline{\nabla \rho \cdot \nabla^{\perp} \nabla^{\perp} \cdot B} + \overline{\nabla \cdot (\nabla \rho \nabla \cdot B - \nu \, \hat{u} \Delta \theta)}\,, \tag{2.41}$$

where $\omega = \nabla^{\perp} \cdot \bar{u}$. The overline represents the average along the characteristics. At the end, the resulting balance model includes a kinematic relationship (2.40) and the evolution equation for the vorticity (2.41).

In our derivations, we use particularly L_1 dynamics and we follow the same strategy to that Oliver and Vasylkevych (2016). In their approach, the velocity field is determined by up to a constant of integration, while Oliver (2006) determined the velocity field completely by geostrophic balance. We do the same without hydrostatic approximation. Our derivations are constructed on the separation of the velocity field, namely mean and mean-free velocity. The mean field is defined using the axis of rotation on which the horizontal component of the Coriolis vector is not ignored. The key point is that the transformation is imposed by considering the mean-free part of the velocity field in the Lagrangian because it causes the degeneracy. We fix it via a transformation vector. Resulting models have two main components. One of them is a transport equation for the prognostic variable which provides the slow dynamics of the system. The other component is the kinematic relationship which corresponds to balance relation. The velocity field can be obtained via them. The resulting model for the mid-latitude is the most generalized model in semi-geostrophic scaling. The other model is the promising work to provide a equatorial balance model.

In this chapter, we give a general overview on how to use Hamilton's principle on rotating flows. Following two chapters derive the balance model using the Hamilton's principle for L_1 model on f- and β-plane for two different regions, mid-latitude and equator.

Chapter 3

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

Reproduced from [Özden, G., & Oliver, M. (2021, Jul). Variational balance models for the three-dimensional Euler–boussinesq equations with full coriolis force. Physics of Fluids, 33(7), 076606.], with the permission of AIP Publishing.

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

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

3.1 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

$$\varepsilon \left(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} \right) + \boldsymbol{\Omega} \times \boldsymbol{u} = -\boldsymbol{\nabla} p - \rho\, \boldsymbol{k} \,, \tag{3.1a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \,, \tag{3.1b}$$

$$\partial_t \rho + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho = 0 \,. \tag{3.1c}$$

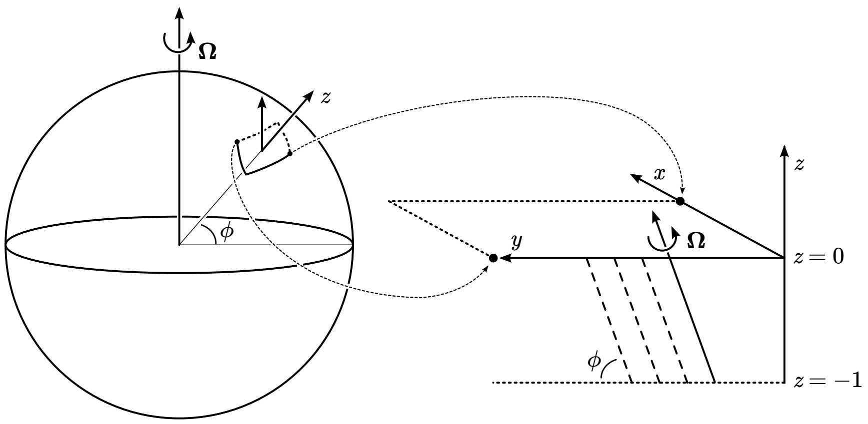


FIGURE 3.1: Geometry of the f-plane approximation at latitude ϕ with the full Coriolis vector. Here, x, y, and z are directed toward the east, the north, and upward, respectively. The dashed lines parallel to the axis of rotation in the y-z plane represent the characteristics of the thermal wind relation.

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

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ \cos\phi \\ \sin\phi \end{pmatrix} \equiv \begin{pmatrix} 0 \\ c \\ s \end{pmatrix} \tag{3.2}$$

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 3.1, 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 et al., 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,

$$\mathcal{D} = \mathbb{T}^2 \times [-1, 0]\,, \tag{3.3}$$

so that

$$\boldsymbol{k} \cdot \boldsymbol{u} = 0 \quad \text{at} \quad z = -1, 0\,. \tag{3.4}$$

The semi-geostrophic scaling used in (3.1) 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 et al., 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 et al., 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, McWilliams, and Rousseau, 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 and 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 et al., 2019 perform a numerical study of balance for a moist atmosphere with phase changes, and Kafiabad and Bartello, 2018; Kafiabad and Bartello, 2017; Kafiabad and Bartello, 2016 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; Salmon, 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, and Oliver, 2017). In other configurations, such as when the Coriolis parameter is spatially varying (Oliver and 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 semi-geostrophic equations. His transformation has subsequently been interpreted as a Legendre transformation, providing a sense of generalized solutions via the theory of optimal transport (Cullen and Purser, 1984; Benamou and Brenier, 1998; Cullen, 2006; Colombo, 2017; Roulstone and 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, and 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 (Gerkema and Shrira, 2005a; Gerkema and Shrira, 2005c; Gerkema et al., 2008b; Brummell, Hurlburt, and Toomre, 1996; Juárez, Fisher, and Orton, 2002; Hayashi and Itoh, 2012) and are also studied in the context of other reduced models such as layers of shallow water (Stewart and Dellar, 2010; Stewart and Dellar, 2012) and models on the β-plane (Dellar, 2011).

Our approach is similar, in principle, to the primitive equation case (Oliver and 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 3.2 recalls the variational derivation for the Euler–Boussinesq equations in the language of the Euler–Poincaré variational principle. Section 3.3 explains the general setting for variational asymptotics in this framework. In Section 3.4, 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 3.5 discusses the leading order balance, the thermal wind relation, on the mid-latitude f-plane with full Coriolis force. In Section 3.6, 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 3.7, we derive the balance model Euler–Poincaré equations from the balance model Lagrangian. In Section 3.8, 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 3.9 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.

3.2 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 $\mathrm{Diff}_\mu(\mathcal{D})$ 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

$$V_{\mathrm{div}} = \{\boldsymbol{u} \in H^s(\mathcal{D}, \mathbb{R}^3) \colon \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0,\ \boldsymbol{k} \cdot \boldsymbol{u} = 0 \text{ at } z = -1, 0\}\,. \tag{3.5}$$

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

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

$$\partial_t \boldsymbol{\eta}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{\eta}(\boldsymbol{a}, t), t) \tag{3.6}$$

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 (3.1c) is equivalent to

$$\rho \circ \boldsymbol{\eta} = \rho_0 \,, \tag{3.7}$$

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

$$\begin{aligned} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_0) &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \frac{\varepsilon}{2} \, |\dot{\boldsymbol{\eta}}|^2 - \rho \, \boldsymbol{k} \cdot \boldsymbol{\eta} \, d\boldsymbol{a} \\ &= \int_{\mathcal{D}} \boldsymbol{R} \cdot \boldsymbol{u} + \frac{\varepsilon}{2} \, |\boldsymbol{u}|^2 - \rho \, z \, d\boldsymbol{x} \equiv \ell(\boldsymbol{u}, \rho) \,. \end{aligned} \tag{3.8}$$

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 and Lifshitz, 1976. Here, on the f-plane, we can choose

$$\boldsymbol{R} = \frac{1}{2} \mathsf{J} \boldsymbol{x} = \frac{1}{2} \begin{pmatrix} cz - sy \\ sx \\ -cx \end{pmatrix} , \tag{3.9}$$

where J is the skew-symmetric matrix

$$\mathsf{J} = \begin{pmatrix} 0 & -s & c \\ s & 0 & 0 \\ -c & 0 & 0 \end{pmatrix}. \tag{3.10}$$

As we see in the second line of (3.8), 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 3.2.1 (Holm, Marsden, and Ratiu, 1998)**.** *Consider a curve $\boldsymbol{\eta}$ in $\mathrm{Diff}_\mu(\mathcal{D})$ with Lagrangian velocity $\dot{\boldsymbol{\eta}}$ and Eulerian velocity $\boldsymbol{u} \in V_{\mathrm{div}}$. Then the following are equivalent.*

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

$$\delta \int_{t_1}^{t_2} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_0)\, dt = 0 \tag{3.11}$$

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

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

$$\delta \int_{t_1}^{t_2} \ell(\boldsymbol{u}, \rho)\, dt = 0 \tag{3.12}$$

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_{div} vanishing at the temporal end points.

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

$$\int_{\mathcal{D}} (\partial_t + \mathcal{L}_{\boldsymbol{u}})\boldsymbol{m} \cdot \boldsymbol{v} + \frac{\delta \ell}{\delta \rho}\, \mathcal{L}_{\boldsymbol{v}} \rho \, d\boldsymbol{x} = 0 \tag{3.13}$$

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

$$\boldsymbol{m} = \frac{\delta \ell}{\delta \boldsymbol{u}} \tag{3.14}$$

is the momentum one-form.

In the language of vector fields on a region of $\mathbb{R}^3$, (3.13) reads

$$\int_{\mathcal{D}} \Big(\partial_t \boldsymbol{m} + (\boldsymbol{\nabla} \times \boldsymbol{m}) \times \boldsymbol{u} + \boldsymbol{\nabla}(\boldsymbol{m} \cdot \boldsymbol{u}) + \frac{\delta \ell}{\delta \rho} \boldsymbol{\nabla} \rho \Big) \cdot \boldsymbol{v} \, d\boldsymbol{x} = 0 \tag{3.15}$$

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

For the Euler–Boussinesq Lagrangian (3.8),

$$\boldsymbol{m} = \frac{\delta \ell}{\delta \boldsymbol{u}} = \boldsymbol{R} + \varepsilon \, \boldsymbol{u} \qquad \text{and} \qquad \frac{\delta \ell}{\delta \rho} = -z \, . \tag{3.16}$$

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 (3.15) implies the Euler–Boussinesq momentum equation (3.1).

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

$$H = \int_{\mathcal{D}} \frac{\delta \ell}{\delta \boldsymbol{u}} \cdot \boldsymbol{u} \, d\boldsymbol{x} - \ell(\boldsymbol{u}, \rho) = \int_{\mathcal{D}} \frac{\varepsilon}{2} \, |\boldsymbol{u}|^2 + z \, \rho \, d\boldsymbol{x} \tag{3.17}$$

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 $\operatorname{Diff}_\mu(\mathcal{D})$: Ripa, 1981 and Salmon, 1982 have shown that this symmetry implies material conservation of the Ertel *potential vorticity*

$$q = (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla} \rho = (\boldsymbol{\Omega} + \varepsilon \, \boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{\nabla} \rho \, , \tag{3.18}$$

that is, q satisfies the advection equation

$$\partial_t q + \boldsymbol{u} \cdot \boldsymbol{\nabla} q = 0 \, . \tag{3.19}$$

3.3 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,

$$\dot{\boldsymbol{\eta}}_\varepsilon = \boldsymbol{u}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \,. \tag{3.20}$$

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

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta} \,. \tag{3.21}$$

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

$$\boldsymbol{\eta}_\varepsilon' = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \,, \qquad \boldsymbol{\eta}_\varepsilon\big|_{\varepsilon=0} = \boldsymbol{\eta} \,. \tag{3.22}$$

Here and in the following, we use the prime symbol to denote a derivative with respect to ε. Cross-differentiation of (3.20) and (3.22) gives

$$\boldsymbol{u}_\varepsilon' = \dot{\boldsymbol{v}}_\varepsilon + \boldsymbol{u}_\varepsilon \cdot \boldsymbol{\nabla} \boldsymbol{v}_\varepsilon - \boldsymbol{v}_\varepsilon \cdot \boldsymbol{\nabla} \boldsymbol{u}_\varepsilon \,. \tag{3.23}$$

Likewise, differentiation in ε of the Lagrangian density gives

$$\rho_\varepsilon' + \boldsymbol{v}_\varepsilon \cdot \boldsymbol{\nabla} \rho_\varepsilon = 0 \,. \tag{3.24}$$

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 3.2.1.

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 (3.23) and (3.24) 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 $v = v_\varepsilon|_{\varepsilon=0}$ 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, Gottwald, and Oliver, 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 and Oliver, 2014). For partial differential equations, this question is open; it is clear, though, that additional conditions are necessary.

3.4 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

$$\mathsf{Q} = \boldsymbol{\Omega}\boldsymbol{\Omega}^\mathsf{T} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & c^2 & cs \\ 0 & cs & s^2 \end{pmatrix} \tag{3.25}$$

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

$$\mathsf{P} = \mathsf{I} - \mathsf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & s^2 & -cs \\ 0 & -cs & c^2 \end{pmatrix} . \tag{3.26}$$

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

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \zeta\,\boldsymbol{\Omega} \equiv \mathsf{A}\boldsymbol{\xi} \equiv \boldsymbol{\chi}(\boldsymbol{\xi}) \tag{3.27}$$

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

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & s \end{pmatrix} . \tag{3.28}$$

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

$$\partial_i (f \circ \boldsymbol{\chi}) = \begin{cases} (\partial_i f) \circ \boldsymbol{\chi} & \text{for } i = x, y \\ (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} f) \circ \boldsymbol{\chi} & \text{for } i = \zeta \end{cases} \tag{3.29}$$

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

$$\boldsymbol{\nabla} \cdot (\boldsymbol{v} \circ \chi) = (\boldsymbol{\nabla} \cdot \mathsf{A}\boldsymbol{v}) \circ \chi \,, \tag{3.30a}$$

$$(\boldsymbol{\nabla} \cdot \mathsf{P}\boldsymbol{u}) \circ \chi = \boldsymbol{\nabla} \cdot (\mathsf{SP}\boldsymbol{u} \circ \chi) \,, \tag{3.30b}$$

where (3.30b) follows from (3.30a) with $\boldsymbol{v} = \mathsf{A}^{-1}\mathsf{P}\boldsymbol{u}$ since $\mathsf{A}^{-1}\mathsf{P} = \mathsf{SP}$ with

$$\mathsf{S} = \operatorname{diag}(1, s^{-2}, s^{-1}) \,. \tag{3.31}$$

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

$$\bar{\phi} \circ \chi = s \int_{-s^{-1}}^{0} \phi \circ \chi \, d\zeta \,. \tag{3.32}$$

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, $\bar{\psi} \circ \chi$ is independent of ζ and $\det \mathsf{A} = s$, we see that

$$\int_{\mathcal{D}} \hat{\phi} \, \bar{\psi} \, d\boldsymbol{x} = s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \hat{\phi} \circ \chi \, d\zeta \, \bar{\psi} \circ \chi \, dx = 0 \,. \tag{3.33}$$

In other words, mean and fluctuating components in the sense of (3.32) are L^2-orthogonal.

3.5 Thermal wind

The formal leading order balance in the Euler–Boussinesq momentum equation (3.1a) gives an expression for the *thermal* or *geostrophic wind* $\boldsymbol{u}_g$,

$$\boldsymbol{\Omega} \times \boldsymbol{u}_g = -\boldsymbol{\nabla} p - \rho \, \boldsymbol{k} \,. \tag{3.34}$$

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

$$\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{u}_g = \begin{pmatrix} -\nabla^{\perp}\rho \\ 0 \end{pmatrix} \,. \tag{3.35}$$

The characteristics of this first order equation are the lines parallel to the axis of rotation. With the notation set up in Section 3.4, it is easy to integrate (3.35) along its characteristic lines. Splitting $\boldsymbol{u}_g = \hat{\boldsymbol{u}}_g + \bar{\boldsymbol{u}}_g$, we first note that (3.35) determines only the mean-free component $\hat{\boldsymbol{u}}_g$. Indeed,

$$\partial_\zeta(\hat{\boldsymbol{u}}_g \circ \chi) = (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \hat{\boldsymbol{u}}_g) \circ \chi = \begin{pmatrix} -\nabla^\perp \rho \circ \chi \\ 0 \end{pmatrix} . \tag{3.36}$$

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

$$0 = s \int_{-s^{-1}}^{0} \hat{\boldsymbol{u}}_g \circ \chi \, d\zeta = \hat{\boldsymbol{u}}_g \circ \chi(-s^{-1}) - s \int_{-s^{-1}}^{0} \zeta \, \partial_\zeta(\hat{\boldsymbol{u}}_g \circ \chi) \, d\zeta \,, \tag{3.37}$$

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

$$\begin{aligned} \hat{\boldsymbol{u}}_g \circ \chi &= \hat{\boldsymbol{u}}_g \circ \chi(-s^{-1}) + \int_{-s^{-1}}^{\zeta} \partial_\zeta(\hat{\boldsymbol{u}}_g \circ \chi') \, d\zeta' \\ &= s \int_{-s^{-1}}^{0} \zeta \, \partial_\zeta(\hat{\boldsymbol{u}}_g \circ \chi) \, d\zeta + \int_{-s^{-1}}^{\zeta} \partial_\zeta(\hat{\boldsymbol{u}}_g \circ \chi') \, d\zeta' \,. \end{aligned} \tag{3.38}$$

Inserting the thermal wind relation (3.36), 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

$$\hat{u}_g = \nabla^\perp \theta \tag{3.39a}$$

with

$$\theta \circ \chi = -s \int_{-s^{-1}}^{0} \zeta \, \rho \circ \chi \, d\zeta - \int_{-s^{-1}}^{\zeta} \rho \circ \chi' \, d\zeta' \,. \tag{3.39b}$$

Equation (3.39) shows that $\hat{u}_g$ is horizontally divergence-free. Since $\boldsymbol{k} \cdot \hat{\boldsymbol{u}}_g = 0$, we also have $\boldsymbol{\nabla} \cdot \boldsymbol{u}_g = 0$.

3.6 Derivation of the first-order balance model

We now implement the procedure outlined in Section 3.3 at first order in ε. Using the Lin constraints for the ε-flow, (3.23) and (3.24), we can write

$$\boldsymbol{u}_\varepsilon = \boldsymbol{u} + \varepsilon \, (\dot{\boldsymbol{v}} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u}) + O(\varepsilon^2) \tag{3.40}$$

and

$$\rho_\varepsilon = \rho - \varepsilon \, \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) + O(\varepsilon^2) \,. \tag{3.41}$$

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 (3.8), expanded in powers of ε and truncated to $O(\varepsilon)$, we obtain after a short computation

$$\begin{aligned} L &= \int_{\mathcal{D}} \boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon + \tfrac{\varepsilon}{2} \, |\dot{\boldsymbol{\eta}}_\varepsilon|^2 \, d\boldsymbol{a} - \int_{\mathcal{D}} \rho_\varepsilon \, z \, d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{x} - \rho \, z \, d\boldsymbol{x} + \varepsilon \int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v} + \tfrac{1}{2} \, |\boldsymbol{u}|^2 - \rho \, \boldsymbol{v} \cdot \boldsymbol{k} \, d\boldsymbol{x} \\ &= L_0 + \varepsilon \, L_1 \,. \end{aligned} \tag{3.42}$$

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

$$\int_{\mathcal{D}} |\boldsymbol{u}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} |\hat{\boldsymbol{u}}|^2 + 2 \, \hat{\boldsymbol{u}} \cdot \bar{\boldsymbol{u}} + |\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \,. \tag{3.43}$$

The cross term in (3.43) vanishes due to (3.33). 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

$$\int_{\mathcal{D}} |\hat{\boldsymbol{u}}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}} \cdot \mathsf{Q}\hat{\boldsymbol{u}} \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}(\hat{\boldsymbol{u}} + \boldsymbol{C}[\hat{\boldsymbol{u}}]) \, d\boldsymbol{x} \,, \tag{3.44}$$

using Lemma 3.13.1 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

$$\begin{aligned}\int_{\mathcal{D}} |\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} &= \int_{\mathcal{D}} (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}} \cdot (\mathsf{I} - \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}} + \mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} \cdot \mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}} \, d\boldsymbol{x} \\ &= -s^{-1} \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\boldsymbol{\nabla}(\boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}})\, z \, d\boldsymbol{x} + \int_{\mathcal{D}} |\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \,. \end{aligned} \tag{3.45}$$

The second equality is based on Lemma 3.13.2 with $\bar{\phi} = s^{-1}\,\boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}}$. Note that Lemma 3.13.2 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 3.12.1, by the fact that the remainder integral in (3.45) 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

$$\int_{\mathcal{D}} |\boldsymbol{u}|^2 \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{\boldsymbol{u}} \cdot \mathsf{P}\hat{\boldsymbol{V}} \, d\boldsymbol{x} + \int_{\mathcal{D}} |\mathsf{S}_0\mathsf{P}\bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \tag{3.46}$$

with

$$\boldsymbol{V} = \hat{\boldsymbol{u}} + \boldsymbol{C}[\hat{\boldsymbol{u}}] - s^{-1}\,\boldsymbol{\nabla}\big(\boldsymbol{\Omega} \cdot (\mathsf{I} + \mathsf{S}_0\mathsf{P})\bar{\boldsymbol{u}}\big)\, z \,. \tag{3.47}$$

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

$$\mathsf{J}\boldsymbol{v} = -\frac{1}{2}\mathsf{P}\hat{\boldsymbol{V}} \,, \tag{3.48}$$

up to terms that only depend on ρ, we can make L_1 affine in $\hat{\boldsymbol{u}}$: Noting that J is invertible on $\operatorname{Range}\mathsf{P}$ 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

$$\boldsymbol{v} = \mathsf{P}\hat{\boldsymbol{v}} + \mathsf{Q}\hat{\boldsymbol{v}} + \bar{\boldsymbol{v}} \tag{3.49}$$

with

$$\mathsf{P}\hat{\boldsymbol{v}} = -\frac{1}{2}\mathsf{J}^\mathsf{T}\,\hat{\boldsymbol{V}} + \lambda\,\mathsf{J}^\mathsf{T}\hat{\boldsymbol{V}}_g\,. \tag{3.50}$$

In this expression, λ is a free parameter and

$$\boldsymbol{V}_g = \hat{\boldsymbol{u}}_g + \boldsymbol{C}[\hat{\boldsymbol{u}}_g] \tag{3.51}$$

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 (3.47) into the ansatz for $\boldsymbol{V}_g$.

The terms $\mathsf{Q}\hat{\boldsymbol{v}}$ and $\bar{\boldsymbol{v}}$ in (3.49) are chosen such that $\boldsymbol{v}$ becomes divergence free and tangent to the top and bottom boundaries. Using the construction from Lemma 3.12.4, we write $\mathsf{Q}\hat{\boldsymbol{v}} \equiv \boldsymbol{\Omega}\,\hat{g}$, where

$$\hat{g}\circ\chi = \hat{g}\circ\chi(-s^{-1}) - \int_{-s^{-1}}^{\zeta} \boldsymbol{\nabla}\cdot(\mathsf{S}\mathsf{P}\hat{\boldsymbol{v}}\circ\chi')\,d\zeta'\,. \tag{3.52}$$

By Lemma 3.12.5, $\hat{\boldsymbol{v}}$ is divergence-free. By Lemma 3.13.3, $\bar{\boldsymbol{v}}$ can be chosen such that the zero flux boundary condition

$$\begin{aligned} \bar{v}_3\circ\chi &= -\boldsymbol{k}\cdot\mathsf{P}\hat{\boldsymbol{v}}\circ\chi(-s^{-1}) - s\,\hat{g}\circ\chi(-s^{-1}) \\ &= -\boldsymbol{k}\cdot\mathsf{P}\hat{\boldsymbol{v}}\circ\chi(0) - s\,\hat{g}\circ\chi(0)\,. \end{aligned} \tag{3.53}$$

is satisfied and, moreover, $\bar{\boldsymbol{v}}$ is divergence free. Lemma 3.13.4 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 (3.53), so that the final expression will not contain any references to $\bar{\boldsymbol{v}}$.

Combining the contributions from rotation and kinetic energy, we find

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

$$= \int_{\mathcal{D}} \lambda\, \hat{u} \cdot \hat{u}_g + \tfrac{1}{2}\, |\mathsf{S}_0 \mathsf{P} \bar{\boldsymbol{u}}|^2 \, d\boldsymbol{x} \tag{3.54}$$

where, in the last equality, we have used Lemma 3.13.1 and the fact that the geostrophic velocity is exclusively horizontal. The contribution to the L_1-Lagrangian from the potential energy term is

$$- \int_{\mathcal{D}} \rho\, \boldsymbol{v} \cdot \boldsymbol{k} \, d\boldsymbol{x} = \int_{\mathcal{D}} \hat{u}_g \cdot \left(\tfrac{1}{2}\, \hat{u} - \lambda\, \hat{u}_g\right) d\boldsymbol{x}\,; \tag{3.55}$$

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

$$\begin{aligned} L_1 &= \int_{\mathcal{D}} \lambda\, \hat{u} \cdot \hat{u}_g + \tfrac{1}{2}\, |\mathsf{S}_0 \mathsf{P} \bar{\boldsymbol{u}}|^2 + \hat{u}_g \cdot (\tfrac{1}{2}\, \hat{u} - \lambda\, \hat{u}_g)\, d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \nu\, \hat{u} \cdot \hat{u}_g + \tfrac{1}{2}\, \bar{\boldsymbol{u}} \cdot \mathsf{M} \bar{\boldsymbol{u}} - \lambda\, |\hat{u}_g|^2 \, d\boldsymbol{x}\,, \end{aligned} \tag{3.56}$$

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

$$\mathsf{M} = \mathsf{P}\mathsf{S}_0^2\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c/s \\ 0 & -c/s & c^2/s^2 \end{pmatrix} . \tag{3.57}$$

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

3.7 Derivation of the balance model equations of motion

Taking the variation of (3.56), we obtain

$$\begin{aligned} \delta L_1 &= \int_{\mathcal{D}} \delta \boldsymbol{u} \cdot (\nu\, \hat{\boldsymbol{u}}_g + \mathsf{M} \bar{\boldsymbol{u}}) + \nu\, u \cdot \delta \hat{u}_g - 2\lambda\, \delta \hat{u}_g \cdot \hat{u}_g \, d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \delta \boldsymbol{u} \cdot \boldsymbol{p} + \delta \hat{u}_g \cdot \hat{b}\, d\boldsymbol{x}\,, \end{aligned} \tag{3.58}$$

where

$$\boldsymbol{p} = \mathsf{M} \bar{\boldsymbol{u}} + \nu\, \hat{\boldsymbol{u}}_g \qquad \text{and} \qquad \hat{b} = \nu\, \hat{u} - 2\lambda\, \hat{u}_g\,, \tag{3.59}$$

To rewrite the second term on the right of (3.58), we insert the expression for the geostrophic velocity (3.39), change variables, and recall that horizontal derivatives and composition with χ commute, so that

$$\begin{aligned}
\int_{\mathcal{D}} \delta \hat{u}_g \cdot \hat{b}\, d\boldsymbol{x} &= -s \int_{\mathcal{D}} \int_{-s^{-1}}^{\zeta} \nabla^{\perp} \delta\rho \circ \boldsymbol{\chi}'\, d\zeta' \cdot \hat{b} \circ \boldsymbol{\chi}\, d\boldsymbol{\xi} \\
&= s \int_{\mathcal{D}} \nabla^{\perp} \delta\rho \circ \boldsymbol{\chi} \cdot \int_{-s^{-1}}^{\zeta} \hat{b} \circ \boldsymbol{\chi}'\, d\zeta'\, d\boldsymbol{\xi} \\
&= -s \int_{\mathcal{D}} \delta\rho \circ \boldsymbol{\chi}\, \nabla^{\perp} \cdot \int_{-s^{-1}}^{\zeta} \hat{b} \circ \boldsymbol{\chi}'\, d\zeta'\, d\boldsymbol{\xi} \\
&= - \int_{\mathcal{D}} \delta\rho\, \nabla^{\perp} \cdot B\, d\boldsymbol{x}\,,
\end{aligned} \tag{3.60}$$

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

$$B = \nu\, U - 2\lambda\, U_g \tag{3.61}$$

where

$$U \circ \boldsymbol{\chi} = \int_{-s^{-1}}^{\zeta} \hat{u} \circ \boldsymbol{\chi}'\, d\zeta' \tag{3.62}$$

and U_g is defined likewise. Inserting (3.60) into (3.58), we have

$$\delta L_1 = \int_{\mathcal{D}} \delta \boldsymbol{u} \cdot \boldsymbol{p} - \delta\rho\, \nabla^{\perp} \cdot B\, d\boldsymbol{x}\,. \tag{3.63}$$

We recall the general Euler–Poincaré equation (3.15), which can be written

$$\partial_t \boldsymbol{m} + (\boldsymbol{\nabla} \times \boldsymbol{m}) \times \boldsymbol{u} + \frac{\delta \ell}{\delta \rho}\, \boldsymbol{\nabla} \rho = \boldsymbol{\nabla} \tilde{\phi} \tag{3.64}$$

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

$$\boldsymbol{m} = \frac{\delta \ell}{\delta \boldsymbol{u}} = \boldsymbol{R} + \varepsilon\, \boldsymbol{p} \qquad \text{and} \qquad \frac{\delta \ell}{\delta \rho} = -z - \varepsilon\, \nabla^{\perp} \cdot B\,, \tag{3.65}$$

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

$$\boldsymbol{\Omega} \times \boldsymbol{u} + \rho \boldsymbol{k} + \varepsilon\, (\partial_t \boldsymbol{p} + (\boldsymbol{\nabla} \times \boldsymbol{p}) \times \boldsymbol{u} - \boldsymbol{\nabla}\rho\, \nabla^{\perp} \cdot B) = -\boldsymbol{\nabla} \phi\,. \tag{3.66}$$

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,

$$\boldsymbol{\xi} \equiv \boldsymbol{\nabla} \times \boldsymbol{p} = \boldsymbol{\Omega}\, \omega + \nu\, \gamma \tag{3.67}$$

with

$$\gamma = \nabla \times \hat{\boldsymbol{u}}_g = -\nabla \times \operatorname{curl} \theta = \begin{pmatrix} -\nabla \partial_z \theta \\ \Delta \theta \end{pmatrix}, \tag{3.68}$$

where we recall that $\hat{u}_g = \nabla^\perp \theta$, write $\operatorname{curl} f = \nabla \times (0, 0, f)$ 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 (3.66) and noting that $\nabla \times (\boldsymbol{\xi} \times \boldsymbol{u}) = \boldsymbol{u} \cdot \nabla \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \nabla \boldsymbol{u}$ as both $\boldsymbol{u}$ and $\boldsymbol{\xi}$ are divergence-free, we find

$$-\boldsymbol{\Omega} \cdot \nabla \boldsymbol{u} + \nabla \times (\rho \boldsymbol{k}) + \varepsilon \left(\partial_t \boldsymbol{\xi} + \boldsymbol{u} \cdot \nabla \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \nabla \boldsymbol{u} + \nabla \rho \times \nabla \nabla^\perp \cdot B \right) = 0 \,. \tag{3.69}$$

The corresponding balance model Hamiltonian, via (3.17), is given by

$$H = \int_{\mathcal{D}} \rho \, z + \varepsilon \left(\tfrac{1}{2} \, \bar{\boldsymbol{u}} \cdot \mathsf{M} \bar{\boldsymbol{u}} + \lambda \, |\hat{\boldsymbol{u}}_g|^2 \right) d\boldsymbol{x} \,. \tag{3.70}$$

The potential vorticity for the balance model reads

$$\begin{aligned} q &= (\nabla \times \boldsymbol{m}) \cdot \nabla \rho \\ &= (\boldsymbol{\Omega} + \varepsilon \, \boldsymbol{\Omega} \, \omega + \varepsilon \nu \, \gamma) \cdot \nabla \rho \\ &= (s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) \, \partial_z \rho + (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) \cdot \nabla \rho \\ &= -(s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) \, \partial_z (\boldsymbol{\Omega} \cdot \nabla \theta) - (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) \cdot \nabla (\boldsymbol{\Omega} \cdot \nabla \theta) \,, \end{aligned} \tag{3.71}$$

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

$$q = \begin{vmatrix} -(s + \varepsilon s \, \omega + \varepsilon \nu \, \Delta \theta) & \nabla (\boldsymbol{\Omega} \cdot \nabla \theta) \\ (\Omega + \varepsilon \, \Omega \, \omega - \varepsilon \nu \, \partial_z \nabla \theta) & \partial_z (\boldsymbol{\Omega} \cdot \nabla \theta) \end{vmatrix} . \tag{3.72}$$

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

$$q = F(\omega, D^2 \theta) \,, \tag{3.73}$$

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

$$[F_{ij}] = \frac{\partial F}{\partial(D^2\theta)} \tag{3.74}$$

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

$$[F_{ij}] = \begin{pmatrix} \varepsilon\nu\partial_z\rho & -\frac{1}{2}\varepsilon c\,\xi_1 & -\frac{1}{2}\varepsilon\,(\xi_1+\nu\partial_x\rho) \\ -\frac{1}{2}\varepsilon c\,\xi_1 & \varepsilon\nu\partial_z\rho - c(c+\varepsilon\xi_2) & -\frac{1}{2}\big(2cs+\varepsilon(c\xi_3+s\xi_2+\nu\partial_y\rho)\big) \\ -\frac{1}{2}\varepsilon\,(\xi_1+\nu\partial_x\rho) & -\frac{1}{2}\big(2cs+\varepsilon(c\xi_3+s\xi_2+\nu\partial_y\rho)\big) & -s^2-\varepsilon s\,\xi_3 \end{pmatrix}, \tag{3.75}$$

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

$$-\varepsilon\nu\partial_z\rho > 0\,, \tag{3.76a}$$

$$-\varepsilon\nu\,\partial_z\rho + c\big[c+\varepsilon\big(\xi_2+\tfrac{1}{4}\tfrac{c\xi_1^2}{\nu\partial_z\rho}\big)\big] > 0\,, \tag{3.76b}$$

$$\det F < 0\,. \tag{3.76c}$$

Condition (3.76c) can be written more explicitly as

$$-\,(\varepsilon s\nu\partial_z\rho)^2 + \varepsilon^2 f(\boldsymbol{\xi},\nabla\rho;\boldsymbol{\xi},\rho) < 0\,, \tag{3.77}$$

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.

3.8 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

$$\partial_t\xi_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla}\xi_3 = \boldsymbol{\xi}\cdot\boldsymbol{\nabla}u_3 + \nabla\rho\cdot\nabla^\perp\nabla^\perp\cdot B + \varepsilon^{-1}\,\boldsymbol{\Omega}\cdot\boldsymbol{\nabla}\hat{u}_3\,. \tag{3.78}$$

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 (3.78) along the axis of rotation and using Lemma 3.11.1 and Lemma 3.11.2 from the appendix to commute averaging with directional derivatives where possible, we find that the evolution equation for ζ_3 reduces to a prognostic equation for ω,

$$\partial_t \omega + \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega = \overline{\nabla\rho \cdot \nabla^\perp \nabla^\perp \cdot B} - \nu \, \overline{(\partial_z \nabla\theta \cdot \nabla u_3 - \Delta\theta \, \partial_z u_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} \Delta\theta)} \,. \tag{3.79}$$

All other contributions to (3.69) are entirely kinematic. To see this, we start with multiplying (3.79) by $\boldsymbol{\Omega}$, then subtract this expression from (3.69) to obtain

$$\begin{aligned} 0 = & - \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \times (\rho \boldsymbol{k}) + \varepsilon \left[\nu \, \partial_t \boldsymbol{\gamma} + \boldsymbol{\Omega} \, \hat{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega + \nu \, \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{\gamma} - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} \boldsymbol{u} \right. \\ & \left. + \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \nabla^\perp \cdot B + \boldsymbol{\Omega} \, \overline{\nabla\rho \cdot \nabla^\perp \nabla^\perp \cdot B} + \nu \, \boldsymbol{\Omega} \, \overline{(\boldsymbol{\gamma} \cdot \boldsymbol{\nabla} u_3 - \boldsymbol{u} \cdot \boldsymbol{\nabla} \Delta\theta)} \right] . \end{aligned} \tag{3.80}$$

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

$$\partial_z \rho = -\alpha + \partial_z \tilde{\rho} \tag{3.81}$$

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 (3.81) makes the structure of the problem more transparent.

Inserting (3.81) into the definition of γ and noting that $\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta = -\rho$, we find

$$\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \dot{\gamma}_3 = \Delta(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \dot{\theta}) = \Delta(\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho) = -\alpha \, \Delta u_3 + \Delta(\boldsymbol{u} \cdot \boldsymbol{\nabla} \tilde{\rho}) \,. \tag{3.82}$$

Decomposing

$$U = \nabla^\perp \Psi + \nabla \Phi \,, \tag{3.83}$$

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

$$\boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \nabla^\perp \cdot B = -\alpha\nu \, \boldsymbol{k} \times \boldsymbol{\nabla} \zeta + \nu \, \boldsymbol{\nabla} \tilde{\rho} \times \boldsymbol{\nabla} \zeta - 2\lambda \, \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \nabla^\perp \cdot U_g \,. \tag{3.84}$$

Inserting (3.82) and (3.84) into (3.80), 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:

$$(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})^2 \zeta + \varepsilon\alpha\nu\, \Delta\phi\zeta = -\Delta\tilde{\rho} + \varepsilon f\,, \tag{3.85a}$$

$$(\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})^2 u_3 + \varepsilon\alpha\nu\, \Delta\phi u_3 = \varepsilon g \tag{3.85b}$$

with

$$\begin{aligned} f = {} & \Omega \cdot \nabla^\perp (\hat{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega) + \nu\, \nabla^\perp \cdot (\boldsymbol{u} \cdot \boldsymbol{\nabla}\gamma) - \nabla^\perp \cdot (\boldsymbol{\xi} \cdot \boldsymbol{\nabla} u) - \nu\, \nabla \cdot (\nabla\tilde{\rho}\, \partial_z \zeta) \\ & + \nu\, \nabla \cdot (\nabla\zeta\, \partial_z \tilde{\rho}) + 2\lambda\, \nabla \cdot (\nabla\rho\, \partial_z \nabla^\perp \cdot U_g) - 2\lambda\, \nabla \cdot (\nabla\nabla^\perp \cdot U_g\, \partial_z \rho) \\ & - 2\lambda\, \Omega \cdot \nabla^\perp\, \overline{\nabla\tilde{\rho} \cdot \nabla^\perp\nabla^\perp \cdot U_g} + \nu\, \Omega \cdot \nabla^\perp\, \overline{(\nabla\tilde{\rho} \cdot \nabla^\perp\zeta + \boldsymbol{\gamma} \cdot \boldsymbol{\nabla} u_3 - \hat{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\Delta\theta)}\,, \end{aligned} \tag{3.85c}$$

$$g = \nu\, \Delta(\boldsymbol{u} \cdot \boldsymbol{\nabla}\tilde{\rho}) + s\, \hat{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega + \nu\, \boldsymbol{\Omega} \cdot \boldsymbol{\nabla}(\boldsymbol{u} \cdot \boldsymbol{\nabla}\gamma_3 - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} u_3 - \nabla\tilde{\rho} \cdot \nabla^\perp\nabla^\perp \cdot B)\,. \tag{3.85d}$$

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.

3.9 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,

$$u = \nabla^\perp \psi + \nabla\phi\,. \tag{3.86}$$

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

$$\Delta\bar{\psi} = s\,\omega + \tfrac{c}{s}\,\partial_x \bar{u}_3 \,. \tag{3.87a}$$

Next, due to (3.83),

$$\Delta\Psi = \zeta \,. \tag{3.87b}$$

Finally, by incompressibility,

$$\Delta\phi = -\partial_z u_3 \tag{3.87c}$$

so that, altogether,

$$u = \nabla^\perp \bar{\psi} + \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \nabla^\perp \Psi + \nabla\phi \,. \tag{3.87d}$$

This expression for the horizontal components of the velocity field, first, proves that the kinematic balance relation (3.85) 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

$$\boldsymbol{\xi} \cdot \boldsymbol{\nabla} (\partial_x \bar{u}_3 + \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \zeta) \,, \tag{3.88}$$

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 (3.85) 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 (3.85), the reconstruction equations (3.87), the continuity equation in three dimensions,

$$\partial_t \rho + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho = 0 \,, \tag{3.89a}$$

and the evolution equation in two spatial dimensions for ω,

$$\partial_t \omega + \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega = \overline{\nabla \rho \cdot \nabla^\perp \nabla^\perp \cdot B} - \nu \, \overline{(\partial_z \nabla \theta \cdot \nabla u_3 - \Delta \theta \, \partial_z u_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} \Delta \theta)} \,. \tag{3.89b}$$

where B is defined in (3.61) and θ is the geostrophic stream function which depends on ρ via (3.39b).

3.10 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 semigeostrophic 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 semigeostrophic 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 (3.47). 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 (3.50), which formally cancels all terms at $O(\varepsilon)$. For the Euler–Boussinesq system, it is the last term in (3.47) 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 3.3 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 3.9. For balance models in two dimensions, potential vorticity based methods provide an alternative, numerically robust setting (Dritschel, Gottwald, and Oliver, 2017). As for primitive equation balance models, potential vorticity based numerics require the solution of a nonlinear elliptic equation (cf. Oliver and Vasylkevych, 2016; Akramov and Oliver, 2020). 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.

3.11 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 3.11.1. *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 $\bar{\phi} \circ \chi$ is independent of ζ, so that

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

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 $\chi(0)$ and $\chi(-s^{-1})$ 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 3.11.2. *Let ϕ be a function with $\phi \circ \chi(0) = \phi \circ \chi(-s^{-1})$ 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,

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

Under the condition stated, the boundary terms cancel. This implies (i). For (ii), we use Lemma 3.11.1 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. □

3.12 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 3.12.1. *Let $\boldsymbol{v} \in V_{\text{div}}$. 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 (3.30b) turns into

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

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 3.12.1.

Lemma 3.12.2. *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 $\bar{v}_3 \circ \chi$ is independent of ζ, this implies $\boldsymbol{\nabla} \cdot (\mathsf{A}^{-1} \bar{\boldsymbol{v}} \circ \chi) = 0$. By (3.30a), this implies that $\boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$. □

Corollary 3.12.3. *If $\boldsymbol{v} \in V_{\text{div}}$, then $\boldsymbol{\nabla} \cdot \hat{\boldsymbol{v}} = \boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$.*

Proof. Lemma 3.12.1 followed by Lemma 3.12.2 yields $\boldsymbol{\nabla} \cdot \bar{\boldsymbol{v}} = 0$. Then $\hat{\boldsymbol{u}} = \boldsymbol{u} - \bar{\boldsymbol{u}}$ is also divergence-free. □

Lemma 3.12.4. *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*

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

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 (3.29),

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

The third equality in (B.6) is due to (3.30b) and the last equality is due to $\boldsymbol{u} = \hat{\boldsymbol{u}} + \bar{\boldsymbol{u}}$, where $\bar{\boldsymbol{u}} \circ \chi$ is independent of ζ, so that the entire contribution from $\bar{\boldsymbol{u}}$ vanishes by Lemma 3.12.1. 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

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

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

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

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

The next lemma provides a converse statement to Lemma 3.12.4.

Lemma 3.12.5. *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 3.12.4, that is,*

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

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 3.12.4.

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

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

This implies that

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

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

3.13 Appendix C. Inner product identities for decomposed vector fields

Lemma 3.13.1. *Let* $\boldsymbol{u} \in V_{\mathrm{div}}$*; 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*

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

Proof. By Lemma 3.12.4,

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

where $\hat{g} \circ \chi$ 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

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

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 3.13.2. *Let $\boldsymbol{u} = \hat{\boldsymbol{u}} + \bar{\boldsymbol{u}} \in V_{\mathrm{div}}$ 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 3.12.4, $\mathsf{Q}\hat{\boldsymbol{u}} = \boldsymbol{\Omega}\hat{g}$ with

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

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

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

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 3.13.3. *Under the conditions of Lemma 3.12.5, 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 3.12.5, $\hat{\boldsymbol{u}}$ 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

$$\begin{aligned} \hat{u}_3 \circ \boldsymbol{\chi}(0) &= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s\, \hat{g} \circ \boldsymbol{\chi}(0) \\ &= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s\, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - s \int_{-s^{-1}}^{0} \boldsymbol{\nabla} \cdot (\mathsf{SP}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi})\, d\zeta \\ &= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + s\, \hat{g} \circ \boldsymbol{\chi}(-s^{-1}) - \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(0) + \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) \\ &= \boldsymbol{k} \cdot \mathsf{P}\hat{\boldsymbol{u}} \circ \boldsymbol{\chi}(-s^{-1}) + s\, \hat{g} \circ \boldsymbol{\chi}(-s^{-1})\,. \end{aligned} \tag{C.10}$$

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 3.12.2, 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 3.13.4. *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 3.12.1, $\nabla \cdot (\mathsf{S}_h \mathsf{P} \bar{\boldsymbol{u}}) = 0$ 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 Range 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

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

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

3.14 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 (3.53) and the representation of $\mathsf{Q}\bar{\boldsymbol{v}}$ via (3.52), we compute:

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

Further, inserting (3.50), 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 χ commute, we obtain

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

Inserting the thermal wind relation (3.36), integrating by parts with respect to ζ, and changing variables, we continue the computation:

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

where, in the last step, we have made use of Lemma 3.13.1, Lemma 3.13.2, and the fact that $\boldsymbol{k} \cdot \boldsymbol{u}_g = 0$.

Chapter 4

Variational Balance model for the equatorial long-wave dynamics

This chapter is in preparation for the submission.

Variational Balance model for the equatorial long-wave dynamics

Abstract

We consider the motion of the rotating fluid in the long-wave scaling regime. It is governed by the three-dimensional Boussinesq equations on the equatorial β-plane. The full Coriolis force is taken into account to keep the rotational effect on the model because the axis of rotation of the Earth aligns the horizontal plane at the equator. This is contrary to the so-called "traditional approximation" in which the horizontal component of the Coriolis vector is zero. The model is constructed via variational principle which depends on the Lagrangian dynamics. The reduced model is called as L_1 model which is derived firstly by Salmon (1985) using Hamilton's principles. On the construction, the similar strategy to that Oliver and Vasylkevych (2016) is followed to derive Euler-Poincare equations. Derivations are done considering the anisotropy on the horizontal scaling. In the long term, we aim to obtain the balance model which includes slow equatorial long-wave dynamics.

Keywords

equatorial β-plane, Boussinesq equations, full Coriolis force, variational approximation, long-wave dynamics

4.1 Introduction

We consider the Euler–Boussinesq equations on the β-plane,

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\frac{1}{\rho_0} \nabla p - \frac{g\rho}{\rho_0} \boldsymbol{k} \,, \tag{4.1a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \,, \tag{4.1b}$$

$$\partial_t \rho + \boldsymbol{u} \cdot \nabla \rho = 0 \,, \tag{4.1c}$$

where $\boldsymbol{u}$ is the three-dimensional velocity field,

$$2\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ f_0 \\ \beta y \end{pmatrix} \tag{4.2}$$

is the angular velocity describing the rotation of the Earth on the equatorial β-plane, ρ_0 a constant reference density, p the departure from hydrostatic pressure, g the constant of gravity, ρ the departure from the constant reference density, and $\boldsymbol{k}$ the unit vector in the vertical. We write $\boldsymbol{x} = (x, y, z)$ with x the zonal, and y the meridional coordinate, and the velocity is $\boldsymbol{u} = (u_1, u_2, u_3)$.

The domain is defined as $\mathcal{D} = \mathbb{T} \times [-\infty, \infty] \times [-\infty, 0]$ on which we have infinite layer depth, and similarly we have infinite boundary in the meridional direction for simplicity. In the zonal direction, we have periodic boundary and

$$u_2 = 0 \quad \text{as} \quad y \to \pm\infty\,, \tag{4.3a}$$

$$u_3 = 0 \quad \text{at} \quad z = 0\,, \tag{4.3b}$$

which implies that we have rigid lid approximation on the surface, and we assume that the meridional velocity decays sufficiently rapid towards the boundary.

We apply the variational principle to the reduced Lagrangian (e.g. Franzke et al., 2019)

$$\ell = \int_{\mathcal{D}} \boldsymbol{R} \cdot \boldsymbol{u} + \frac{1}{2}\,|\boldsymbol{u}|^2 - \frac{g}{\rho_0}\,\rho z\, d\boldsymbol{x} \tag{4.4}$$

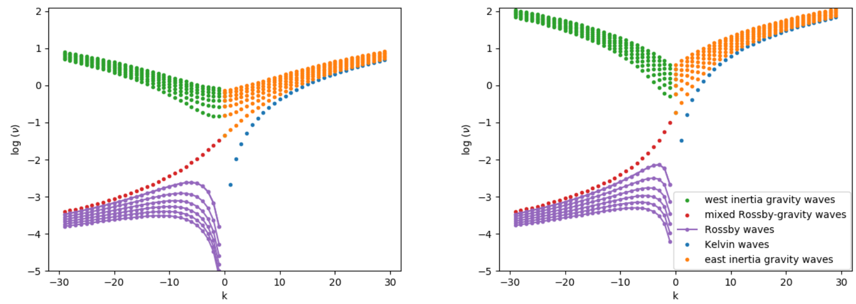
on the Lie algebra of divergence free vector fields. Here, $\boldsymbol{R}$ is a vector potential for the Coriolis vector which satisfies $2\boldsymbol{\Omega} = \boldsymbol{\nabla} \times \boldsymbol{R}$. A convenient choice is

$$\boldsymbol{R} = (f_0\, z - \tfrac{1}{2}\,\beta\, y^2)\,\boldsymbol{i}\,, \tag{4.5}$$

where $\boldsymbol{i}$ is the unit vector in the zonal direction. Alternatively, we can encode the incompressibility constraint as a Lagrange multiplier, so that

$$\ell = \int_{\mathcal{D}} \boldsymbol{R} \cdot \boldsymbol{u} + \frac{1}{2}\,|\boldsymbol{u}|^2 - \frac{g}{\rho_0}\,\rho z\, d\boldsymbol{x} + \int_{\mathcal{D}} p\, \boldsymbol{\nabla} \cdot \boldsymbol{u}\, d\boldsymbol{x}\,. \tag{4.6}$$

The dispersion relation for the equatorial waves on the β-plane was shown by Matsuno (1966). We illustrate it in Figure 4.1 using TIGAR which solves the shallow water equation on the globe (Vasylkevych and Žagar, 2021). The scale separation is clearly seen between Rossby waves and inertia-gravity waves. However, two extra waves appear in this region, namely Kelvin wave and mixed Rossby-gravity wave. Both waves fills the frequency gap between inertia-gravity waves and Rossby waves. For instance, Kelvin waves span the whole frequency range. Similarly, mixed Rossby-gravity waves with positive wavenumbers move fast as inertia-gravity waves and the rest with negative wavenumbers are slow as Rossby waves. As a consequence of this, some waves are in the same frequency range, but for a certain equivalent height.



(A) $D = 400m$ (B) $D = 4000m$

FIGURE 4.1: Dispersion curves of linear wave solutions at different equivalent heights for a shallow water model.

One of the main characteristic elements to decide which waves appear on the balance model is the equivalent height for shallow water models. The range of the frequency gap varies with the equivalent height as shown in Figure 4.1 and detail analysis is done by Žagar et al. (2015). Then, it is expected to have Kelvin waves but with mixed Rossby-gravity waves on a very shallow layer (Wheeler and Kiladis, 1999). Besides, these waves propagate vertically (Wallace and Kousky, 1968; Yanai and Maruyama, 1966) and so the three-dimensional model might be an advantage to capture them.

The scaling around the equator is a challenge for reduced models because of Kelvin waves and mixed Rossby-gravity waves. In addition, Charney (1948) is the pioneer of the quasi-geostrophic scaling on the mid-latitude which is a very powerful tool in the scale analysis. Rossby number is used as a small parameter for the scale separation to obtain a balance model there, but it vanishes due to the vertical component of the Coriolis vector on the equator. Thus, it is not applicable in the vicinity of the equator. On the other hand, a reduced model can be obtain without using any scaling parameter as Theiss and Mohebalhojeh (2009) obtained one via non-linear normal mode initialization. On their model, there is no Kelvin wave, but Rossby wave with mixed-Rossby-gravity wave on the generalized model. They mention that Kelvin wave can be obtained as a separate model as Gill (1982) declared because it is characterized by having zero meridional velocity. They claimed that their model is suitable to extend to a bounded, forced and stratified fluids.

Chan and Shepherd (2013) applied the long-wave scaling to the weakly nonlinear shallow water equation. They obtained Rossby waves and Kelvin waves together in their model excluding mixed Rossby-gravity and inertia-gravity waves. They used the slaving method by which the prognostic equation variable is based on the height field. Thus, it is not slow as in potential vorticity (PV) based models. This helps to keep Kelvin wave in their model because it is invisible in PV-based balance models (Hoskins, McIntyre, and Robertson, 1985). In addition to this weakly non-linear model, they showed two more cases. One is for a fully linear shallow water equations and the other one is for the stratified flow with primitive equations.

We infer that equatorial waves differ into two main categories as fast and slow by their speed. Fast motions are split off regarding the direction of the motion east/west. Slow motions are separated as rotational and divergent. In a general view, balance models include only the slow part of this dynamics. Because of the geometry, we work on the infinite depth and our model covers the vertical motion. We expect to obtain a balance model covering the slow part of the dynamics.

The non-traditional terms have a central role in equatorial wave dynamics. Hence, we have quite large motivation to work on this region with the full Coriolis force. From this perspective, Dutrifoy and Majda (2006) have a new PDE analysis using raising and lowering operator which are defined for the quantum harmonic oscillator. Their linear long-wave regime provides the existence of the solution. Additionally, these terms are used to explain some phenomena in the atmosphere. For instance, Kohma and Sato (2013b) used it to examine atmospheric waves in the equator. Hayashi and Itoh (2012) used it with "quasi-hydrostatic" approximation in the scaling (on spherical coordinates) following Dellar (2011) to explain some important circulations like Madden–Julian oscillation. They imply that non-traditional components might interact with the diabatic heating and these terms might be helpful to explain such influential phenomena. Besides, these phenomena are related with equatorial waves mainly like Kelvin waves (Dunkerton, 1995) and so corresponding works definitely reveal the wave generation for this specific region. Similarly, Ong and Roundy (2019) show the behavior of the zonally symmetric heating source to mimics Intertropical Convergence Zone (ITCZ) for the large scale flows with/without non-traditional Coriolis term on their model. It is constructed via anelastic equations set on the β-plane forced by a heating source. Therefore, it is worth to include horizontal component of the Coriolis vector into the model.

Verdière and Schopp (1994) claim that hydrostatic approximation can be used for the upper part of the ocean (above characteristics) and for the rest non-hydrostatic approximation must be used with the full Coriolis force. Similarly, Kuo (1977) explains some scaling relationship for the upper part of the atmosphere and ocean. He implies that there is no couple model for Kelvin waves and it exists independently in both atmosphere and ocean. Then, we go with the non-traditional approximation for non-hydrostatic flows. While doing this, we take the ratio between Coriolis vector components, κ given in (1.66) in our model, large as suggested by Verdière and Schopp (1994).

In our previous paper (Özden and Oliver, 2021), we observed the effect of the non-zero horizontal component of the Coriolis vector on the balance model for a stratified fluid. In the present draft, we investigate the variational model reduction in the long-wave regime on the equatorial β-plane. We expect the reduced model is simpler than the mid-latitude model, but it will bring another complexity as a consequence of the singularity. Here, the special case comes from the characteristics which do not intercept the bottom boundary. This work is important to complete the work on characteristics because derivations change completely in the equator.

Yano and Bonazzola (2009) performed the scale analysis considering a parameter which is the ratio of the acceleration of the flow to the Coriolis force on their diabatic models. Then, the horizontal length scale is the main character on their scaling options. To have the balance on the large scale, they claimed that the initial conditions should be chosen as a strong constraint. Otherwise, thermodynamic balance is well-posed only initially. The scaling on our model is done considering their specified parameter as large on which the rotation is dominated on the dynamics. Specifically, we add the anisotropy as a small parameter because Rossby number vanishes.

We use the variational principle to obtain the equatorial balance model. This principle is based on the Lagrangian dynamics. It was identified by Salmon (1982; 1985), using Hamilton's principles. In our derivations, equations of motion are obtained following the similar strategy to that Oliver and Vasylkevych (2016). The reduced model is called L_1 model. In our derivations, the anisotropy on the horizontal scale is taken as a small parameter because it is one of the main characteristic features on the equator. We choose Case 2 which is one of the scaling options defined in the main part of the thesis for the derivations. We suggest to work on other cases for the theoretical interest.

The Section 4.2 provides a brief derivation on how to obtain the equations of motion via variational principle for Case 2. In Section 4.3, we introduce the notation considering the axis of rotation. Next, we show the geostrophic velocity under the chosen scaling in Section 4.4. Truncated Lagrangian is given in Section 4.5. In Section 4.6, we apply the variational asymptotic to obtain L_1 model for Case 2. In Section 4.7, we show how the conservation laws work on the chosen dynamics. Section 4.8 gives the conclusion and the outlook for the problem.

4.2 Derivations for Case 2

We rewrite the momentum equations

$$\hat{\varepsilon}\,(\partial_t u_1 + \boldsymbol{u}\cdot\boldsymbol{\nabla} u_1) - y\,u_2 + \kappa\,u_3 = -\partial_x p\,, \tag{4.7a}$$

$$\hat{\varepsilon}\,\varepsilon^2\,(\partial_t u_2 + \boldsymbol{u}\cdot\boldsymbol{\nabla} u_2) + y\,u_1 = -\partial_y p\,, \tag{4.7b}$$

$$\hat{\varepsilon}\,\varepsilon^2\,\alpha^2\,(\partial_t u_3 + \boldsymbol{u}\cdot\boldsymbol{\nabla} u_3) - \kappa\,u_1 = -\partial_z p - \rho\,. \tag{4.7c}$$

In the leading order, we obtain

$$\boldsymbol{\Omega}\times\boldsymbol{u} = -\boldsymbol{\nabla} p - \rho\boldsymbol{k}\,, \tag{4.8}$$

where the vector in the direction of rotation axis is

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 \\ \kappa \\ y \end{pmatrix} . \tag{4.9}$$

The corresponding Lagrangian reads

$$\begin{aligned} \ell = & \int_{\mathcal{D}} R\, u_1 + \frac{1}{2}\,\hat{\varepsilon}\left(u_1^2 + \varepsilon^2\, u_2^2 + \alpha^2\,\varepsilon^2\, u_3^2\right) - \rho\, z\, d\boldsymbol{x} \\ & + \frac{VP}{L}\int_{\mathcal{D}} p\,\boldsymbol{\nabla}\cdot\boldsymbol{u}\, d\boldsymbol{x}\,, \end{aligned} \tag{4.10}$$

where $R = \kappa\, z - \frac{1}{2}\, y^2$.

Since the incompressiblity condition is unchanged by the scaling, we can easily incorporate incompressibility by restricting to flow map to be volume preserving. Therefore, we do not need the Lagrange multiplier in the Lagrangian. Further, dropping terms beyond first order and renaming $\hat{\varepsilon}$ to ε, we obtain

$$\ell = \int_{\mathcal{D}} \boldsymbol{R}\cdot\boldsymbol{u} + \frac{\varepsilon}{2}\, u_1^2 - \rho\, z\, d\boldsymbol{x}\,. \tag{4.11}$$

For the simplest model, we take the variation of (4.11) in the leading order with the Lin constraints $\delta\boldsymbol{u} = \dot{\boldsymbol{v}} + \boldsymbol{\nabla}\boldsymbol{v}\,\boldsymbol{u} - \boldsymbol{\nabla}\boldsymbol{u}\,\boldsymbol{v}$ and $\delta\rho = -\boldsymbol{u}\cdot\boldsymbol{\nabla}\rho$. Here, we set the parameter given in (1.62) as $\gamma = 1$. Then, we have

$$\begin{aligned} \delta\ell &= \int_0^{t_1}\int_{\mathcal{D}} \boldsymbol{R}\cdot\delta\boldsymbol{u} + \varepsilon\,\delta\boldsymbol{u}_1\cdot\boldsymbol{u}_1 - \delta\rho\, z\, d\boldsymbol{x}\, dt \\ &= \int_0^{t_1}\int_{\mathcal{D}} \boldsymbol{R}\cdot\dot{\boldsymbol{v}} + \boldsymbol{R}\cdot\boldsymbol{\nabla}\boldsymbol{v}\,\boldsymbol{u} - \boldsymbol{R}\cdot\boldsymbol{\nabla}\boldsymbol{u}\,\boldsymbol{v} + \varepsilon\,\boldsymbol{u}_1\cdot\dot{\boldsymbol{v}} + \varepsilon\,\boldsymbol{u}_1\cdot\boldsymbol{\nabla}\boldsymbol{v}\,\boldsymbol{u} \\ &\qquad\qquad - \varepsilon\,\boldsymbol{u}_1\cdot\boldsymbol{\nabla}\boldsymbol{u}\,\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\rho\, z\, d\boldsymbol{x}\, dt \\ &= -\int_0^{t_1}\int_{\mathcal{D}} \big((\boldsymbol{\nabla}\boldsymbol{R}^{\mathsf{T}} - \boldsymbol{\nabla}\boldsymbol{R})\boldsymbol{u} + \varepsilon\,\dot{\boldsymbol{u}}_1 + \varepsilon\,(\boldsymbol{\nabla}\boldsymbol{u}_1^{\mathsf{T}} - \boldsymbol{\nabla}\boldsymbol{u}_1)\boldsymbol{u} + \rho\boldsymbol{k}\big)\cdot\boldsymbol{v}\, d\boldsymbol{x}\, dt \\ &= -\int_0^{t_1}\int_{\mathcal{D}} \big(\partial_t(\boldsymbol{R} + \varepsilon\,\boldsymbol{u}_1) + (\boldsymbol{\nabla}\times(\boldsymbol{R} + \varepsilon\,\boldsymbol{u}_1))\times\boldsymbol{u} + \rho\boldsymbol{k}\big)\cdot\boldsymbol{v}\, d\boldsymbol{x}\, dt\,, \end{aligned} \tag{4.12}$$

where $\boldsymbol{u}_1 = (u_1, 0, 0)$ and

$$\boldsymbol{m} = \frac{\delta\ell}{\delta\boldsymbol{u}} = \boldsymbol{R} + \varepsilon\,\boldsymbol{u}_1\,, \quad \text{and} \quad \frac{\delta\ell}{\delta\rho} = -z\,. \tag{4.13}$$

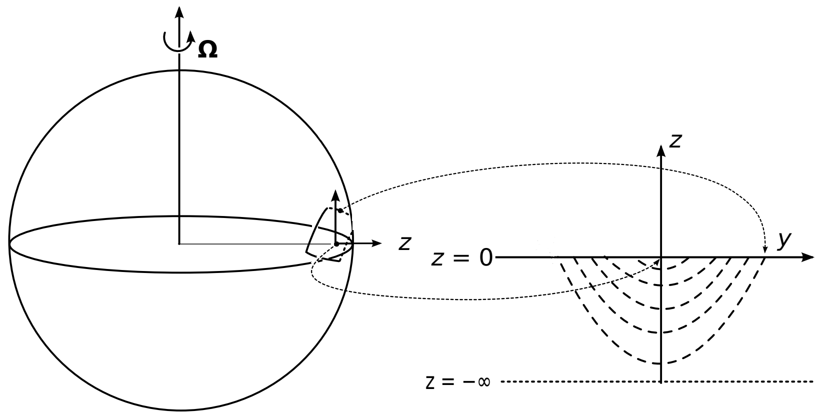


FIGURE 4.2: Geometry of the β-plane approximation around the equator with the full Coriolis vector. Here, x, y, and z are directed toward the east, the north, and the upward, respectively. The dashed lines represent the characteristics of the thermal wind relation, inspired from Verdière and Schopp (1994), and the axis of rotation is tangent to them.

Thus, the Euler–Poincaré equation reads

$$\varepsilon\,(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u}_1 + \boldsymbol{\Omega}\times\boldsymbol{u} = -\boldsymbol{\nabla}p - \rho\boldsymbol{k}\,, \tag{4.14}$$

which coincides with (4.7) up to terms $\mathcal{O}(\varepsilon)$. Taking the curl of (4.14), noting that $\boldsymbol{\Omega}$ and $\boldsymbol{u}$ are divergence free, we find

$$-\varepsilon\begin{pmatrix} 0 \\ \nabla^{\perp}_{\mathrm{m}}(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})u_1 \end{pmatrix} + \boldsymbol{u}\cdot\boldsymbol{\nabla\Omega} - \boldsymbol{\Omega}\cdot\boldsymbol{\nabla u} = \begin{pmatrix} \nabla^{\perp}\rho \\ 0 \end{pmatrix}, \tag{4.15}$$

where $\nabla^{\perp} = (-\partial_y, \partial_x)$ and $\nabla^{\perp}_{m} = (-\partial_z, \partial_y)$.

4.3 Notation on the tilted axis

While the geostrophic balance holds the horizontal velocity in the mid-latitude, this situation vanishes in the equator. The leading order gives the dominant balance only in the meridional direction. For the solution of (4.33), we need to integrate the velocity field along characteristics which are shown in 4.2. These lines are not in the same direction as the axis of rotation. Additionally, they are not straight as we have on $f-$plane model because characteristics are deformed when we transform them to the Cartesian coordinate.

Around the equator, there are four types of characteristics having own specific features. Two of them intercept the bottom boundary. Then, the zero-flux boundary condition is valid for the flow at the bottom and at the top. However, these two types appear on each side of the equator, respectively. The other type of characteristics is the equator itself. It appears as a point on the meridional-vertical plane. In this special case, $y = 0$. The last type is defined for lines which are very close to the equator and so the value of y is very small. The main feature of this type is that the characteristics do not intercept the bottom boundary in the vicinity of the equator. Then, the derivation of the model is different from the mid-latitude model (Özden and Oliver, 2021).

We are mainly interested in the dynamics near the equator. Thus, the characteristics that we work on are curves that have vertical boundaries only on the surface.

We repeat the parameterization of the characteristics as we did in the previous paper (Özden and Oliver, 2021). The corresponding mapping is

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \boldsymbol{\chi}(\boldsymbol{\xi}) = \begin{pmatrix} \xi \\ \kappa\,\eta \\ \zeta + \frac{1}{2}\,\kappa\,\eta^2 \end{pmatrix}, \tag{4.16}$$

where η is a parameter along the characteristic curve and ζ is its depth at the equator. To keep the Jacobian is free from the coordinate, we use the meridional coordinate for the scaling. The characteristic lines will touch the surface at

$$\eta_{\pm}(\eta) = \pm\sqrt{\tfrac{2}{\kappa}\,|\zeta|}\,. \tag{4.17}$$

Then

$$\partial_\xi(f \circ \boldsymbol{\chi}) = (\partial_x f) \circ \boldsymbol{\chi}\,, \tag{4.18a}$$

$$\partial_\zeta(f \circ \boldsymbol{\chi}) = (\partial_z f) \circ \boldsymbol{\chi}\,, \tag{4.18b}$$

$$\partial_\eta(f \circ \boldsymbol{\chi}) = \kappa\,(\partial_y f) \circ \boldsymbol{\chi} + \kappa\,\eta\,(\partial_z f) \circ \boldsymbol{\chi} = (\kappa\,\partial_y f) \circ \boldsymbol{\chi} + (y\,\partial_z f) \circ \boldsymbol{\chi} = (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} f) \circ \boldsymbol{\chi}\,. \tag{4.18c}$$

For the general representation, any scalar function f with the mapping defined in (4.16) satisfies

$$\nabla(f \circ \chi) = (\mathsf{A}^{\mathsf{T}} \nabla f) \circ \chi\,, \tag{4.19a}$$

$$(\nabla f) \circ \chi = \mathsf{A}^{-\mathsf{T}} \circ \chi \, \nabla(f \circ \chi)\,, \tag{4.19b}$$

where the matrix A is

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & y & 1 \end{pmatrix}, \tag{4.19c}$$

and we note that $\det \mathsf{A} = \kappa$. For the vector field $\boldsymbol{v}$,

$$\nabla \cdot (\boldsymbol{v} \circ \chi) = (\mathsf{A} : \nabla \boldsymbol{v}^T) \circ \chi\,, \tag{4.20}$$

and

$$(\nabla \cdot \boldsymbol{v}) \circ \chi = \mathsf{A}^{-1} : \nabla(\boldsymbol{v}^T \circ \chi)\,. \tag{4.21}$$

We define the average of a function f along the characteristics that crosses the equator at $(\xi, 0, \zeta)$ via

$$\bar{f} \circ \chi = \fint_{\eta_-(\zeta)}^{\eta_+(\zeta)} f \circ \chi \, d\eta\,, \tag{4.22}$$

where

$$\fint_a^b \phi \, d\eta = \frac{1}{b-a} \int_a^b \phi \, d\eta\,. \tag{4.23}$$

We now introduce an operator Λ acting on vector fields via $\Lambda \boldsymbol{u} = \nabla \times (\boldsymbol{\Omega} \times \boldsymbol{u})$ which, if $\boldsymbol{u}$ is divergence free, reads

$$\Lambda \boldsymbol{u} = \boldsymbol{u} \cdot \nabla \boldsymbol{\Omega} - \boldsymbol{\Omega} \cdot \nabla \boldsymbol{u}\,. \tag{4.24}$$

We now show that $\boldsymbol{u} \in \operatorname{Ker} \Lambda$ if and only if $\boldsymbol{u} = \bar{\boldsymbol{u}}$ with $u_2 = 0$. Suppose first that $\boldsymbol{u} \in \operatorname{Ker} \Lambda$. Then, for the horizontal components, $\Lambda_{\mathrm{h}} \boldsymbol{u} = \boldsymbol{\Omega} \cdot \nabla u = 0$, so $u = \bar{u}$. For the vertical component, we note that

$$0 = \overline{\Lambda_3 \boldsymbol{u}} = \int_{\eta_-(\zeta)}^{\eta_+(\zeta)} u_2 \circ \chi - (\boldsymbol{\Omega} \cdot \nabla u_3) \circ \chi \, d\eta\,. \tag{4.25}$$

The second term under the integral does not contribute and so we have $\bar{u}_2 = 0$ because $u_3 = 0$ at the vertical boundaries. This implies that $u_2 = \bar{u}_2$. Therefore, we also have $u_2 = 0$ and further, from the vertical component equation, $u_3 = \bar{u}_3$. This proves the forward implication. The converse implication is obvious.

Since $\hat{\boldsymbol{u}}$ is mean-free,

$$\int_{\eta_-(\zeta)}^{\eta_+(\zeta)} \eta\, \partial_\eta (\hat{\boldsymbol{u}} \circ \chi)\, d\eta = \eta_+(\zeta)\, \hat{\boldsymbol{u}} \circ \chi(\eta_+(\zeta)) - \eta_-(\zeta)\, \hat{\boldsymbol{u}} \circ \chi(\eta_-(\zeta))\,. \tag{4.26}$$

We use the following lemma on integration by parts along the characteristics.

Lemma 4.3.1. *Let f an arbitrary function and $\hat{g}$ is mean free along characteristics. Then,*

$$\int_{\mathcal{D}} f\, T[\hat{g}]\, d\boldsymbol{x} = -\int_{\mathcal{D}} T[f]\, \hat{g}\, d\boldsymbol{x}\,, \tag{4.27}$$

where T denotes the antiderivative along characteristics, defined as

$$T[f] \circ \chi = \int_{\eta_+(\zeta)}^{\eta} f \circ \chi'\, d\eta'\,. \tag{4.28}$$

Proof. Integrating by parts in $\boldsymbol{\xi}$-coordinates, we compute

$$\begin{aligned}
\int_{\mathcal{D}} f\, T[\hat{g}]\, d\boldsymbol{x} &= \kappa \int_{\mathbb{T}} \int_{-\infty}^{0} \int_{\eta_-(\zeta)}^{\eta_+(\zeta)} f \circ \chi\, T[\hat{g}] \circ \chi\, d\eta\, d\zeta\, d\xi \\
&= \kappa \int_{\mathbb{T}} \int_{-\infty}^{0} T[f] \circ \chi\, T[\hat{g}] \circ \chi \Big|_{\eta=\eta_-(\zeta)}^{\eta=\eta_+(\zeta)} d\zeta\, d\xi \\
&\quad - \kappa \int_{\mathbb{T}} \int_{-\infty}^{0} \int_{\eta_-(\zeta)}^{\eta_+(\zeta)} T[f] \circ \chi\, \hat{g} \circ \chi\, d\eta\, d\zeta\, d\xi \\
&= -\int_{\mathcal{D}} T[f]\, \hat{g}\, d\boldsymbol{x}\,,
\end{aligned} \tag{4.29}$$

where boundary terms vanish so long as $\hat{g}$. □

Remark 1. *If the function f has an odd symmetry with respect to the equator, i.e,*

$$\int_0^{\eta_+(\zeta)} f \circ \chi\, d\eta = -\int_{\eta_-(\zeta)}^{0} f \circ \chi\, d\eta\,, \tag{4.30}$$

then it is mean-free along the characteristics by the definition in (4.22)

$$
\begin{aligned}
0 &= \int_0^{\eta_+(\zeta)} f \circ \chi \, d\eta + \int_{\eta_-(\zeta)}^0 f \circ \chi \, d\eta \\
&= \int_{\eta_-(\zeta)}^{\eta_+(\zeta)} f \circ \chi \, d\eta \,,
\end{aligned} \tag{4.31}
$$

or

$$
T[f] \circ \chi(\eta_-(\zeta)) = 0 \,, \tag{4.32}
$$

which means that it vanishes at the boundary.

4.4 Geostophic velocity

The leading order of (4.15) reads $\Lambda \boldsymbol{u} = \boldsymbol{u} \cdot \boldsymbol{\nabla\Omega} - \boldsymbol{\Omega} \cdot \boldsymbol{\nabla u} = -\boldsymbol{\nabla} \times (\rho \boldsymbol{k})$. In components,

$$
-\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} u = \nabla^\perp \rho \,, \tag{4.33a}
$$

$$
-\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} u_3 + u_2 = 0 \,. \tag{4.33b}
$$

Integrating (4.33a) along characteristics, we obtain

$$
\hat{u}_g \circ \chi = \hat{u}_g \circ \chi(\eta_+(\zeta)) + \int_\eta^{\eta_+(\zeta)} (\nabla^\perp \rho) \circ \chi' \, d\eta' \,. \tag{4.34}
$$

Setting $\theta = -T[\rho]$ so that $-\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta = \rho$, we can express the geostrophic velocity as

$$
\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \hat{u}_g = \nabla^\perp (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) = \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \nabla^\perp \theta - \boldsymbol{i} \, \partial_z \theta \,. \tag{4.35}
$$

Integrating this expression along characteristics, we obtain

$$
\hat{u}_g = \nabla^\perp \theta - T[\partial_z \theta] \, \boldsymbol{i} + \phi \,, \tag{4.36}
$$

where ϕ is constant along characteristics. Therefore, the geostrophic velocity is not automatically mean-free and so θ is not mean-free as we have on the f-plane model.

4.5 Truncated Lagrangian

We adapt variational framework developed by Oliver (2006) and extended to the f-plane Boussinesq equations (Özden and Oliver, 2021) to the equatorial β-plane here. We expand all terms up to $\mathcal{O}(\varepsilon)$, dropping contributions at $\mathcal{O}(\varepsilon^2)$ without further mention. We write

$$\boldsymbol{u}_\varepsilon = \boldsymbol{u} + \varepsilon\, \boldsymbol{u}'\,, \tag{4.37a}$$

$$\rho_\varepsilon = \rho + \varepsilon\, \rho'\,, \tag{4.37b}$$

where the first order terms are subject to the Lin constraints $\boldsymbol{u}' = \dot{\boldsymbol{v}} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{v} - \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{u}$ and $\rho' = -\boldsymbol{v}\cdot\boldsymbol{\nabla}\rho$, where $\boldsymbol{v}$ is the vector field generating the transformation. It is assumed to be divergence free and to leave the domain invariant. Then,

$$\begin{aligned} L_\varepsilon &= \int_{\mathcal{D}} \boldsymbol{R}\cdot\boldsymbol{u}_\varepsilon + \frac{1}{2}\,\varepsilon\,(\boldsymbol{i}\cdot\boldsymbol{u}_\varepsilon)^2 - \rho_\varepsilon\, z\, d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \boldsymbol{R}\cdot\boldsymbol{u} - \rho\, z\, d\boldsymbol{x} + \varepsilon \int_{\mathcal{D}} \boldsymbol{R}\cdot\boldsymbol{u}' + \frac{1}{2}\, u_1^2 - \rho'\, z\, d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \boldsymbol{R}\cdot\boldsymbol{u} - \rho\, z\, d\boldsymbol{x} + \varepsilon \int_{\mathcal{D}} (\boldsymbol{\nabla}\boldsymbol{R}^{\mathsf{T}} - \boldsymbol{\nabla}\boldsymbol{R})\,\boldsymbol{u}\cdot\boldsymbol{v} + \frac{1}{2}\, u_1^2 + z\,\boldsymbol{v}\cdot\boldsymbol{\nabla}\rho\, d\boldsymbol{x} \\ &\equiv L_0 + \varepsilon\, L_1\,. \end{aligned} \tag{4.38}$$

Since $\boldsymbol{R} = (\kappa z - \frac{1}{2}y^2, 0, 0)$, cf. (4.10). Introducing the skew-symmetric matrix

$$\mathsf{J} = \boldsymbol{\nabla}\boldsymbol{R} - \boldsymbol{\nabla}\boldsymbol{R}^{\mathsf{T}} = \begin{pmatrix} 0 & \partial_y R_x & \partial_z R_x \\ -\partial_y R_x & 0 & 0 \\ -\partial_z R_x & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -y & \kappa \\ y & 0 & 0 \\ -\kappa & 0 & 0 \end{pmatrix}, \tag{4.39}$$

and integrating by parts in the last term in (4.38), we can write

$$L_1 = \int_{\mathcal{D}} \boldsymbol{u}\cdot\mathsf{J}\boldsymbol{v} + \frac{1}{2}\, u_1^2 - \rho\,\boldsymbol{k}\cdot\boldsymbol{v}\, d\boldsymbol{x}\,. \tag{4.40}$$

We seek to determine $\boldsymbol{v}$ such that L_1 becomes affine in $\hat{u}_1$. This requires that $\mathsf{J}\boldsymbol{v} = -\frac{1}{2}\,\hat{u}_1\,\boldsymbol{i}$ up to terms that do not depend on $\hat{u}_1$. We choose

$$\mathsf{J}\boldsymbol{v} = -\frac{1}{2}\,\hat{u}_1\,\boldsymbol{i} + \lambda\,\hat{u}_{g_1}\,\boldsymbol{i}\,, \tag{4.41}$$

so that $v_1 = 0$ and

$$- y\, v_2 + \kappa\, v_3 = -\frac{1}{2}\, \hat{u}_1 + \lambda\, \hat{u}_{g_1}\,, \tag{4.42}$$

where λ is the free parameter. We refer Oliver and Vasylkevych, 2016 for the detail analysis on the effect of λ.

A solution is given by

$$\boldsymbol{v} = \frac{1}{2} \begin{pmatrix} 0 \\ -\partial_z \psi \\ \partial_y \psi \end{pmatrix} - \lambda \begin{pmatrix} 0 \\ -\partial_z \psi_g \\ \partial_y \psi_g \end{pmatrix}, \tag{4.43}$$

where

$$\psi \circ \chi = -\int_{\eta_+(\zeta)}^{\eta} \hat{u}_1 \circ \chi'\, d\eta' = -T[\hat{u}_1] \circ \chi\,, \tag{4.44}$$

and

$$\psi_g \circ \chi = -\int_{\eta_+(\zeta)}^{\eta} \hat{u}_{g_1} \circ \chi'\, d\eta' = -T[\hat{u}_{g_1}] \circ \chi\,, \tag{4.45}$$

where T is defined in (4.28). As required, $\boldsymbol{v}$ is divergence free by the construction and the usage of the stream function on the construction of $\boldsymbol{v}$ guarantees this property. Moreover, ψ and ϕ are zero on the top boundaries, so that $\boldsymbol{k} \cdot \boldsymbol{v} = 0$ on the surface. Similarly, they are vanishing towards the meridional boundaries, $\boldsymbol{j} \cdot \boldsymbol{v} = 0$ as $T[\hat{u}_1]$ and $T[\hat{u}_{g_1}]$ are zero at the meridional boundaries. Therefore, $\hat{\boldsymbol{u}}$ tends to zero fast enough as $y \to \infty$ in the infinite domain.

With this choice, the kinetic energy and rotational contribution to the L_1-Lagrangian (4.40) read

$$\int_{\mathcal{D}} \boldsymbol{u} \cdot \mathsf{J}\boldsymbol{v} + \tfrac{1}{2}\, u_1^2\, d\boldsymbol{x} = \int_{\mathcal{D}} \lambda\, \hat{u}_1\, \hat{u}_{g_1} + \tfrac{1}{2}\, \bar{u}_1^2\, d\boldsymbol{x}\,. \tag{4.46}$$

The potential energy term is

$$\begin{aligned} -\int_{\mathcal{D}} \rho\, \boldsymbol{k} \cdot \boldsymbol{v}\, d\boldsymbol{x} &= -\int_{\mathcal{D}} \rho \left(\tfrac{1}{2}\, \partial_y \psi - \lambda\, \partial_y \psi_g \right) d\boldsymbol{x} \\ &= \int_{\mathcal{D}} \rho \left(\tfrac{1}{2}\, \partial_y T[\hat{u}_1] - \lambda\, \partial_y T[\hat{u}_{g_1}] \right) d\boldsymbol{x} \\ &= -\int_{\mathcal{D}} \partial_y \rho \left(\tfrac{1}{2}\, T[\hat{u}_1] - \lambda\, T[\hat{u}_{g_1}] \right) d\boldsymbol{x} \\ &= \int_{\mathcal{D}} T[\partial_y \rho] \left(\tfrac{1}{2}\, \hat{u}_1 - \lambda\, \hat{u}_{g_1} \right) d\boldsymbol{x}\,. \end{aligned} \tag{4.47}$$

We have applied the integration by parts on the third line. The boundary term yields zero because $T[\hat{u}_1]$ and $T[\hat{u}_{g_1}]$ vanish at the boundary. For the last equality, we have used Lemma 4.3.1. Now, inserting the first component of the thermal wind relation (4.33a), we continue:

$$
\begin{aligned}
-\int_{\mathcal{D}} \rho\, \boldsymbol{k} \cdot \boldsymbol{v}\, d\boldsymbol{x} &= \int_{\mathcal{D}} T[\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \hat{u}_{g_1}] \left(\tfrac{1}{2}\, \hat{u}_1 - \lambda\, \hat{u}_{g_1}\right) d\boldsymbol{x} \\
&= \int_{\mathcal{D}} \tfrac{1}{2}\, \hat{u}_{g_1} \hat{u}_1 - \lambda\, \hat{u}_{g_1}^2\, d\boldsymbol{x}\,.
\end{aligned}
\tag{4.48}
$$

We note that, in general, $T[\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \hat{u}_{g_1}] = \hat{u}_{g_1}$ only up to a constant of integration on each characteristic. Here, as we are integrating against a mean-free quantity. This constant of integration does not contribute to the integral. Hence it can be omitted.

Altogether,

$$
L_1 = \int_{\mathcal{D}} \nu\, \hat{u}_1\, \hat{u}_{g_1} + \tfrac{1}{2}\, \bar{u}_1^2 - \lambda\, \hat{u}_{g_1}^2\, d\boldsymbol{x}\,, \tag{4.49}
$$

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

4.6 Variational principle

The variation of (4.49) is

$$
\delta L_1 = \int_{\mathcal{D}} \delta u_1\, (\bar{u}_1 + \nu\, \hat{u}_{g_1}) + \delta \hat{u}_{g_1}\, (\nu\, \hat{u}_1 - 2\lambda\, \hat{u}_{g_1})\, d\boldsymbol{x} \equiv \int_{\mathcal{D}} \delta u_1\, p_1 + \delta \hat{u}_{g_1}\, \hat{b}_1\, d\boldsymbol{x}\,, \tag{4.50}
$$

where

$$
p_1 = \bar{u}_1 + \nu\, \hat{u}_{g_1}\,, \quad \text{and} \quad b_1 = \nu\, \hat{u}_1 - 2\lambda\, \hat{u}_{g_1}\,. \tag{4.51}
$$

We rewrite the last term of (4.50) with the help of (4.33a) as

$$
\int_{\mathcal{D}} \delta \hat{u}_{g_1}\, \hat{b}_1\, d\boldsymbol{x} = -\int_{\mathcal{D}} \delta T[\partial_y \rho]\, \hat{b}_1\, d\boldsymbol{x} = \int_{\mathcal{D}} \partial_y \delta\rho\, T[\hat{b}_1]\, d\boldsymbol{x} = -\int_{\mathcal{D}} \delta\rho\, \partial_y B_1\, d\boldsymbol{x}\,, \tag{4.52}
$$

where we define $B_1 = T[\hat{b}_1]$. The general Euler–Poincaré equation reads

$$
\partial_t \boldsymbol{m} + (\boldsymbol{\nabla} \times \boldsymbol{m}) \times \boldsymbol{u} + \frac{\delta l}{\delta \rho}\, \boldsymbol{\nabla} \rho = \boldsymbol{\nabla} \tilde{\phi}\,. \tag{4.53}
$$

We have

$$\boldsymbol{m} = \frac{\delta l}{\delta \boldsymbol{u}} = \boldsymbol{R} + \varepsilon\, \boldsymbol{p}\,, \qquad \text{and} \qquad \frac{\delta l}{\delta \rho} = -z - \varepsilon\, \partial_y B_1 \tag{4.54}$$

with $\boldsymbol{p} = (p_1, 0, 0)$, so that, setting $\tilde{\phi} = -\phi - z\rho$, we obtain

$$\boldsymbol{\Omega} \times \boldsymbol{u} + \rho \boldsymbol{k} + \varepsilon \left(\partial_t \boldsymbol{p} + (\boldsymbol{\nabla} \times \boldsymbol{p}) \times \boldsymbol{u} - \boldsymbol{\nabla} \rho\, \partial_y B_1 \right) = -\boldsymbol{\nabla} \phi\,. \tag{4.55}$$

Setting

$$\boldsymbol{\xi} \equiv \boldsymbol{\nabla} \times \boldsymbol{p} = \begin{pmatrix} 0 \\ \partial_z \bar{u}_1 \\ -\partial_y \bar{u}_1 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ \partial_z \hat{u}_{g_1} \\ -\partial_y \hat{u}_{g_1} \end{pmatrix} \equiv \boldsymbol{\omega} + \nu\, \boldsymbol{\zeta}\,, \tag{4.56}$$

and taking the curl of (4.55) to remove the pressure term, we find

$$u_2\, \boldsymbol{k} - \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \times (\rho \boldsymbol{k}) + \varepsilon \left(\partial_t \boldsymbol{\xi} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \partial_y B_1 \right) = 0\,. \tag{4.57}$$

The vertical component of (4.57) is

$$0 = u_2 - \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} u_3 + \varepsilon \left(\partial_t \xi_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} \xi_3 - \boldsymbol{\xi} \cdot \boldsymbol{\nabla} u_3 - \nabla \rho \cdot \nabla^{\perp} \partial_y B_1 \right). \tag{4.58}$$

We rewrite it,

$$\partial_t \omega_3 + \boldsymbol{u} \cdot \boldsymbol{\nabla} \omega_3 = -\nu\, \partial_t \zeta_3 - \nu\, \boldsymbol{u} \cdot \boldsymbol{\nabla} \zeta_3 + \boldsymbol{\xi} \cdot \boldsymbol{\nabla} u_3 + \nabla \rho \cdot \nabla^{\perp} \partial_y B_1 - \varepsilon^{-1} \left(u_2 - \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} u_3 \right). \tag{4.59}$$

We note that the vertical component of the geostrophic velocity is zero and so $\mathcal{O}(\varepsilon^{-1})$-term is actually contributing at $\mathcal{O}(1)$. Then, we take the average of (4.59) and we obtain

$$\begin{aligned} \partial_t \omega_3 + \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega_3 = {} & \overline{\boldsymbol{\omega} \cdot \boldsymbol{\nabla} u_3} + \overline{\nu \left(-\boldsymbol{u} \cdot \boldsymbol{\nabla} \zeta_3 + \boldsymbol{\zeta} \cdot \boldsymbol{\nabla} u_3 \right)} + \overline{\nabla \rho \cdot \nabla^{\perp} \partial_y B_1} \\ & - \left(\overline{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega_3} - \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \omega_3 \right) - \varepsilon^{-1} \bar{u}_2\,, \end{aligned} \tag{4.60}$$

where $\bar{u}_2$ is small enough and so it is the (slow) prognostic equation. We note that the commutation idea as given in Lemma 3.11.1-3.11.2 do not work here. Coriolis vector depends on y-coordinate and so it does not allow to take some terms out of the integral that we define the average directly.

To obtain the balance relation, we proceed as follows. Firstly, we check the time derivative term on (4.57) which appears on the meridional-vertical plane given by

$$\begin{aligned}\partial_t \xi_{\mathrm{m}} &= -\partial_t \nabla^{\perp}_{\mathrm{m}} \bar{u}_1 - \nu \partial_t \nabla^{\perp}_{\mathrm{m}} \hat{u}_{g_1} \\ &= -\partial_t \nabla^{\perp}_{\mathrm{m}} \bar{u}_1 + \nu\, \nabla^{\perp}_{\mathrm{m}} T[\partial_y (\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho)] \\ &= -\partial_t \nabla^{\perp}_{\mathrm{m}} \bar{u}_1 + \tfrac{\nu}{\kappa}\, \nabla^{\perp}_{\mathrm{m}} (\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho) - \nu\, \nabla^{\perp}_{\mathrm{m}} \big(T[\tfrac{y}{\kappa} \partial_z (\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho)]\big)\,, \end{aligned} \tag{4.61}$$

since $\hat{u}_{g_1} = T[\partial_y \rho]$.

We assume the strong stratification with small $\tilde{\rho}$ via

$$\partial_z \rho = -\alpha + \partial_z \tilde{\rho}\,, \tag{4.62}$$

and so the last term on (4.57) is

$$\boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} \partial_y B_1 = \begin{pmatrix} -\nabla^{\perp} \tilde{\rho}\, \partial_{yz} B_1 - \alpha \nabla^{\perp} \partial_y B_1 \\ \nabla^{\perp} \tilde{\rho} \cdot \nabla \partial_y B_1 \end{pmatrix}. \tag{4.63}$$

We insert U into $\hat{u}$, i.e, $\hat{u} = \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} U$, and we obtain the second order differential equation for U on the horizontal component of (4.57)

$$\begin{aligned} LU = -\nabla^{\perp} \rho + \varepsilon \left(\partial_{tz} \bar{u}_1 - \tfrac{\nu}{\kappa}\, \partial_z (\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho) + \nu\, \partial_z \big(T[\tfrac{y}{\kappa} \partial_z (\boldsymbol{u} \cdot \boldsymbol{\nabla} \rho)]\big) \right) \boldsymbol{j} + \varepsilon\, \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \boldsymbol{U} \cdot \boldsymbol{\nabla} \xi \\ - \varepsilon\, \bar{\boldsymbol{u}} \cdot \boldsymbol{\nabla} (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \xi) - \varepsilon\, \nabla^{\perp} \tilde{\rho}\, \partial_{yz} B_1 - \varepsilon\, \alpha \nu\, \nabla^{\perp} \partial_y U_{g_1} + \kappa\, \partial_z U - \varepsilon\, \xi_2 \partial_z U\,, \end{aligned} \tag{4.64}$$

where the second order differential operator acting on U is

$$LU = (\kappa^2 + \varepsilon\, \xi_2)\, \partial_y^2 U + (2\kappa\, y + \varepsilon\, \boldsymbol{\Omega} \cdot \boldsymbol{\xi})\, \partial_{yz} U + (y^2 + \varepsilon\, \xi_3 y)\, \partial_z^2 U + \varepsilon\, \alpha \nabla^{\perp} \partial_y U_1 \tag{4.65}$$

which is not elliptic equation because there is a time-derivative term in the meridional component. Besides, there are some higher order terms on the right-hand side of the meridional equation and they are not easily split because of the operator T. Before making some more assumptions, we check the meridional-vertical plane component for another possible balance relation. We apply the curl on the meridional-vertical plane onto (4.61), we have

$$\begin{aligned}\partial_t \nabla_m^\perp \cdot \xi_{\mathrm{m}} &= -\partial_t \Delta_{\mathrm{m}} \bar{u}_1 + \nu \Delta_{\mathrm{m}} T[\partial_y(\boldsymbol{u}\cdot\boldsymbol{\nabla}\rho)] \\ &= -\partial_t \Delta_{\mathrm{m}} \bar{u}_1 - \alpha\nu \Delta_{\mathrm{m}} T[\partial_y u_3] + \nu \Delta_{\mathrm{m}} T[\partial_y(\boldsymbol{u}\cdot\boldsymbol{\nabla}\tilde{\rho})] \end{aligned} \tag{4.66}$$

which suggest a balance relation for u_3. However, it is not enough for the construction of the velocity field.

4.6.1 A special case on the construction of the balance model

We go on an extreme case on the meridional velocity which is taken small as before. We choose that $u_2 = \bar{u}_2 = 0$ and the density is free from the zonal dependence, i.e, $\rho = \rho(y,z,t)$ and so $\theta = \theta(y,z,t)$. Therefore, the velocity field is $\boldsymbol{u} = (u_1, 0, u_3)$. Then, we rewrite the elliptic equation

$$LU_1 = -\partial_y \rho - \varepsilon\, \partial_y \tilde{\rho}\, \partial_{yz} B_1 - \varepsilon\, \alpha\nu\, \partial_y^2 U_{g_1} - \kappa\, \partial_z U_1 + \varepsilon\, \xi_2 \partial_z U_1\,, \tag{4.67}$$

where the operator L acting on U_1 is

$$LU_1 = (\kappa^2 + \varepsilon\, \xi_2)\, \partial_y^2 U_1 + (2\kappa\, y + \varepsilon\, \boldsymbol{\Omega}\cdot\boldsymbol{\xi})\, \partial_{yz} U_1 + (y^2 + \varepsilon\, \xi_3 y)\, \partial_z^2 U_1\,. \tag{4.68}$$

The related prognostic equation becomes

$$\begin{aligned}\partial_t \omega_3 + \bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega_3 &= \overline{\boldsymbol{\omega}\cdot\boldsymbol{\nabla} u_3} + \overline{\nu\left(-\boldsymbol{u}\cdot\boldsymbol{\nabla}\zeta_3 + \boldsymbol{\zeta}\cdot\boldsymbol{\nabla} u_3\right)} + \overline{\nabla\rho\cdot\nabla^\perp \partial_y B_1} \\ &\quad - \left(\overline{\boldsymbol{u}\cdot\boldsymbol{\nabla}\omega_3} - \bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega_3\right). \end{aligned} \tag{4.69}$$

For the first glance, we have two evolution equations for the prognostic variables ω_3 and ρ, one kinematic relationship, incompressibility condition and $u_2 = 0$. It seems that they are enough for the construction of the velocity field. Detail analysis is necessary for this ongoing work.

4.7 Conservation laws

Hamiltonian of the reduced Lagrangian is

$$H = \int_{\mathcal{D}} \varepsilon \left(\frac{1}{2} \bar{u}_1^2 + \lambda \, \hat{u}_{g_1}^2 \right) + \rho z \, d\boldsymbol{x} \, . \tag{4.70}$$

Therefore, the energy is conserved in the combination of the potential energy and the kinetic energy from the contribution of the zonal component of the velocity field on the reduced model. Then, the extreme choice in Subsection 4.6.1 might be reasonable.

The other conserved quantity is potential vorticity given by

$$\begin{aligned} q &= (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla} \rho \\ &= (\boldsymbol{\nabla} \times \boldsymbol{R}) \cdot \boldsymbol{\nabla} \rho + \varepsilon (\boldsymbol{\nabla} \times \boldsymbol{p}) \cdot \boldsymbol{\nabla} \rho \\ &= \Omega_{\mathrm{m}} \cdot \nabla_{\mathrm{m}} \rho - \varepsilon \, \nabla_{\mathrm{m}}^{\perp} \bar{u}_1 \cdot \nabla_{\mathrm{m}} \rho - \varepsilon \nu \, \nabla_{\mathrm{m}}^{\perp} \hat{u}_{g_1} \cdot \nabla_{\mathrm{m}} \rho \\ &= (\kappa + \varepsilon \, \partial_z \bar{u}_1 + \varepsilon \nu \, \partial_z \hat{u}_{g_1}) \, \partial_y \rho + (y - \varepsilon \, \partial_y \bar{u}_1 - \varepsilon \nu \, \partial_y \hat{u}_{g_1}) \, \partial_z \rho \, . \end{aligned} \tag{4.71}$$

The potential vorticity appears on the meridional-vertical plane. We know that $\hat{u}_{g_1} = T[\partial_y \rho]$ and $\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta = -\rho$ so it can be written as

$$q = \begin{vmatrix} y - \varepsilon \, \partial_y \bar{u}_1 + \varepsilon \nu \, \partial_y T[\partial_y (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta)] & -\partial_y (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) \\ \kappa + \varepsilon \, \partial_z \bar{u}_1 - \varepsilon \nu \, \partial_z T[\partial_y (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta)] & -\partial_z (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta) \end{vmatrix} \, , \tag{4.72}$$

on which it is expected to have a non-linear second order operator for θ. However, terms with the operator T brings complexity and we cannot write directly $T[\partial_y (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \theta)] = \partial_y \theta$ because $\Omega_z = y$.

4.8 Conclusion and Outlook

In this draft, we have studied a balance model on the equatorial β-plane with the full Coriolis force. The tangent layer has periodic boundary conditions along x-axis and rigid-lid approximation is assumed vertically. For simplicity, we assume that it is infinite on yz-plane and so the meridional velocity vanishes towards the boundary. We have constructed derivations via imposing degeneracy to the truncated Lagrangian which is called variational model reduction.

The identical setting was used in primitive equations by Salmon (1996) and Oliver and Vasylkevych (2016). Then, it was extended to the three-dimensional Boussinesq equation by Özden and Oliver (2021). This work provides the derivations for the long-wave regime constructed on the same idea following these works.

The ellipticity of the balance relation for the general model fails in our derivations. To handle it, we assume that the strong stratification on which the horizontal variation might be ignored. We lost the control on the geometry because of the anisotropy and the nearly horizontal the axis of rotation in the vicinity of the equator. Then, we worked on more extreme case with assumptions on the meridional velocity and the density. With this special case, it makes the construction of the fields possible and the detail analysis must be done for this ongoing work. It is consistent with the energy conservation because only the zonal component of the velocity field contributes to the conservation. In addition, we cannot write directly the ellipticity on the potential vorticity equation because of the complexity on variables with the operator T.

The equator is the special part of the Earth. The axis of the rotation aligns parallel to the horizontal plane. Then, we have not expected to have a similar result as we have in higher latitudes. β-plane approximation seems very convenient for the equatorial region. However, it causes a complexity in our derivations because of the y dependence on the vertical component. We expect to obtain a valid equatorial balance model in the long term.

Chapter 5

Conclusion

In this dissertation, we considered three dimensional stratified Euler-Boussinesq equations with the full Coriolis force. We derived the balance model on the f-plane for the mid-latitude and worked on the β-plane for an equatorial balance model. We used the non-hydrostatic approximation on each model and so our models include the vertical velocity. We pursued our derivations using variational principle because conservation laws follow automatically.

First, we investigate the balance model on the f-plane for the mid-latitude. It is an extended version of the model obtained by Oliver and Vasylkevych (2016) using primitive equations. Then, it is the most general model in the semi-geostrophic scaling with the full Coriolis force. The axis of rotation is not parallel to the local vertical. This makes the derivations difficult because of the non-zero horizontal component of the Coriolis vector. As a result of working on tilted lines, we obtain extra terms on the resulting model which are not seen on the primitive equation model.

We further examine the balance model on the β-plane in the long-wave scaling. In this case, the horizontal aspect ratio between meridional and zonal length scales is used as a small parameter. This approximation provides the geostrophic balance in the meridional direction. Thus, it supports Kelvin wave which appears on dynamics with zero-meridional velocity (Majda, 2003). According to the theory, the mixed Rossby-gravity wave is not seen on this scale because this wave requires asymmetry in the meridional direction (Boyd, 2018). This wave is not meridionally geostrophic. Therefore, the long-wave approximation implies the dynamics which contains only Rossby waves and Kelvin waves because meridional geostrophy does not support mixed Rossby-gravity wave even it is in Rossby-like regime. Kelvin waves require zero meridional velocity, but mixed Rossby-gravity waves need asymmetry in the meridional direction. Thus, the equatorial balance model which is free from only gravity waves needs a different approach.

The scaling is the main challenge on the equator. We considered three different scaling assumptions. We followed one of them for which the rotation is dominant. We choose it because we are interested the large-scale motions with rotation. On this region, the Coriolis vector depends on y-coordinate because of the β-plane approximation and the characteristics are not straight lines. So far, we have not obtained any convincing result in the general model, but it is promising work in more extreme case on which the meridional velocity is zero.

The following items remain open for the equatorial balance model:

- We have briefly checked the f-plane model for the equator assuming that the axis of the rotation horizontally aligns. However, it seems that the ellipticity of the balance relation fails there, but a full proof is necessary. We believe that it is due to the geometry. The coordinate dependence of the Coriolis vector presents an extra difficulty and then we suggest to construct the model on f-plane with constant latitude ϕ.

- The scaling that is used by Majda (2003) does not provide a solvable model, a brief derivation of the kinematic relation is given in Appendix A. On the other hand, Chan and Shepherd (2013) have a model with Kelvin waves in addition to Rossby waves. However, the variable of the prognostic equation is the height field and so it is not slow. In conclusion, we did not find any well-posed balance model in the literature. We expect to have a valid model with our ongoing work.

When we look at the literature, we observe that one needs to sacrifice one of the wave type from the resulting model depending on the method. For instance, Theiss and Mohebalhojeh (2009) obtain an equatorial balance model, but without Kelvin waves. Similarly, Chan and Shepherd, 2013 do not have mixed Rossby-gravity waves in their model. Therefore, a balance model near the equator covering all slow equatorial waves is still missing. Mathematically, it is expected to have interesting features (Majda, 2003). We hope to combine the variational approach with long-wave scaling in the future work.

Appendix A

Towards the solution of the equatorial balance model for Case 2

In this part, non-linear equatorial long-wave equations on the β-plane is briefly discussed. It was derived by Majda (2003) for the shallow water equations. In the leading order, we have

$$\partial_t u_1 + u \cdot \nabla u_1 - y\, u_2 = -\partial_x h\,, \tag{A.1a}$$

$$y\, u_1 = -\partial_y h\,, \tag{A.1b}$$

$$\partial_t h + \nabla \cdot (hu) = 0\,. \tag{A.1c}$$

where h denotes the total height field. We apply the cross-differentiation to the momentum equations (A.1a) and (A.1b), and we obtain

$$\partial_{yt} u_1 + u \cdot \nabla \partial_y u_1 + \partial_y u \cdot \nabla u_1 = \nabla \cdot (yu)\,, \tag{A.2}$$

Taking two y-derivatives of (A.1b), we have

$$2\,\partial_y u_1 + y\,\partial_{yy} u_1 = -\partial_{yyy} h\,. \tag{A.3}$$

We differentiate (A.3) in time, and we get

$$(2 + y\,\partial_y)\partial_{yt} u_1 = -\partial_{yyy}\partial_t h\,. \tag{A.4}$$

Then, we insert (A.2) and (A.1c) to (A.4), and we find

$$(2 + y\,\partial_y)(\nabla \cdot (yu) - u \cdot \nabla \partial_y u_1 - \partial_y u \cdot \nabla u_1) = \partial_{yyy} \nabla \cdot (hu)\,. \tag{A.5}$$

We now sort all terms containing u_2 onto the left, everything else onto the right hand side:

$$
\begin{aligned}
&(2+y\,\partial_y)\big(\partial_y(yu_2) - u_2\,\partial_{yy}u_1 - \partial_y u_1\,\partial_y u_2\big) - \partial_{yyyy}(hv) \\
&\qquad = -(2+y\,\partial_y)\big(\partial_x(yu_1) - u_1\partial_{xy}u_1 - \partial_y u_1\,\partial_x u_1\big) + \partial_{xyyy}(hu)\,. \qquad \text{(A.6)}
\end{aligned}
$$

Separating out linear terms (assuming that h is a small perturbation about $h = 1$), we find

$$
2\,u_2 + 3\,y\,\partial_y u_2 + y^2\,\partial_{yy}u_2 - \partial_{yyyy}u_2 + \text{NL} = \text{RHS}\,, \qquad \text{(A.7)}
$$

where "NL" denotes the whole non-linear terms on the left-hand side of (A.6) and "RHS" represents terms on the right-hand side of the same equation. Unfortunately, the fourth-order term has the "wrong" sign, and so we are not going to get a coercive bilinear form from the linear operator. Otherwise the expression would not look too bad: On a sufficiently small strip, the left-hand linear operator would be invertible, e.g. with homogeneous Dirichlet *and* Neumann conditions. But as things stand, there will be resonances for particular meridional widths of the channel. We conclude that the model with the scaling given in Case 1 does not provide a solvable kinematic balance relation.

Appendix B

Horizontal mean field on the mid-latitude

The following lemma gives the derivation of the horizontal mean field. Derivations are done for the mid-latitude and the coefficients are defined in Chapter 3. Here, we draw attention to the main difference between the Boussinesq and the shallow water models. The Boussinesq equations describe incompressible flow while and we do not have such a property for the shallow water models.

Lemma B.0.1. *Let $\langle \phi \rangle$ denote the average over the entire three-dimensional domain, $\langle \phi \rangle = 0$ if and only if the horizontal mean of $\bar{\phi}$ is zero for any function ϕ.*

Proof. The average over the entire domain $\mathcal{D}$ of some function ϕ

$$\langle \phi \rangle = s \int_{\mathbb{T}^2} \int_{-s^{-1}}^{0} \phi \circ \chi \, d\boldsymbol{x} = \int_{\mathbb{T}^2} \bar{\phi} \circ \chi \, dx = 0 \,, \tag{B.1}$$

so $\langle \phi \rangle = 0$ if and only if the horizontal mean of $\bar{\phi}$ is zero. $\square$

We check this argument of the Euler–Boussinesq equations. Initially, we assume $\langle \boldsymbol{u} \rangle = 0$. Thus, we have to look at the Euler–Boussinesq momentum equations term by term. The advective term is

$$\begin{aligned} \langle \boldsymbol{u} \cdot \nabla \boldsymbol{u} \rangle &= \int_{\mathcal{D}} \boldsymbol{u} \cdot \nabla \boldsymbol{u} \, d\boldsymbol{x} \\ &= - \int_{\mathcal{D}} \boldsymbol{u} \, \nabla \cdot \boldsymbol{u} \, d\boldsymbol{x} + \int_{\partial \mathcal{D}} \boldsymbol{u} \, \boldsymbol{u} \cdot \boldsymbol{n} \, dS \\ &= 0 \,, \end{aligned} \tag{B.2}$$

where the divergence theorem is used on the second line. The first term on it is zero by the incompressibility, and the second term vanishes by the zero-flux and the periodic boundary conditions. We note that for the shallow water equations, this argument does not work as the divergence of $\boldsymbol{u}$ is not necessarily zero, so the advective term can change the mean.

We know that the horizontal mean cannot be zero. Then, we split it into a steady vertical profile and a perturbation density, $\rho = \bar{\rho}(z) + \rho'(\boldsymbol{x}, t)$. Setting

$$p_{\text{static}} = \int_z^0 \bar{\rho}(z')\, dz' \,, \tag{B.3}$$

and

$$\begin{aligned} \langle \rho \rangle &= \int_{\mathcal{D}} \rho \circ \boldsymbol{\chi}\, d\boldsymbol{x} \\ &= \int_{\mathbb{T}^2} \int_{-H}^0 \bar{\rho} + \rho'\, d\boldsymbol{x} = \int_{\mathbb{T}^2} \int_{-H}^0 \rho'\, d\boldsymbol{x} \,, \end{aligned} \tag{B.4}$$

which contributes to the hydrostatic part of the pressure term which is just a constant. Thus, the mean of ρ does not change in time. Altogether, we rewrite the Boussinesq momentum equations as

$$\varepsilon \left(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} \right) + \boldsymbol{\Omega} \times \boldsymbol{u} = -\boldsymbol{\nabla} (p - p_{\text{static}}) - \rho' \boldsymbol{k} \,. \tag{B.5}$$

The average of the horizontal momentum equations over the entire domain is zero, but that of the vertical momentum equation needs more detail analysis.

Bibliography

Akramov, Ibrokhimbek and Marcel Oliver (2020). "On the existence of solutions to a bi-planar Monge–Ampère equation". In: *Acta Math. Sci.* 40.2, pp. 379–388. ISSN: 0252-9602. DOI: 10.1007/s10473-020-0206-6. URL: https://doi.org/10.1007/s10473-020-0206-6.

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

Arnold, Vladimir I and Boris A Khesin (1999). *Topological Methods in Hydrodynamics*. Springer.

Babin, A, A Mahalov, and B Nicolaenko (2002). "Fast singular oscillating limits of stably-stratified 3D Euler and Navier–Stokes equations and ageostrophic wave fronts". In: *Large-scale atmosphere–ocean dynamics*. Ed. by John Norbury and Ian Roulstone. Vol. 1. Cambridge University Press, pp. 126–201.

Benamou, J. D. and Y. Brenier (1998). "Weak existence for the semigeostrophic equations formulated as a coupled Monge–Ampère/transport problem". In: *SIAM J. Appl. Math.* 58.5, pp. 1450–1461.

Boyd, John P. (2018). *Dynamics of the Equatorial Ocean*. 1st ed. Springer-Verlag Berlin Heidelberg. ISBN: 9783662554760; 3662554763; 9783662554746; 3662554747.

Bretherton, Francis P. (1970). "A note on Hamilton's principle for perfect fluids". In: *J. Fluid Mech.* 44.1, pp. 19–31. DOI: 10.1017/S0022112070001660.

Brummell, Nicholas H., Neal E. Hurlburt, and Juri Toomre (1996). "Turbulent compressible convection with rotation. I. Flow structure and evolution". In: *Astrophys. J.* 473, pp. 494–513. DOI: 10.1086/178161.

Busse, F. H. (June 1994). "Convection driven zonal flows and vortices in the major planets". In: *Chaos An Interdisciplinary Journal of Nonlinear Science vol. 4 iss. 2* 4 (2). DOI: 10.1063/1.165999.

Çalık, Mahmut, Marcel Oliver, and Sergiy Vasylkevych (2013). "Global well-posedness for the generalized large-scale semigeostrophic equations". In: *Arch. Ration. Mech. Anal.* 207.3, pp. 969–990. ISSN: 0003-9527. DOI: 10.1007/s00205-012-0587-3. URL: http://dx.doi.org/10.1007/s00205-012-0587-3.

Cane, Mark A. and Edward S. Sarachik (1977). "Forced baroclinic ocean motions. II - The linear equatorial bounded case". In.

Chan, Ian H and Theodore G Shepherd (2013). "Balance model for equatorial long waves". In: *J. Fluid Mech.* 725, pp. 55–90.

Charney, J. G. (1948). "On the scale of atmospheric motions". In: *Geofys. Publ.* 4.17, pp. 1–17.

Colombo, Maria (2017). *Flows of non-smooth vector fields and degenerate elliptic equations. With applications to the Vlasov-Poisson and semigeostrophic systems*. Vol. 22. Tesi. Scuola Normale Superiore di Pisa (Nuova Series) [Theses of Scuola Normale Superiore di Pisa (New Series)]. Edizioni della Normale, Pisa, pp. xxiii+269. ISBN: 978-88-7642-606-3; 978-88-7642-607-0. DOI: 10.1007/978-88-7642-607-0. URL: https://doi.org/10.1007/978-88-7642-607-0.

Cullen, M. J. P. (2006). *A Mathematical Theory of Large-scale Atmosphere/Ocean Flow*. Imperial College Press. ISBN: 9781860945182. URL: https://books.google.de/books?id=vtSAZLYg6iEC.

Cullen, Michael J. P. and R. James Purser (1984). "An extended Lagrangian theory of semi-geostrophic frontogenesis". In: *J. Atmos. Sci.* 41.9, pp. 1477–1497.

Dellar, Paul J. (2011). "Variations on a beta-plane: derivation of non-traditional beta-plane equations from Hamilton's principle on a sphere". In: *J. Fluid Mech.* 674, pp. 174–195. DOI: 10.1017/S0022112010006464.

Dritschel, David G., Georg A. Gottwald, and Marcel Oliver (2017). "Comparison of variational balance models for the rotating shallow water equations". In: *J. Fluid Mech.* 822, pp. 689–716. DOI: 10.1017/jfm.2017.292.

Dunkerton Timothy J.; Crum, Francis X. (1995). "Eastward propagating $\sim$ 2- to 15-day equatorial convection and its relation to the tropical intraseasonal oscillation". In: *Journal of Geophysical Research vol. 100 iss. D12* 100 (D12). DOI: 10.1029/95jd02678.

Dutrifoy, Alexandre and Andrew Majda (2006). "The dynamics of equatorial long waves: a singular limit with fast variable coefficients". In: *Communications in Mathematical Sciences* 4.2, pp. 375 –397. DOI: `cms/1154635529`.

Ebin, David G. and Jerrold Marsden (1970). "Groups of diffeomorphisms and the motion of an incompressible fluid". In: *Ann. of Math. (2)* 92, pp. 102–163. ISSN: 0003-486X. DOI: `10.2307/1970699`. URL: `https://doi.org/10.2307/1970699`.

Eckart, Carl (1960). *Hydrodynamics of Oceans and Atmospheres*. First Edition. Pergamon Press. ISBN: 9780080092485; 0080092489.

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

Embid, Pedro F. and Andrew J. Majda (1996). "Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity". In: *Comm. Partial Differential Equations* 21.3-4, pp. 619–658. ISSN: 0360-5302. DOI: `10.1080/03605309608821200`. URL: `https://doi.org/10.1080/03605309608821200`.

— (1998). "Low Froude number limiting dynamics for stably stratified flow with small or finite Rossby numbers". In: *Geophys. Astrophys. Fluid Dyn.* 87 (1-2), pp. 1–50. DOI: `10.1080/03091929808208993`.

Ertel, Hans (Sept. 1942). "Ein neuer hydrodynamischer Erhaltungssatz". In: *Naturwissenschaften* 30, pp. 543–544. DOI: `10.1007/BF01475602`.

Evans, Lawrence C. (1998). *Partial Differential Equations*. Graduate Studies in Mathematics 19. American Mathematical Society. ISBN: 0821807722; 9780821807729.

Franzke, Christian L. E. et al. (2019). "Multi-scale methods for geophysical flows". In: *Energy Transfers in Atmosphere and Ocean*. Ed. by Carsten Eden and Armin Iske. Cham: Springer, pp. 1–51.

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

Gerkema, T. et al. (2008a). "Geophysical and astrophysical fluid dynamics beyond the traditional approximaion". In: *Reviews of Geophysics* 46.

Gerkema, T. et al. (2008b). "Geophysical and astrophysical fluid dynamics beyond the traditional approximation". In: *Rev. Geophys.* 46.2, RG2004. DOI: `https://doi.org/10.1029/2006RG000220`.

Gerkema, Theo and Victor Shrira (Apr. 2005b). “Near-inertial waves in the ocean: Beyond the ’traditional approximation’”. In: *Journal of Fluid Mechanics* 529, pp. 195–219. DOI: `10.1017/S0022112005003411`.

Gerkema, Theo and Victor I Shrira (2005c). “Near-inertial waves in the ocean: beyond the ‘traditional approximation’”. In: *J. Fluid Mech.* 529, pp. 195–219.

Gilbarg, David and Neil S. Trudinger (2001). *Elliptic Partial Differential Equations of Second Order*. Springer.

Gilbert, Andrew D. and Jacques Vanneste (2018). “Geometric generalised Lagrangian-mean theories”. In: *J. Fluid Mech.* 839, pp. 95–134. DOI: `10.1017/jfm.2017.913`.

Gill, Adrian E (1982). *Atmosphere-Ocean Dynamics*. Academic Press.

Gottwald, Georg A. and Marcel Oliver (2014). “Slow dynamics via degenerate variational asymptotics”. In: *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 470.2170, p. 20140460. ISSN: 1364-5021. DOI: `10.1098/rspa.2014.0460`.

Hayashi, Michiya and Hisanori Itoh (2012). “The Importance of the Nontraditional Coriolis Terms in Large-Scale Motions in the Tropics Forced by Prescribed Cumulus Heating”. In: *J. Atmos. Sci.* 69 (9), pp. 2699–2716. DOI: `10.1175/jas-d-11-0334.1`.

Holm, Darryl D. (2011). *Geometric Mechanics - Part I: Dynamics And Symmetry (2Nd Edition)*. 2nd Revised ed. ICP. ISBN: 184816775X; 9781848167759.

Holm, Darryl D, Jerrold E Marsden, and Tudor S Ratiu (1998). “The Euler–Poincaré equations and semidirect products with applications to continuum theories”. In: *Adv. Math.* 137.1, pp. 1–81.

Hoskins, B. J. and F. P. Bretherton (Jan. 1972). “Atmospheric Frontogenesis Models: Mathematical Formulation and Solution”. In: *J. Atmos. Sci.* 29.1, pp. 11–37. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1972)029<0011:AFMMFA>2.0.CO;2. URL: http://dx.doi.org/10.1175/1520-0469(1972)029<0011:AFMMFA>2.0.CO;2.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson (1985). “On the use and significance of isentropic potential vorticity maps”. In: *Quarterly Journal of the Royal Meteorological Society* 111 (470). DOI: `10.1002/qj.49711147002`.

Hoskins, Brian J (1975). “The geostrophic momentum approximation and the semi-geostrophic equations”. In: *J. Atmos. Sci.* 32.2, pp. 233–242.

Ju, Qiangchang and Pengcheng Mu (2019). "Low Froude and Rossby number three-scale singular limits of the rotating stratified Boussinesq equations". In: *Z. Angew. Math. Phys.* 70.6, Paper No. 161, 20. ISSN: 0044-2275. DOI: 10.1007/s00033-019-1208-x. URL: https://doi.org/10.1007/s00033-019-1208-x.

Juárez, M. de laTorre, B.M. Fisher, and G.S. Orton (2002). "Large Scale Geostrophic Winds with a Full Representation of the Coriolis Force: Application to IR Observations of the Upper Jovian Troposphere". In: *Geophys. Astrophys. Fluid Dyn.* 96 (2), pp. 87–114. DOI: 10.1080/03091920290027943.

Julien, Keith et al. (2006). "Generalized quasi-geostrophy for spatially anisotropic rotationally constrained flows". In: *J. Fluid Mech.* 555, pp. 233–274.

Kafiabad, Hossein A. and Peter Bartello (2016). "Balance dynamics in rotating stratified turbulence". In: *J. Fluid Mech.* 795, pp. 914–949. ISSN: 0022-1120. DOI: 10.1017/jfm.2016.164. URL: https://doi.org/10.1017/jfm.2016.164.

— (2018). "Spontaneous imbalance in the non-hydrostatic Boussinesq equations". In: *J. Fluid Mech.* 847, pp. 614–643. ISSN: 0022-1120. DOI: 10.1017/jfm.2018.338. URL: https://doi.org/10.1017/jfm.2018.338.

Kafiabad, Hossein A., Jacques Vanneste, and William R. Young (2021). "Wave-averaged balance: a simple example". In: *J. Fluid Mech.* 911, Paper No. R1. ISSN: 0022-1120. DOI: 10.1017/jfm.2020.1032. URL: https://doi.org/10.1017/jfm.2020.1032.

Kafiabad, Hossein Amini and Peter Bartello (2017). "Rotating stratified turbulence and the slow manifold". In: *Comput. & Fluids* 151, pp. 23–34. ISSN: 0045-7930. DOI: 10.1016/j.compfluid.2016.10.020. URL: https://doi.org/10.1016/j.compfluid.2016.10.020.

Kiladis, George N et al. (2009). "Convectively coupled equatorial waves". In: *Reviews of Geophysics* 47.2.

Kohma, Masashi and Kaoru Sato (Jan. 2013a). "Kelvin and Rossby Waves Trapped at Boundaries under the Full Coriolis Force". In: *SOLA* 9, pp. 9–14. DOI: 10.2151/sola.2013-003.

— (2013b). "Kelvin and Rossby waves trapped at boundaries under the full Coriolis force". In: *SOLA* 9, pp. 9–14. DOI: 10.2151/sola.2013-003.

Kosmann-Schwarzbach, Yvette (2011). *The Noether Theorems: Invariance and Conservation Laws in the Twentieth Century*. 1st ed. Sources and Studies in the History of

Mathematics and Physical Sciences. Springer-Verlag New York. ISBN: 9780387878683; 0387878688; 9780387878676; 038787867X.

Kuo, H. L. (1977). "Characteristics of disturbances in the atmosphere and oceans". In: *Pure Appl. Geophys.* 115 (4). DOI: `10.1007/bf00881216`.

Landau, L. D. and E. M. Lifshitz (1976). *Mechanics*. 3rd ed. Butterworth-Heinemann. ISBN: 0750628960.

LeBlond, Paul H. and Lawrence A. Mysak (1978). *Waves in the Ocean*. Elsevier. Elsevier Oceanography Series 20. Elsevier Science. ISBN: 9780444416025; 0444416021; 9780444419262; 0444419268; 9780080879772; 0080879772.

Leith, C. E. (May 1980). "Nonlinear Normal Mode Initialization and Quasi-Geostrophic Theory". In: *Journal of the Atmospheric Sciences vol. 37 iss. 5* 37 (5). DOI: 10.1175/1520-0469(1980)037<0958:NNMIAQ>2.0.CO;2.

Lucas, Carine, James C. McWilliams, and Antoine Rousseau (2017). "On nontraditional quasi-geostrophic equations". In: *ESAIM Math. Model. Numer. Anal.* 51.2, pp. 427–442. ISSN: 0764-583X. DOI: `10.1051/m2an/2016041`. URL: `https://doi.org/10.1051/m2an/2016041`.

Majda, A. (2003). *Introduction to PDEs and Waves for the Atmosphere and Ocean*. American Mathematical Society. ISBN: 9780821883495. URL: `https://books.google.co.uk/books?id=GdGTHgaEC7UC`.

Matsuno, Taroh (1966). "Quasi-geostrophic motions in the equatorial area". In: *J. Meteor. Soc. Japan* 44.1, pp. 25–43.

McIntyre, M. E. and I. Roulstone (2002). "Are there higher-accuracy analogues of semi-geostrophic theory?" In: *Large-scale atmosphere–ocean dynamics*. Ed. by John Norbury and Ian Roulstone. Vol. 2. Cambridge University Press, pp. 301–364.

McIntyre, M.E. (2015). "Dynamical Meteorology – Balanced Flow". In: *Encyclopedia of Atmospheric Sciences*. Ed. by John Pyle and Fuqing Zhang. Second. Oxford: Academic Press, pp. 298–303. ISBN: 978-0-12-382225-3. DOI: `http://dx.doi.org/10.1016/B978-0-12-382225-3.00484-9`. URL: `http://www.sciencedirect.com/science/article/pii/B9780123822253004849`.

McIntyre, Michael E and Warwick A Norton (2000). "Potential vorticity inversion on a hemisphere". In: *J. Atmos. Sci.* 57.9, pp. 1214–1235.

Mohebalhojeh, Ali R. and Jürgen Theiss (2011). "The assessment of the equatorial counterpart of the quasi-geostrophic model". In: *Quart. J. R. Meteorol. Soc.* 137.658, pp. 1327–1339. ISSN: 1477-870X. DOI: `10.1002/qj.835`. URL: `http://dx.doi.org/10.1002/qj.835`.

Nieves, David et al. (2016). "Investigations of non-hydrostatic, stably stratified and rapidly rotating flows". In: *J. Fluid Mech.* 801, pp. 430–458. DOI: `10.1017/jfm.2016.443`.

Noether, E. (1918). "Invariante Variationsprobleme". In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918, pp. 235–257.

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

Oliver, Marcel (2006). "Variational asymptotics for rotating shallow water near geostrophy: a transformational approach". In: *J. Fluid Mech.* 551, pp. 197–234. ISSN: 0022-1120. DOI: `10.1017/S0022112005008256`. URL: `http://dx.doi.org/10.1017/S0022112005008256`.

— (2014). "A variational derivation of the geostrophic momentum approximation". In: *J. Fluid Mech.* 751, R2 (10 pages). ISSN: 0022-1120. DOI: `10.1017/jfm.2014.309`. URL: `http://dx.doi.org/10.1017/jfm.2014.309`.

Oliver, Marcel and Sergiy Vasylkevych (2016). "Generalized large-scale semigeostrophic approximations for the f-plane primitive equations". In: *J. Phys. A: Math. Theor.* 49, p. 184001. DOI: `10.1088/1751-8113/49/18/184001`.

Ong, Hing and Paul Roundy (May 2019). "Linear effects of nontraditional Coriolis terms on intertropical convergence zone forced large-scale flow". In: *Quarterly Journal of the Royal Meteorological Society* 145, pp. 2445–2453. DOI: `10.1002/qj.3572`.

Palais, Richard S. (1968). *Foundations of Global Non-Linear Analysis*. W. A. Benjamin, Inc., New York-Amsterdam, pp. vii+131.

Pedlosky, Joseph (May 1965). "A Study of the Time Dependent Ocean Circulation". In: *Journal of the Atmospheric Sciences vol. 22 iss. 3* 22 (3). DOI: `10.1175/1520-0469(1965)022<0267:asottd>2.0.co;2`.

— (1987). *Geophysical Fluid Dynamics*. Second. Springer.

Pozrikidis, C. (2011). *Introduction to Theoretical and Computational Fluid Dynamics.* 2nd ed. Oxford University Press. ISBN: 9780199752072; 0199752079.

Ripa, P. (1981). "Symmetries and conservation laws for internal gravity waves". In: *AIP Conference Proceedings* 76.1, pp. 281–306. DOI: `10.1063/1.33180`.

Roulstone, I. and M. J. Sewell (1997). "The mathematical structure of theories of semigeostrophic type". In: *Philos. Trans. Roy. Soc. London Ser. A* 355.1734, pp. 2489–2517. ISSN: 0962-8428. DOI: `10.1098/rsta.1997.0144`. URL: `https://doi.org/10.1098/rsta.1997.0144`.

Roulstone, Ian and Michael J. Sewell (Apr. 1996). "Potential vorticities in semi-geostrophic theory". In: *Quarterly Journal of the Royal Meteorological Society vol. 122 iss. 532* 122 (532). DOI: `10.1002/qj.49712253210`.

Salmon, Rick (1982). "Hamilton's principle and Ertel's theorem". In: *AIP Conference Proceedings* 88.1, pp. 127–135. DOI: `10.1063/1.33631`.

— (1983). "Practical use of Hamilton's principle". In: *J. Fluid Mech.* 132, pp. 431–444. DOI: `10.1017/S0022112083001706`.

— (1985). "New equations for nearly geostrophic flow". In: *J. Fluid Mech.* 153, pp. 461–477. ISSN: 1469-7645. DOI: `10.1017/S0022112085001343`. URL: `http://journals.cambridge.org/article_S0022112085001343`.

— (1988). "Semigeostrophic theory as a Dirac-bracket projection". In: *J. Fluid Mech.* 196, pp. 345–358. DOI: `10.1017/S0022112088002733`.

— (1996). "Large-scale semigeostrophic equations for use in ocean circulation models". In: *J. Fluid Mech.* 318, pp. 85–105. ISSN: 1469-7645. DOI: `10.1017/S0022112096007045`. URL: `http://journals.cambridge.org/article_S0022112096007045`.

Stewart, Andrew L and Paul J Dellar (2010). "Multilayer shallow water equations with complete Coriolis force. Part 1. Derivation on a non-traditional beta-plane". In: *J. Fluid Mech.* 651, p. 387.

— (2012). "Multilayer shallow water equations with complete Coriolis force. Part 2. Linear plane waves". In: *J. Fluid Mech.* 690, pp. 16–50.

Sverdrup, H. U. (Nov. 1947). "Wind-Driven Currents in a Baroclinic Ocean; with Application to the Equatorial Currents of the Eastern Pacific". In: *Proceedings of the National Academy of Sciences vol. 33 iss. 11* 33 (11). DOI: `10.2307/87657`.

Theiss, Jürgen and Ali R Mohebalhojeh (2009). "The equatorial counterpart of the quasi-geostrophic model". In: *J. Fluid Mech.* 637, pp. 327–356.

Vallis, Geoffrey K (2017). *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*. Second. Cambridge University Press.

Vasylkevych, S. and N. Žagar (2021). "High accuracy global prognostic model for the simulation of the Rossby and gravity wave dynamics". In: *Quart. J. Roy. Meteor. Soc.* DOI: 10.1002/qj.4006.

Verdière, A. Colin de and R. Schopp (Oct. 1994). "Flows in a rotating spherical shell: the equatorial case". In: *J. Fluid Mech.* 276. DOI: 10.1017/S0022112094002545.

Verkley, WTM and IR van der Velde (2010). "Balanced dynamics in the tropics". In: *Quart. J. R. Meteorol. Soc.* 136.646, pp. 41–49.

Wallace, John M and VE Kousky (1968). "Observational evidence of Kelvin waves in the tropical stratosphere". In: *Journal of Atmospheric Sciences* 25.5, pp. 900–907.

Wallace John M.; Kousky, V. E. (Sept. 1968). "Observational Evidence of Kelvin Waves in the Tropical Stratosphere". In: *Journal of the Atmospheric Sciences vol. 25 iss. 5* 25 (5). DOI: 10.1175/1520-0469(1968)025<0900:OEOKWI>2.0.CO;2.

Wedi, Nils (June 2014). "Increasing horizontal resolution in numerical weather prediction and climate simulations: Illusion or panacea?" In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 372. DOI: 10.1098/rsta.2013.0289.

Wetzel, Alfredo N. et al. (2019). "Balanced and unbalanced components of moist atmospheric flows with phase changes". In: *Chin. Ann. Math. Ser. B* 40.6, pp. 1005–1038. ISSN: 0252-9599. DOI: 10.1007/s11401-019-0170-4. URL: https://doi.org/10.1007/s11401-019-0170-4.

Wheeler, Matthew and George N Kiladis (1999). "Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain". In: *J. Atmos. Sci.* 56.3, pp. 374–399.

White, A A (2002). "A view of the equations of meteorological dynamics and various approximations". In: *Large-scale atmosphere–ocean dynamics*. Ed. by John Norbury and Ian Roulstone. Vol. 1. Cambridge University Press, pp. 1–100.

Whitehead, Jared P., Terry Haut, and Beth A. Wingate (2018). "The effect of two distinct fast time scales in the rotating, stratified Boussinesq equations: variations from quasi-geostrophy". In: *Theor. Comput. Fluid Dyn.* 32.6, pp. 713–732. ISSN: 0935-4964. DOI: 10.1007/s00162-018-0472-2. URL: https://doi.org/10.1007/s00162-018-0472-2.

Whitehead, Jared P. and Beth A. Wingate (Oct. 2014). "The influence of fast waves and fluctuations on the evolution of the dynamics on the slow manifold". In: *J. Fluid Mech.* 757, pp. 155–178. ISSN: 1469-7645. DOI: 10.1017/jfm.2014.467. URL: http://journals.cambridge.org/article_S0022112014004674.

Yanai, Michio and Takio Maruyama (1966). "Stratospheric wave disturbances propagating over the equatorial Pacific". In: *Journal of the Meteorological Society of Japan. Ser. II* 44.5, pp. 291–294.

Yano, Jun-Ichi (Jan. 1998). "Deep convection in the interior of major planets: a review." In: *Australian Journal of Physics* 51, p. 875. DOI: 10.1071/P97079.

Yano, Jun-Ichi and Marine Bonazzola (Jan. 2009). "Scale Analysis for Large-Scale Tropical Atmospheric Dynamics". In: *American Meteorological Society* 66 (1). DOI: 10.1175/2008JAS2687.1.

Yoshida, Kozo (1959). "A Theory of the Cromwell Current (the Equatorial Undercurrent) and of the Equatorial Upwelling". In: *Journal of the Oceanographical Society of Japan* 15.4, pp. 159–170. DOI: 10.5928/kaiyou1942.15.159.

Zeitlin, Vladimir (Nov. 2008). "Decoupling of Balanced and Unbalanced Motions and Inertia–Gravity Wave Emission: Small versus Large Rossby Numbers". In: *Journal of the Atmospheric Sciences* 65.11, pp. 3528–3542. ISSN: 0022-4928. DOI: 10.1175/2008JAS2481.1.

Özden, Gözde and Marcel Oliver (June 2021). "Variational balance models for the three-dimensional Euler–Boussinesq equations with full Coriolis force". In: *Physics of Fluids* 33.7, p. 076606. ISSN: 1089-7666. DOI: 10.1063/5.0053092.

Žagar N.and Kasahara, A. et al. (Apr. 2015). "Normal-mode function representation of global 3-D data sets: open-access software for the atmospheric research community". In: *Geoscientific Model Development* 8 (4). DOI: 10.5194/gmd-8-1169-2015.