Abstract

The ocean surface boundary layer (OSBL) tends to be vertically well-mixed, but it can be horizontally inhomogeneous. For example, buoyancy fronts can develop as steep horizontal gradients of temperature, salinity and density in the OSBL. Misalignment of these steep horizontal buoyancy gradients with either the horizontal gradients of surface elevation or bathymetry is known to generate circulation which can deflect ocean currents which transport fluid properties, as well as pollution and debris. In turn, the generation of circulation itself can break up these fronts, thereby cascading horizontally circulating structures to smaller scales.

Taking advantage of the vertically well-mixed property of the OSBL and working in the stochastic Euler–Poincaré variational framework introduced in [HL19], we derive the thermal rotating shallow water (TRSW) equations with stochastic advection by Lie transport (SALT), as a theoretical foundation for uncertainty quantification and data assimilation in computational models of the effects of submesoscale oceanic circulation using data-driven stochastic parametrisations, as in [CCH+18, CCH+19]. The key feature of SALT for geophysical fluid dynamics (GFD) is that SALT respects the Kelvin circulation theorem, which is the essence of cyclogenesis. Asymptotic expansion in the three small parameters present in the TRSW model in the neighbourhood of thermal quasi-geostrophic balance among the flow velocity and the gradients of free surface elevation and buoyancy leads first to the deterministic thermal versions of the Lagrangian 1 (TL1) equations and then to the thermal quasi-geostrophic (TQG) theory. We illustrate the instabilities of TQG which cascade circulation to smaller, typically unresolvable scales. We derive the stochastic version of this hierarchy of models TRSW/TL1/TQG in the framework of the stochastic Euler–Poincaré variational principle. Finally, we indicate the next steps in applying these results for uncertainty quantification and data assimilation of the cascading cyclogenetic effects of buoyancy fronts using the data-driven stochastic parametrisation algorithm based on SALT at these three levels of description.

1 Introduction

In this paper we are dealing with the thermal rotating shallow water (TRSW) equations, which can be regarded as the vertically averaged version of the primitive equations with a buoyancy variable [Zei18]. In the balanced 2D model hierarchy of TRSW, TL1 and TQG, we are investigating a certain stochastic model of cyclogenesis as a potential basis for stochastic parametrisation of the dynamical creation of unresolved degrees of freedom in computational simulations of upper ocean dynamics. Specifically, we have chosen the SALT (Stochastic Advection by Lie Transport) algorithm introduced in [HL19] and applied in [CCH+18, CCH+19] as our modelling approach. The SALT approach preserves the Kelvin circulation theorem and an infinite family of integral conservation laws. The goal of the SALT algorithm is to quantify the uncertainty in the process of upscaling, or coarse graining of either observed or synthetic data at fine scales, for use in computational simulations at coarser scales. The present work prepares us to take the next step from (ii) to (iii) in the well-known path of discovery in oceanography, weather prediction and climate science, which is

(i) driven by large datasets and new methods for its analysis;
(ii) informed by rigorous mathematical derivations and analyses of stochastic geophysical fluid equations;
(iii) quantified using computer simulations, evaluated for uncertainty, variability and model error;
(iv) optimized by cutting edge data assimilation techniques, then
(v) compared with new observation datasets to determine what further analysis and improvements will be needed.

The objective in applying the SALT algorithm to coarse grained simulations is to answer the following question, enunciated in [CCH+18, CCH+19]: “How can one use computationally simulated surrogate data at highly resolved scales, in combination with the mathematics of stochastic processes in nonlinear dynamical systems, to estimate and model the effects on the simulated variability at much coarser scales of the computationally unresolvable, small, rapid, scales of motion at the finer scales?” The present paper will lay the theoretical foundations for addressing this question in the 2D context of the thermal rotating shallow water (TRSW) model and its balanced thermal quasi-geostrophic (TQG) model. Our eventual goal is to apply the SALT algorithm to calibrate our stochastic models for assimilating data, e.g., from satellite observations of the cascade in the OSBL of horizontally circulating structures to smaller scales, as shown in Figure 1 below.

Figure 1: This image of the Lofoten Vortex, courtesy of https://ovl.oceandatalab.com/ illustrates the configurations of submesoscale currents obtained from ESA Sentinel-3 OLCI instrument observation of chlorophyll on the surface of the Norwegian Sea in the Lofoten Basin, near the Faroe Islands. Fueled by warm saline Atlantic waters crossing the Norwegian Sea, the Lofoten Basin is a major reservoir of heat whose buoyancy gradient interacts with the bathymetry gradient to sustain a large anticyclonic vortex exhibiting intense mesoscale and submesoscale activity. In particular, the figure shows many of the features of submesoscale currents surveyed in [McW19]. High resolution (4 km) computational simulations of the Lofoten Vortex have recently discovered that its time-mean circulation is primarily barotropic, [VKL15], thereby making the Lofoten Vortex a reasonable candidate for investigation using vertically averaged dynamics such as the TQG approach. For more information about the Lofoten Vortex, see, e.g., [FSL18, BSB+17, BKP+20, BBK+18, FB20] and references therein.

The TRSW equations. The TRSW equations arise in a series of nested approximations of the rotating thermal Green-Naghdi equations, which were derived and investigated in [HL19]. Upon neglecting the nonhydrostatic pressure effects which are present in the Green-Naghdi model, the TRSW model is obtained. The TRSW equations and their quasi-geostrophic counterpart, the TQG equations, comprise standard models of thermal effects in GFD, as reviewed, e.g., in [BCC+18, Zei18]. The TQG equations are discussed in the GFD literature by Warneford and Dellar [WD13], for example, following earlier work by Ripa [Rip95]. In fact, the TRSW and TQG equations have been rederived several times, as recounted in [Zei18]. In this paper, we will derive the stochastic versions of TRSW and TQG, as well as TL1, which is a thermal version of an intermediate theory known in the GFD literature as Salmon’s L1 model [Sal83]. For this purpose, we follow the geometric approach of [Hol15] which is based on Hamilton’s variational principle for Eulerian fluid flows [HMR98].

The TRSW, TL1 and TQG models separate the wave and current aspects of their flows into gravity waves
and Rossby waves on the free surface, and fluid circulation in the region between the free surface and the bottom topography. The Kelvin circulation theorems for TRSW, TL1 and TQG show that horizontal gradients of the buoyancy in the fluid region (e.g., at thermal fronts) can couple to gradients on the upper and lower interfaces to produce horizontal circulation of the fluid whenever either of those two surface gradients are misaligned with the horizontal gradient of the buoyancy. Thus, both waves on the surface and variations of the bottom topography can create fluid circulation when the buoyancy is spatially inhomogeneous.

The generation of submesoscale circulations involves a wide range of time scales, as well as many couplings among the various degrees of freedom and the boundaries. The ‘irreducible imprecision’ of numerical simulations [McW07] and the sparsity of observed data in both space and time produce uncertainty in forecasts of submesoscale cyclogenesis and thereby present a ‘grand challenge’ for data assimilation.

In preparation for meeting this challenge, the present paper develops the stochastic variational principles for the TRSW, TL1 and TQG models of submesoscale cyclogenesis. This mathematical framework is needed in applying the SALT (Stochastic Advection by Lie Transport) approach to the derivation of stochastic fluid equations which preserve the geometric structure of fluid dynamics [Hol15]. The physical effects of stochasticity in the SALT approach for deriving the stochastic TRSW, TL1 and TQG models are revealed in their Kelvin circulation theorems. Namely, the corresponding material loops defining the circulation integrals are shown to move along stochastic Lagrangian paths. The motivations and recent applications of the SALT approach for uncertainty quantification and for data assimilation are laid out in [CCH+18, CCH+19].

Content of this paper

1. In section 2 we review the deterministic TRSW model by re-deriving its equations in the Euler–Poincaré variational framework of [HMR98]. In the Euler–Poincaré framework, we prove the Kelvin–Noether circulation theorem and discuss steady solution properties of the deterministic TRSW equations. This derivation of the TRSW equations with stochastic advection by Lie transport (SALT) is intended to be used as a systematic means of introducing data-driven stochastic parametrisations for uncertainty quantification and data assimilation. This uncertainty quantification and data assimilation has already been accomplished using this approach for the 2D Euler equations in a square domain with fixed boundaries and the 2-layer QG equations in a periodic channel, in [CCH+18] and [CCH+19], respectively.

2. In section 3 we discuss the deterministic thermal Eliassen approximation, or TL1 model, of TRSW, as derived from a combination of the Euler–Poincaré variational approach and asymptotic expansions in the vorticity–divergence representation of the fluid velocity. In the derivation of TL1, we use a modified version
of the Euler–Poincaré framework introduced in [AH96] which expresses the approximate momentum in terms of gradients of advected quantities.

3. In section 4 we derive the TL1 equations with stochastic advection by Lie transport (SALT) in the modified Euler–Poincaré variational framework. These stochastic TL1 equations would be useful for uncertainty quantification and data assimilation at this intermediate level of approximation.

4. In section 5 we take the next step in the asymptotic expansion to derive and discuss the deterministic and stochastic versions of the thermal quasi-geostrophic (TQG) equations respectively in sections 5.1 and 5.2. Section 5.3 specifies the details of the numerical implementation of the TQG solution shown in Figure 2.

5. In section 6 we conclude by outlining a few next steps and open problems to which the present work has led, but which are beyond the scope of the present paper.

2 The thermal rotating shallow water (TRSW) model

The thermal rotating shallow water (TRSW) model describes the motion of a single two dimensional layer of fluid with horizontally varying buoyancy and bottom topography (or, bathymetry). The TRSW model is an extension of the rotating shallow water model and a simplification of the various three dimensional models such as the Primitive Equations and the Euler-Boussinesq model commonly used for computationally simulating large-scale ocean and atmosphere circulation dynamics. The thermal rotating shallow water equations may also be interpreted as a model for an upper active layer of fluid on top of a lower inert layer. For that reason the TRSW model is sometimes called a 1.5 layer model [WD13]. A stochastic version of this model has already been derived from a variational point of view in Appendix B of [HL19]. For a related discussion of a fully multilayer variational model with nonhydrostatic pressure, see [CHP10].

2.1 Deterministic TRSW equations

The deterministic TRSW equations in a rotating planar domain $\mathcal{D} \subset \mathbb{R}^2$ with boundary $\partial \mathcal{D}$ are expressed using the following definitions. The depth is denoted $\eta = \eta(x, t)$, where $x = (x, y)$ is the horizontal vector position, and $t$ is time. The (nonnegative) horizontal buoyancy is written as $b(x, t) = (\bar{\rho} - \rho(x, t))/\bar{\rho}$, where $\rho(x, t)$ is the mass density, $\bar{\rho}$ is the uniform reference mass density. The nondimensional deterministic TRSW equations for the Eulerian horizontal vector velocity $u(x, t)$, thickness $\eta(x, t)$, and buoyancy $b(x, t)$ of the active fluid layer are given by

$$
\frac{D}{Dt} u + \frac{1}{Ro} f \mathbf{\hat{z}} \times u = -\frac{1}{Fr^2} \nabla (b \zeta) + \frac{1}{2Fr^4} (\zeta - h) \nabla b , \quad \frac{\partial \eta}{\partial t} + \nabla \cdot (\eta \mathbf{u}) = 0 , \quad \frac{D}{Dt} b = 0 .
$$

(2.1)

The other notation is $f$ for the Coriolis parameter, $\zeta = \eta - h$ for the free surface elevation, where $h(x)$ is the time-independent mean depth, $Ro = U/(f_0 L)$ for the Rossby number and $Fr^2 = U^2/(gH)$ for the Froude number. The material time derivative for scalar advected quantities is given by $\frac{D}{Dt} = \partial_t + \mathbf{u} \cdot \nabla$. In defining the dimensionless numbers, $U$ denotes the horizontal velocity scale, $L$ the horizontal length scale, $f_0$ the typical rotation frequency, $g$ is the gravitational acceleration and $H$ is the typical depth. The notation $\mathbf{\hat{z}}$ is used to denote the unit vector perpendicular to $\mathcal{D}$. The boundary conditions are

$$
\mathbf{\hat{n}} \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{\hat{n}} \times \nabla b = 0 \quad \text{on the boundary } \partial \mathcal{D},
$$

(2.2)

meaning that fluid velocity $\mathbf{u}$ is tangential and buoyancy $b$ is constant on the boundary $\partial \mathcal{D}$, fixed in the frame rotating at time-independent angular frequency $\mathbf{\hat{z}} f(x)/2$. In the boundary conditions $\mathbf{\hat{n}}$ denotes the outward unit normal. Periodic boundary conditions may also be considered.

Variational formulation. The TRSW equations (2.1) can be derived by means of the Euler–Poincaré variational principle, as is shown in [BCC18]. When the equations of motion are derived in this framework, there is a natural way to express a number of fundamental relations. The first fundamental relation is the Kelvin circulation theorem, the second one is the advection equation for potential vorticity and the third one is an infinity of conserved integral quantities arising from Noether’s theorem for the symmetry of Eulerian fluid quantities under Lagrangian particle relabelling. For example, in rotating shallow water, without a buoyant scalar, the enstrophy is among these integral quantities. This framework turns out to be ideal for introducing stochasticity, as shown
2.2 The Euler–Poincaré theorem

Variational derivatives of functionals. The Euler–Poincaré theorem relies on variational derivatives of functionals. This type of derivative is given by the following definition.

**Definition 2.1.** A functional \( F[\rho] \) is defined as a map \( F : \rho \in C^\infty(D) \rightarrow \mathbb{R} \). The variational derivative of \( F(\rho) \), denoted \( \delta F/\delta \rho \), is defined by

\[
\delta F[\rho] := \lim_{\varepsilon \to 0} \frac{F[\rho + \varepsilon \phi] - F[\rho]}{\varepsilon} = \left. \frac{d}{d\varepsilon} F[\rho + \varepsilon \phi] \right|_{\varepsilon=0} = \int_D \frac{\delta F}{\delta \rho}(x) \phi(x) \, dx =: \left\langle \frac{\delta F}{\delta \rho}, \phi \right\rangle .
\]

In this definition, \( \varepsilon \in \mathbb{R} \) is a real parameter, \( \phi \) is an arbitrary smooth function, and the angle brackets \( \langle \cdot, \cdot \rangle \) indicate \( L^2 \) real symmetric pairing of integrable smooth functions on the flow domain \( D \). The function \( \phi(x) \) above is called the ‘variation of \( \rho \)’ and will be denoted as \( \delta \rho := \phi(x) \). Since the variation is a linear operator on functionals, we can denote the functional derivative \( \delta \) operationally as

\[
\delta F[\rho] = \left\langle \frac{\delta F}{\delta \rho}, \delta \rho \right\rangle .
\]

**Euler–Poincaré theorem.** Given the boundary conditions and definitions above, the following form of the Euler–Poincaré theorem with stochastic variations provides the corresponding stochastic equations of motion derived from Hamilton’s principle with a deterministic Lagrangian functional \( \ell : X \times V^* \rightarrow \mathbb{R} \) defined on the domain of flow, \( D \). Here \( X \) denotes the space of smooth vector fields on \( D \) and \( V^* \) is the vector space of advected quantities. Adveected quantities are tensor fields of different degrees. The space of smooth vector fields is a Lie algebra when endowed with the Jacobi–Lie bracket, which is denoted as \( [\cdot, \cdot] : X \times X \rightarrow X \), and is defined for \( u, v \in X \) by the commutator relation

\[
[u, v] := (u \cdot \nabla)v - (v \cdot \nabla)u \cdot \nabla.
\]

**Theorem 2.1** (Euler–Poincaré equations [HMR98]).
The following two statements are equivalent:

_i) Hamilton’s variational principle in Eulerian coordinates, with \( u \in X(D) \) and \( b, \eta \in V^*(D) \),

\[
\delta S := \delta \int_{t_1}^{t_2} \ell(u, b, \eta) \, dt = 0,
\]

holds on \( X(D) \times V^* \), using variations of the form

\[
\delta u = \frac{\partial \ell}{\partial t} v - [u, v], \quad \delta b = -(v \cdot \nabla)b, \quad \delta \eta = -\nabla \cdot (\eta v),
\]

where the vector field \( v \in X(D) \) is arbitrary and vanishes on the endpoints \( t_1 \) and \( t_2 \).

ii) The Euler–Poincaré equations hold. These equations are

\[
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta u} + (u \cdot \nabla) \frac{\delta \ell}{\delta u} + (\nabla u) \cdot \frac{\delta \ell}{\delta u} + \frac{\delta \ell}{\delta \eta} (\nabla \cdot u) = -\frac{\delta \ell}{\delta b} \nabla b + \eta \nabla \frac{\delta \ell}{\delta \eta},
\]

or, equivalently, in two-dimensional vector calculus notation,

\[
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta u} - u^\perp \left( \nabla^\perp \frac{\delta \ell}{\delta u} \right) + \nabla \left( u \cdot \frac{\delta \ell}{\delta u} \right) + \frac{\delta \ell}{\delta \eta} (\nabla \cdot u) = -\frac{\delta \ell}{\delta b} \nabla b + \eta \nabla \frac{\delta \ell}{\delta \eta},
\]

or, finally, as an embedding in three dimensional space,

\[
\frac{\partial}{\partial t} \frac{\delta \ell}{\delta u} - u \times \left( \nabla \times \frac{\delta \ell}{\delta u} \right) + \nabla \left( u \cdot \frac{\delta \ell}{\delta u} \right) + \frac{\delta \ell}{\delta \eta} (\nabla \cdot u) = -\frac{\delta \ell}{\delta b} \nabla b + \eta \nabla \frac{\delta \ell}{\delta \eta},
\]

with advection equations

\[
\frac{\partial}{\partial t} b = -u \cdot \nabla b \quad \text{and} \quad \frac{\partial}{\partial t} \eta = -\nabla \cdot (\eta u).
\]
Remark 2.2. The abstract statement of the Euler–Poincaré Theorem 2.1, formulated on general semidirect product Lie groups, is presented in [HMR98] deterministically, in [Hol15, dLHLT20] stochastically and finally in [CHLN20] on rough paths.

Remark 2.3. In Theorem 2.1, the operator $\delta$ in (2.6) is the functional derivative defined in (2.3), the brackets $[\cdot, \cdot]$ denote the commutator of vector fields defined in (2.5), and $v \in \mathfrak{X}(\mathcal{D})$ is an arbitrary vector field in two dimensions which vanishes at the endpoints in time, $t_1$ and $t_2$. Equations (2.9) and (2.10) are exactly the same and can be transformed into each other by the conventions $u^\perp = \hat{z} \times u$ and $\nabla^\perp = \hat{z} \cdot \nabla$, where $\hat{z}$ is the outward unit vector perpendicular to the planar domain $\mathcal{D}$.

Remark 2.4. One may interpret the Euler–Poincaré equation (2.8) in terms of Newton’s law for fluid motion. Namely, the rate of change of the covector momentum $P := \delta \ell / \delta u$ equals the sum of forces on the right hand side of equation (2.8).

Proof. Hamilton’s variational principle implies

$$0 = \int_{t_1}^{t_2} \left[ \left\langle \frac{\delta \ell}{\delta u}, \frac{\partial}{\partial t} \delta u \right\rangle_x + \left\langle \frac{\delta \ell}{\delta b}, \delta v \right\rangle_{V^*} + \left\langle \frac{\delta \ell}{\delta \eta}, \delta \eta \right\rangle_{V^*} \right] \, dt$$

$$= \int_{t_1}^{t_2} \left[ \left\langle \frac{\delta \ell}{\delta u}, \frac{\partial}{\partial t} \delta u \right\rangle_x + \left\langle \frac{\delta \ell}{\delta b}, -(v \cdot \nabla) b \right\rangle_{V^*} + \left\langle \frac{\delta \ell}{\delta \eta}, -\nabla \cdot (\eta v) \right\rangle_{V^*} \right] \, dt$$

$$= \int_{t_1}^{t_2} \left[ -\frac{\partial}{\partial t} \frac{\delta \ell}{\delta u} - (u \cdot \nabla) \frac{\delta \ell}{\delta u} - (\nabla u) \cdot \frac{\delta \ell}{\delta u} + \frac{\delta \ell}{\delta u} (\nabla \cdot u), v \right]_x + \left\langle -\frac{\delta \ell}{\delta b}, \nabla b \right\rangle_x$$

$$+ \left\langle \eta \nabla \frac{\delta \ell}{\delta \eta}, v \right\rangle_x \right] \, dt.$$

The subscripts $\mathfrak{X}$ and $V^*$ on the $L^2$ pairings indicate over which space that the pairing is defined. Since $v$ is arbitrary and vanishes at the endpoints $t_1$ and $t_2$ in time, the following equation holds,

$$\frac{\partial}{\partial t} \frac{\delta \ell}{\delta u} + (u \cdot \nabla) \frac{\delta \ell}{\delta u} + (\nabla u) \cdot \frac{\delta \ell}{\delta u} + \frac{\delta \ell}{\delta u} (\nabla \cdot u) = -\frac{\delta \ell}{\delta b} \nabla b + \eta \nabla \frac{\delta \ell}{\delta \eta}.$$ 

This finishes the proof of the stochastic Euler–Poincaré equation in (2.8). The equivalent forms in equations (2.9) and (2.10) follow by means of a standard vector identity. □

2.3 Kelvin–Noether circulation theorem

A straightforward calculation using the second advection equation in (2.11) shows that (2.8) may be written equivalently as follows.

Lemma 2.5. The Euler–Poincaré equation in (2.8) is equivalent to the following,

$$\frac{\partial}{\partial t} \left( \frac{1}{\eta} \frac{\delta \ell}{\delta u} \right) + (u \cdot \nabla) \left( \frac{1}{\eta} \frac{\delta \ell}{\delta u} \right) + (\nabla u) \cdot \left( \frac{1}{\eta} \frac{\delta \ell}{\delta u} \right) = -\frac{1}{\eta} \frac{\delta \ell}{\delta b} \nabla b + \nabla \frac{\delta \ell}{\delta \eta}. \quad (2.12)$$

One of the main features of Theorem 2.1 for fluid dynamics is that its Euler–Poincaré equations satisfy the following Kelvin circulation theorem.

Theorem 2.6 (Kelvin-Noether circulation). For an arbitrary loop $c(t)$ which is advected by the velocity field $u$, the following dynamics holds for the circulation integral $I$, given by

$$I := \oint_{c(t)} \frac{1}{\eta} \frac{\delta \ell}{\delta u} \cdot dx, \quad \frac{\partial}{\partial t} I = -\oint_{c(t)} \left( \frac{1}{\eta} \frac{\delta \ell}{\delta b} \right) \nabla b \cdot dx. \quad (2.13)$$

Remark 2.7. The notation $c(u)$ indicates that the material loop $c$ is transported by the flow $\phi_t$ which is generated by the vector field $u$. To be precise, $c(u) = \phi_t c(0)$, where $\phi_{t*}$ is the pull-back by the inverse of the flow $\phi_t$, also known as the push-forward by $\phi_t$.

Proof. The Kelvin circulation law (2.13) follows from Newton’s law of motion obtained from the stochastic Euler–Poincaré equation (2.12) for the evolution of momentum per unit mass $\eta^{-1} \delta \ell / \delta u$ concentrated on an advecting material loop, $c(t) = \phi_t c(0)$, where $\phi_t$ is the flow map which is generated by the vector field $u$. Upon
changing variables by pulling back the integrand to its initial position, the time derivative can be moved inside the integral and the product rule may be applied. Then, by inverting the pull-back we have the following
\[
\frac{\partial}{\partial t} \oint_{c(u)} \frac{1}{\eta} \left( \frac{\partial \ell}{\partial \eta} + \nabla \cdot (\nabla u) \right) \cdot \frac{\delta \ell}{\delta \eta} \cdot dx = -\oint_{c(u)} \frac{1}{\eta} \nabla b \cdot dx + \oint_{c(u)} \nabla \cdot \frac{\delta \ell}{\delta \eta} \cdot dx
\]
\[
= -\oint_{c(u)} \frac{1}{\eta} \nabla b \cdot dx .
\]
In the second line, we have used the Euler–Poincaré equation (2.8) and the advection equation for the density. The last step applies the fundamental theorem of calculus to prove vanishing of the last loop integral in the second line. For the corresponding proof in the general case, see [HMR98].

\[\square\]

Corollary 2.7.1. By Stokes Law, equation (2.15) in the Kelvin–Noether circulation theorem 2.6 implies
\[
\frac{\partial}{\partial t} I = -\int_{\partial S=c(u)} \hat{z} \cdot \nabla \left( \frac{1}{\eta} \frac{\delta \ell}{\delta b} \right) \times \nabla b \, dx \, dy .
\] (2.15)
Therefore, circulation is created by misalignment of the gradients of buoyancy \(b\) and its dual quantity \(\eta^{-1} \delta \ell / \delta b\).

The thermal rotating shallow water equations (2.1) in a planar domain \(D\) are obtained by applying the Euler–Poincaré theorem 2.1 to the Lagrangian
\[
\ell_{\text{TRSW}}(u, \eta, b) = \int_D \frac{1}{2} \eta |u|^2 + \frac{1}{R_o} \eta u \cdot R(x) - \frac{1}{2 \text{Fr}^2} b(\eta^2 - 2\eta h) \, dx \, dy .
\] (2.16)
This is easily shown once we have computed the variational derivatives of the Lagrangian, since these derivatives can simply be substituted into (2.8), or into one of the equivalent formulations (2.9) or (2.10). These variational derivatives are obtained by using definition 2.1 to find,
\[
\frac{\delta \ell_{\text{TRSW}}}{\delta u} = \eta u + \eta R ,
\]
\[
\frac{\delta \ell_{\text{TRSW}}}{\delta \eta} = \frac{1}{2} |u|^2 + \frac{1}{R_o} u \cdot R - \frac{1}{\text{Fr}^2} b \zeta ,
\]
\[
\frac{\delta \ell_{\text{TRSW}}}{\delta b} = -\frac{1}{2 \text{Fr}^2} (\eta^2 - 2\eta h) .
\] (2.17)

With these variations, we obtain (2.1) and by means of theorem 2.6, we see that the TRSW equations satisfy the following Kelvin circulation theorem.

Theorem 2.8 (Kelvin theorem for deterministic TRSW). The deterministic TRSW equations (2.1) imply the following Kelvin circulation law
\[
\frac{d}{dt} \oint_{c(u)} (u + \frac{1}{R_o} R) \cdot dx = \frac{1}{2 \text{Fr}^2} \oint_{c(u)} (\zeta - h) \nabla b \cdot dx = \frac{1}{2 \text{Fr}^2} \int_{\partial S=c(u)} \hat{z} \cdot \nabla (\zeta - h) \times \nabla b \, dx \, dy ,
\] (2.18)
where \(c(u)\) is any closed loop moving with horizontal fluid velocity \(u(x, t)\) in two dimensions.

Proof. This result follows from the Kelvin–Noether theorem 2.6 for Euler–Poincaré fluid equations. \[\square\]

Remark 2.9. One sees in equation (2.18) that misalignment among the horizontal gradients of free surface elevation \(\zeta\), bathymetry \(h\) or buoyancy \(b\) will generate circulation in a horizontal plane. In the sections to come, we will use theorem 2.8 to interpret the properties of the approximate equations we will derive.

Corollary 2.9.1 (Circulation on the boundary). The circulation \(\oint_{\Gamma_k} (u + \frac{1}{R_o} R) \cdot dx\) on each connected component of the boundary \(\Gamma_k \in \partial D\) is conserved by the deterministic TRSW equations.

Proof. Preservation of circulation on each connected component of the boundary follows from the boundary conditions in (2.2) and Kelvin’s theorem for TRSW in (2.18). The first boundary condition in (2.2) implies that the velocity is tangent to the boundary. Hence, a circuit \(c(u)\) on the boundary remains on the boundary; so, Kelvin’s theorem for TRSW in (2.18) applies to a boundary circuit. The second boundary condition in (2.2) implies that \(\nabla b \cdot dx = 0\) on the boundary \(\partial D\). Hence, the right-hand side of (2.18) vanishes for a circuit \(c(u)\) on the boundary and the circulation \(\oint_{\Gamma_k} (u + \frac{1}{R_o} R) \cdot dx\) is conserved. \[\square\]
The potential vorticity for the thermal rotating shallow water equations (2.1) is defined as
\[
q := \hat{z} \cdot \nabla \times u + \frac{1}{\rho_0} R_0 \eta .
\] (2.19)

Even though this is the same definition of potential vorticity as for the rotating shallow water equations, in thermal rotating shallow water, the potential vorticity is not conserved along Lagrangian paths. Rather, the potential vorticity satisfies the following equation
\[
\frac{\partial}{\partial t} q + u \cdot \nabla q = \frac{1}{2} \frac{1}{Fr^2} \eta q \hat{z} \cdot \nabla (\zeta - h) \times \nabla b .
\] (2.20)

Not unexpectedly, the mechanism responsible for the generation of circulation in the Kelvin circulation theorem 2.8 is also rate of creation of potential vorticity, \(q\), along fluid particle trajectories in equation (2.20).

**Conservation laws for deterministic TRSW.** The deterministic TRSW equations (2.1) conserve the energy
\[
E_{TRSW}(u, \eta, b) = \frac{1}{2} \int_D \eta |u|^2 + \frac{1}{Fr^2} \eta \rho (\eta - 2h) \, dx \, dy .
\] (2.21)

The conservation of energy (2.21) can be proved directly by using the TRSW equations (2.1) and the boundary conditions (2.2). The TRSW equations (2.1) also conserve an infinity of integral conservation laws, determined by two arbitrary differentiable functions of buoyancy \(\Phi(b)\) and \(\Psi(b)\) as
\[
C_{\Phi, \Psi} = \int_D \eta \Phi(b) + \bar{\omega} \Psi(b) \, dx \, dy = \int_D \left( \Phi(b) + q \Psi(b) \right) \eta \, dx \, dy ,
\] (2.22)

where \(\bar{\omega} = \eta q = \hat{z} \cdot \nabla \times (u + \frac{1}{\rho_0} R_0)\) is the total vorticity. That is, for any choice of differentiable \(\Phi\) and \(\Psi\), the quantity \(C_{\Phi, \Psi}\) is conserved in time. The conservation of (2.37) can also be proved as a direct calculation using equations (2.1) and the boundary conditions (2.2).

**Noether’s theorem.** Conservation of the integral quantities in equations (2.21) and (2.22) is associated by Noether’s theorem with smooth transformations which leave invariant the Eulerian fluid quantities in the Lagrangian [AM78]. For example, conservation of energy (2.21) arises from invariance of the Lagrangian in (2.16) under translations in time; since this Lagrangian does not depend explicitly on time. Likewise, the conserved quantities in (2.22) are associated by Noether’s theorem with the smooth flows which translate the fluid parcels along steady solutions of the equations of motion; since, of course these transformations preserve the Eulerian fluid variables in the Lagrangian [HMRW85]. Upon introducing stochasticity via the Euler–Poincaré theorem, the latter transformations and their Noether conservation laws persist. However, energy conservation does not persist because the stochastic Lagrangian depends explicitly on time through the Brownian noise. The geometrical significance of the conservation laws in equation (2.22) which persist for stochastic TRSW will be discussed further in remark 2.15.

**2.4 TRSW with stochastic advection by Lie transport (SALT)**

By modifying the fluid transport vector field in the Euler–Poincaré theorem 2.1, one can derive the stochastic equations of motion which preserve the geometric properties of their deterministic counterparts. Following [Hol15], we introduce the stochastic vector field for fluid transport
\[
d\chi_t := u(x, t) \, dt + \sum_{k=1}^M \xi_k(x) \circ dW_t^k ,
\] (2.23)

where \(\xi_k\) must satisfy the same boundary condition as \(u\). The circle notation \(\circ\) means that the stochastic integral is to be understood in the Stratonovich sense. Stratonovich calculus comes with the ordinary chain rule and product rule. These properties are written in integral form, though, because stochastic equations are not differentiable with respect to time. The sources of the stochasticity are the independent, identically distributed Brownian motions \(W_t^k\) associated to each \(\xi_k\). The Brownian motions are defined with respect to the standard probability space, see [Ito84]. One may regard the \(\xi_k(x)\) as eigenvectors
of the velocity-velocity correlation tensor. In practice, the eigenvectors $\xi_k(x)$ are obtained using the SALT algorithm developed in [CCH+18, CCH+19]. This algorithm is based on empirical orthogonal function analysis, and the number $M$ of $\xi_k$ needed in (2.23) is decided by how much of the variance is required to be represented.

The SALT version of Theorem 2.1 may be stated, as follows.

**Theorem 2.10** (Stochastic Euler–Poincaré equations [Hol15, dLHLT20]). The following two statements are equivalent:

i) The stochastic Hamilton’s variational principle in Eulerian coordinates, with $u \in X(D)$ and $b, \eta \in V^*(D)$,

$$\delta S := \delta \int_{t_1}^{t_2} \ell(u, b, \eta) \, dt = 0, \quad (2.24)$$

holds on $X(D) \times V^*$, using variations of the form

$$\delta u \, dt = dv - [d\chi, v], \quad \delta b \, dt = -(v \cdot \nabla) b \, dt, \quad \delta \eta \, dt = -\nabla \cdot (\eta v) \, dt, \quad (2.25)$$

where the vector field $v \in X(D)$ is arbitrary and vanishes on the endpoints $t_1$ and $t_2$ and $d\chi_t$ is defined in (2.23).

ii) The stochastic Euler–Poincaré equations hold. These equations are

$$\frac{d}{dt} \frac{\delta \ell}{\delta u} + (d\chi_t \cdot \nabla) \frac{\delta \ell}{\delta u} + (\nabla d\chi_t) \cdot \frac{\delta \ell}{\delta u} + \frac{\delta \ell}{\delta \eta} (\nabla \cdot d\chi_t) = -\frac{\delta \ell}{\delta b} \nabla b \, dt + \eta \nabla \frac{\delta \ell}{\delta \eta} \, dt, \quad (2.26)$$

or, equivalently, in two dimensional vector calculus notation,

$$\frac{d}{dt} \frac{\delta \ell}{\delta u} - d\chi_t \times (\nabla \cdot \frac{\delta \ell}{\delta u}) + \nabla \left( d\chi_t \cdot \frac{\delta \ell}{\delta u} \right) + \frac{\delta \ell}{\delta \eta} (\nabla \cdot d\chi_t) = -\frac{\delta \ell}{\delta b} \nabla b \, dt + \eta \nabla \frac{\delta \ell}{\delta \eta} \, dt, \quad (2.27)$$

or as an embedding in three dimensional space,

$$\frac{d}{dt} \frac{\delta \ell}{\delta u} - d\chi_t \times (\nabla \times \frac{\delta \ell}{\delta u}) + \nabla \left( d\chi_t \cdot \frac{\delta \ell}{\delta u} \right) + \frac{\delta \ell}{\delta \eta} (\nabla \cdot d\chi_t) = -\frac{\delta \ell}{\delta b} \nabla b \, dt + \eta \nabla \frac{\delta \ell}{\delta \eta} \, dt, \quad (2.28)$$

with advection equations

$$db = -d\chi_t \cdot \nabla b \quad \text{and} \quad d\eta = -\nabla \cdot (\eta d\chi_t). \quad (2.29)$$

For the proof of this theorem and the technical details we refer to [Hol15, dLHLT20]. By taking the variational derivatives of the Lagrangian for thermal rotating shallow water as in (2.17), we obtain the stochastic TRSW equations

$$du + (d\chi_t \cdot \nabla) u + \sum_k (\nabla \xi_k) \cdot u \circ dW_t^k = -\frac{1}{Fr^2} \nabla (b\zeta) \, dt + \frac{1}{2Fr^2} (\zeta - h) \nabla b \, dt$$

$$- \frac{1}{Ro} f\hat{z} \times d\chi_t - \frac{1}{Ro} \sum_k \nabla (\xi_k \cdot R) \circ dW_t^k, \quad (2.30)$$

$$d\eta + \nabla \cdot (\eta d\chi_t) = 0, \quad db + (d\chi_t \cdot \nabla) b = 0.$$  

The boundary conditions are given by

$$\hat{n} \cdot u = 0 \quad \text{and} \quad \hat{n} \cdot \xi_k = 0 \quad \text{and} \quad \hat{n} \times \nabla b = 0 \quad \text{on the boundary } \partial D. \quad (2.31)$$

The boundary condition on $\xi_k$ is required to be satisfied for any $k$. The Kelvin circulation theorem has now become stochastic, because the circulation loop is transported by the stochastic vector field $d\chi_t$, rather than by the deterministic vector field $u$. Specifically, we have:

**Theorem 2.11.** The stochastic Kelvin circulation law associated to the stochastic Euler–Poincaré theorem is

$$\int_{c(d\chi_t)} \frac{1}{\eta} \frac{\delta \ell}{\delta b} \cdot dx = -\int_{c(d\chi_t)} \frac{1}{\eta} \frac{\delta \ell}{\delta \eta} \nabla b \cdot dx \, dt, \quad (2.32)$$

where $c(d\chi_t)$ is a closed loop that is transported by the flow generated by the stochastic fluid velocity $d\chi_t$ in two dimensions.
Proof. By following the proof (2.14) of the deterministic Kelvin circulation theorem 2.6 and using the product rule and chain rule for the stochastic time differential $d$, we have

$$
\int_{c(d\mathbf{x}_t)} \frac{1}{\eta} \delta \mu_i \cdot dx = \int_{c(d\mathbf{x}_t)} \left( d + d\mathbf{x}_t \cdot \nabla \right) \left( \frac{1}{\eta} \delta \mu_i \right) \cdot dx dt
$$

$$
= -\int_{c(d\mathbf{x}_t)} \frac{1}{\eta} \delta \mu_i \cdot dx dt + \int_{c(d\mathbf{x}_t)} \delta \frac{\mu_i}{\delta \mu_j} \cdot dx dt
$$

$$
= -\int_{c(d\mathbf{x}_t)} \left( \frac{1}{\eta} \delta \mu_i \right) \cdot dx dt.
$$

(2.33)

Remark 2.12. For the stochastic TRSW equations (2.30), we have

$$
\frac{d}{dt} \int_{c(d\mathbf{x}_t)} \left( u + \frac{1}{Ro} R \right) \cdot dx = \frac{1}{2Fr} \int_{c(d\mathbf{x}_t)} \left( \zeta - h \right) \nabla b \cdot dx dt = \frac{1}{2Fr} \int_{\partial S = c(d\mathbf{x}_t)} \hat{z} \cdot \nabla \left( \zeta - h \right) \times \nabla b \ dx \ dy \ dt.
$$

(2.34)

One sees in equation (2.34) that misalignment of the horizontal gradients of free surface elevation $\zeta$, bathymetry $h$ and buoyancy $b$ will generate circulation, cf. the corresponding deterministic TRSW Kelvin circulation theorem in equation (2.18).

Remark 2.13. The evolution of potential vorticity on fluid parcels for the TRSW equations in (2.30) is given by

$$
dq + (d\mathbf{x}_t \cdot \nabla) q = \frac{1}{2Fr^2 \eta} \hat{z} \cdot \nabla (\zeta - h) \times \nabla b \ dt,
$$

(2.35)

where the potential vorticity is defined by

$$
q := \frac{\varpi}{\eta}, \quad \text{and} \quad \varpi := \hat{z} \cdot \nabla \times \left( u + \frac{1}{Ro} R \right).
$$

(2.36)

Remark 2.14. The stochastic TRSW equations (2.30) have an infinite number of conserved integral quantities

$$
C_{\Phi, \Psi} = \int_D \left( \Phi(b) + q \Psi(b) \right) \eta \ dx \ dy,
$$

(2.37)

for the boundary conditions given in (2.2) and any differentiable functions $\Phi$ and $\Psi$.

Remark 2.15. The Legendre transform which determines the Hamiltonian $d\mathbf{h}_{\text{TRSW}}$ for the stochastic TRSW equations is defined as

$$
d\mathbf{h}_{\text{TRSW}}(\mu, \eta, b) := \langle \mu, d\mathbf{x}_t \rangle - \ell_{\text{TRSW}}(u, \eta, b) dt,
$$

(2.38)

in which the angle brackets in the definition of the Legendre transform denote the $L^2$ pairing over the domain $D$. The Hamiltonian form of the stochastic TRSW equations is given for a functional $F(\mu, \eta, b)$ by

$$
dF = \left\{ F, d\mathbf{h}_{\text{TRSW}} \right\} = -\int_D \left[ \begin{array}{c} \delta F/\delta \mu_i \\ \delta F/\delta \eta \\ \delta F/\delta b \\ \\ \mu_j \partial_i + \partial_j \mu_i - \eta \partial_i b_i \end{array} \right] ^T \left[ \begin{array}{ccc} \partial j \partial_i & \eta \partial_i & -b_i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \mu_j \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \eta \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta b \\ \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \mu_j \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \eta \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta b \\ \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \mu_j \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \eta \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta b \\ \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \mu_j \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \eta \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta b \\ \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \mu_j \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta \eta \\ \delta (\mathbf{h}_{\text{TRSW}})/\delta b \end{array} \right] \ dx \ dy.
$$

(2.39)

In this notation, repeated indices are summed over. The conserved integral quantities $C_{\Phi, \Psi}$ defined in (2.37) are Casimirs of the Lie–Poisson bracket in (2.39). That is, the vector of variational derivatives of $C_{\Phi, \Psi}$ comprises a null eigenvector of the Lie–Poisson bracket in (2.39). Consequently, their conservation persists when the Hamiltonian is made stochastic. This means that the solutions of these equations describe stochastic coadjoint motion in function space on level sets of the Casimir functionals $C_{\Phi, \Psi}$. Thus, the introduction of SALT into the TRSW equations preserves the Lie–Poisson bracket in their Hamiltonian formulation and thereby preserves their geometric interpretation as coadjoint motion [HSS09].

\[1\] Notice that the Hamiltonian $d\mathbf{h}$ in (2.38) is a semimartingale. See Street and Crisan [2020].
3 Balanced interpretations of TRSW

There exist several approximations of the rotating shallow water (RSW) equations, the most famous one being the quasi-geostrophic (QG) approximation. By assuming the motion to take place in a particular scaling regime, it can be shown that the largest component of the velocity field, called the geostrophic velocity field, is determined by a diagnostic equation, rather than a prognostic equation. The QG approximation is a small perturbation around this geostrophic velocity field. There exists an intermediate model which is more accurate than QG, but is still an approximation of RSW. In this section, we will derive the thermal geostrophic balance by identifying the correct scaling regime and use asymptotic expansions to simplify the TRSW equations. Next, we will show that the thermal rotating shallow water equations can be approximated geometrically to derive a class of equations which was first proposed by Eliassen [Eli49] and made into a variational theory by [Sal83], where it is called L1. The Lagrangian corresponding to the equations proposed by Eliassen can be obtained via two approaches. The first approach involves the Helmholtz decomposition and the second approach follows [AH96]. The methods of [AH96] will be applied in the Euler–Poincaré framework to derive the corresponding equations of motion. Finally, the stochastic L1 equations will be derived via the stochastic Euler–Poincaré theorem.

3.1 Thermal geostrophic balance

To obtain the thermal geostrophic balance relation, we return to the nondimensional deterministic TRSW equations for the Eulerian horizontal vector velocity $\mathbf{u}(x, t)$, thickness $\eta(x, t)$, buoyancy $b(x, t)$, and free surface elevation $\zeta = \eta - h$, with mean depth $h(x)$, given in (2.1) by

$$
\frac{D}{Dt} \mathbf{u} + \frac{1}{Ro} f \hat{z} \times \mathbf{u} = -\frac{1}{Fr^2} \nabla (b\zeta) + \frac{1}{2Fr^2} (\zeta - h) \nabla b, \quad \frac{\partial \eta}{\partial t} + \nabla \cdot (\eta \mathbf{u}) = 0, \quad \frac{D}{Dt} b = 0, \quad (3.1)
$$

with boundary conditions in (2.2). In order to find an asymptotic balance among these equations, a number of assumptions are necessary. First, we assume that the free surface elevation $\zeta = \eta - h$ is small, denoted as $\zeta = \epsilon \zeta_1$ of order $O(\epsilon)$ with $\epsilon \ll 1$. Second, in line with the Boussinesq approximation in three dimensional fluids, we assume that the buoyancy is also small, denoted as $b(x, t) = 1 + \epsilon b_1(x, t)$. Third, we assume that the following ratio involving the three dimensionless numbers in this approximation is of order $O(1)$:

$$
\frac{\epsilon Ro}{Fr^2} \sim O(1). \quad (3.2)
$$

Assumption (3.2) is equivalent to requiring that $O(\epsilon) = O(Ro^{-1}Fr^2)$. Now, in the QG approximation one also assumes the gradients of bathymetry and Coriolis parameter are small, of order $O(Ro)$. In combination, this amounts to the following set of assumptions:

$$
\zeta(x, t) = \epsilon \zeta_1(x, t), \\
b(x, t) = 1 + \epsilon b_1(x, t), \\
f(x) = 1 + Ro f_1(x), \\
h(x) = 1 + Ro h_1(x). \quad (3.3)
$$

These assumptions were also made in derivation of this model in [HL19] and they are sufficient to derive the TRSW balance relation, which we will show now. First, we multiply the momentum equation in (3.1) by the Rossby number $Ro \ll 1$, then we substitute the assumptions (3.3) into the momentum equation and use the scaling relation (3.2) to find

$$
\z \times \mathbf{u} + Ro \frac{D}{Dt} \mathbf{u} + Ro f_1 \hat{z} \times \mathbf{u} = -\nabla \zeta_1 - \frac{1}{2} \nabla b_1 - \epsilon b_1 \nabla \zeta_1 - \frac{1}{2} (\epsilon \zeta_1 + Ro b_1) \nabla b_1. \quad (3.4)
$$

Note that the ordering in (3.2) is satisfied in the special case when $O(\epsilon) = O(Ro) = O(Fr)$. This assumption allows us to continue the analysis with a single small parameter. We shall write $\epsilon \ll 1$ for this small parameter, so that (3.4) leads to

$$
\z \times \mathbf{u} + \epsilon \frac{D}{Dt} \mathbf{u} + \epsilon f_1 \hat{z} \times \mathbf{u} = -\nabla \zeta_1 - \frac{1}{2} \nabla b_1 - \epsilon b_1 \nabla \zeta_1 - \frac{\epsilon}{2} (\zeta_1 + h_1) \nabla b_1. \quad (3.5)
$$
Thus, equation (3.5) implies the following relation

$$\hat{z} \times u = -\nabla \zeta_1 - \frac{1}{2} \nabla b_1 + O(\epsilon).$$

(3.6)

By operating with $\hat{z} \times$ on the TRSW balance relation (3.6), we find the defining expression for the divergence-free thermal geostrophic velocity field, denoted as $u_{TG}$,

$$u_{TG} = \hat{z} \times \nabla \left( \zeta_1 + \frac{1}{2} b_1 \right) =: \hat{z} \times \psi.$$  

(3.7)

Here, $\psi(x,t)$ is the stream function for the divergence-free leading order thermal geostrophic velocity vector field $u_{TG}$. Relation (3.6) allows the total fluid velocity to be represented as the sum of the leading order thermal geostrophic velocity $u_{TG}$ and higher order terms,

$$u = u_{TG} + \epsilon v$$  

(3.8)

In particular, (3.8) implies that the difference, referred to as the ageostrophic component of the velocity field, satisfies $u - u_{TG} = \epsilon v = O(\epsilon)$. This decomposition of the velocity field into a leading order divergence-free part plus higher order parts is similar to the Helmholtz decomposition, except the divergence free component is allowed to have both a leading order and a higher order part, while the rotation free component has only higher order parts.

3.2 Moving into the balanced frame

One may transform the thermal rotating shallow water equations in (2.1) into a time-dependent local frame moving with the thermal geostrophically balanced velocity, $u_{TG}(x,t)$, by inserting the decomposition (3.8) into the Lagrangian for the TRSW equations (2.16),

$$\ell_{TRSW}(u, b, \eta) = \int_D \frac{1}{2} \eta|u|^2 + \frac{1}{\epsilon} \eta u \cdot R - \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta h) \, dx \, dy$$

$$= \int_D \epsilon \left( \frac{\epsilon}{2} |v|^2 + v \cdot u_{TG} + \frac{1}{\epsilon} v \cdot R \right) + \eta \left( \frac{1}{2} |u_{TG}|^2 + \frac{1}{\epsilon} u_{TG} \cdot R \right)$$

$$- \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta h) \, dx \, dy.$$  

(3.9)

One can apply Hamilton’s variational principle to the Lagrangian (3.9) $0 = \delta S$ with $S = \int_{t_1}^{t_2} \ell_{TRSW} \, dt$. We substitute the velocity decomposition in (3.8) to define the variation $\delta u = \delta u_{TG} + \epsilon \delta v$ with

$$\delta u_{TG} = \hat{z} \times \nabla \left( \delta \eta + \frac{1}{2} \delta b \right).$$  

(3.10)

Hamilton’s principle then yields the Euler–Poincaré equation (2.12) in the form,

$$\frac{\partial}{\partial t} \left( \frac{1}{\eta} \frac{\delta \ell}{\delta \eta} \right) + (u \cdot \nabla) \left( \frac{1}{\eta} \frac{\delta \ell}{\delta \eta} \right) + (\nabla u) \cdot \left( \frac{1}{\eta} \frac{\delta \ell}{\delta \eta} \right) = - \frac{1}{\eta} \frac{\delta \ell}{\delta b} \nabla b + \nabla \frac{\delta \ell}{\delta \eta}.$$  

(3.11)

Thus, the relative motion equation for TRSW dynamics in the frame moving with the thermal geostrophic balance velocity $u_{TG}(x,t)$ in (3.7) keeps its Euler–Poincaré form (2.12). Upon eliminating $\partial_t u_{TG}$ in (3.11) by using the advection equations for $(b, \eta)$ in (3.1), the system closes and thereby transforms the TRSW equations into the new variables $(v, b, \eta)$ in the reference frame moving with velocity $u_{TG}(x,t)$.

**Stationary thermal geostrophic balance as “mean dynamic topography”**. To a good approximation, much of upper ocean dynamics is well-approximated by a mean dynamic topography (MDT), which is monitored continuously with in situ instruments and satellites, see, e.g., [MNC+09]. Ocean dynamics is then envisioned as time-dependent variations in the steady moving frame of the MDT. To apply this idea to TRSW dynamics, we envision TRSW dynamics as taking place in the moving reference frame defined by a time-independent mean thermal geostrophic velocity $\overline{u}_{TG}(x)$. In this regard, the Kelvin theorem 2.18 for deterministic TRSW derived from equation (3.11) takes the following form.
Theorem 3.1 (Kelvin theorem for deterministic TRSW in a stationary balanced frame). The deterministic TRSW equations (2.1) imply the following Kelvin circulation law in a stationary TG balanced frame moving with time-independent velocity $\mathbf{u}_{TG}(x)$,

$$\frac{d}{dt} \int_{c(v)} \left( \epsilon \mathbf{v} + \mathbf{u}_{TG}(x) + \frac{1}{\epsilon} \mathbf{R}(x) \right) \cdot d\mathbf{x} = \frac{1}{2\epsilon^2} \int_{c(v)} (\zeta - h) \nabla b \cdot d\mathbf{x} = \frac{1}{2\epsilon^2} \int_{\partial S = c(v)} \hat{z} \cdot \nabla (\zeta - h) \times \nabla b \, dx \, dy, \quad (3.12)$$

where $c(v)$ is any closed loop moving with horizontal fluid velocity $\mathbf{v}(x,t)$ relative to the frame of motion whose velocity is $\mathbf{u}_{TG}(x) + \epsilon^{-1}\mathbf{R}(x)$ in two horizontal dimensions.

Remark 3.2. Thus, in the mean thermal geostrophic balance frame, the frame velocity $\mathbf{u}_{TG}(x)$ simply adds another contribution to the momentum per unit mass. In turn, this contributes an additional ‘Coriolis’ force in the dynamics of the relative velocity $\mathbf{v} = \mathbf{u} - \mathbf{u}_{TG}$. The corresponding SALT version in this case would simply replace $\mathbf{u}(x,t)$ by $\mathbf{v}(x,t)$ in the drift velocity of the stochastic vector field $d\mathbf{x}_t$ defined in (2.23).

Next, we will consider the thermal versions of two of the classic GFD approximations of RSW developed previously in the absence of buoyancy. Namely, we will consider thermal versions of the Eliassen approximation and the quasigeostrophic approximation.

### 3.3 The Eliassen approximation

The starting point in deriving the thermal Eliassen approximation is the Lagrangian for the thermal rotating shallow water equations (2.16)

$$\ell_{TRSW} = \int \frac{1}{2} \eta |\mathbf{u}|^2 + \frac{1}{\Ro} \mathbf{u} \cdot \mathbf{R} - \frac{1}{2 \Fr^2} b(\eta^2 - 2\eta h) \, dx \, dy. \quad (3.13)$$

Since the thermal geostrophic velocity field (3.7) is divergence free, it is useful to transform the velocity variables inside the Lagrangian to vorticity and divergence by using the Helmholtz decomposition. The forward transformation is $(\mathbf{u}, \eta, b) \mapsto (\omega, D, \eta, b)$ and amounts to

$$\omega = \hat{z} \cdot \nabla \times \mathbf{u},$$
$$D = \nabla \cdot \mathbf{u},$$
$$\eta = \eta,$n
$$b = b. \quad (3.14)$$

The inverse transformation is unique if the kernel of the Laplacian is trivial, which is the same as saying that there are no harmonic functions for the domain $\mathcal{D}$ and the boundary conditions on $\partial \mathcal{D}$. In this case, the inverse transformation $(\omega, D, \eta, b) \mapsto (\mathbf{u}, \eta, b)$ is given by

$$\mathbf{u} = \hat{z} \times \nabla \Delta^{-1} \omega + \nabla \Delta^{-1} D,$n
$$\eta = \eta,$n
$$b = b. \quad (3.15)$$

The inverse transformation (3.15) uniquely defines the vector $\mathbf{u}$ in terms of its divergence and curl. The inverse of the Laplacian can be interpreted in terms of the Green’s function in two dimensions,

$$\Delta^{-1} F = -\frac{1}{2\pi} \int \ln \|x - x’\| F(x') \, dx’. \quad (3.16)$$

Changing variables using the classical Helmholtz decomposition leads to the following formulation of the Lagrangian for thermal rotating shallow water

$$\ell_{TRSW} = \int \frac{1}{2} \eta |\nabla \Delta^{-1} \omega|^2 + \eta J (\Delta^{-1} \omega, \Delta^{-1} D) + \frac{1}{2} \eta |\nabla \Delta^{-1} D|^2$$
$$+ \frac{1}{\Ro} \left( \hat{z} \times \nabla \Delta^{-1} \omega + \nabla \Delta^{-1} D \right) \cdot \mathbf{R} - \frac{1}{2 \Fr^2} b(\eta^2 - 2\eta h) \, dx \, dy. \quad (3.17)$$

It should be noted that the vorticity and divergence are not orthogonal in a weighted $L^2$ space, so the Jacobian term $\int \eta J (\Delta^{-1} \omega, \Delta^{-1} D) \, dx \, dy \neq 0$. The Jacobian term does vanish in the standard $L^2$ space, in which $\int J (\Delta^{-1} \omega, \Delta^{-1} D) \, dx \, dy = 0$. By means of the ordering $\mathcal{O}(\epsilon) = \mathcal{O}(\Ro) = \mathcal{O}(\Fr)$ and the decomposition of $\mathbf{u}$
into a thermal geostrophic part and a higher order part (3.8), we can take the following asymptotic expansions for the vorticity and the divergence

\[ \omega = \omega_0 + \epsilon \omega_1 + o(\epsilon), \]
\[ D = \epsilon D_1 + o(\epsilon). \]  

(3.18)

The thermal geostrophic balance implies a decomposition of the velocity field in terms a leading order divergence free component and a higher order general component. This means that we can identify \( \omega_0 \) with the curl of the thermal geostrophic velocity field (3.7), but it also means that we must keep a higher order vorticity term around. Substituting (3.18) into the Lagrangian yields

\[
\ell_{\text{TRSW}} = \int \frac{1}{2} \eta \left| \nabla \Delta^{-1} (\omega_0 + \epsilon \omega_1) \right|^2 + \epsilon \eta \left| \hat{\omega} \times \nabla \Delta^{-1} (\omega_0 + \epsilon \omega_1) + \epsilon \nabla \Delta^{-1} D_1 \right| \cdot \mathbf{R} - \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta \theta) \, dx \, dy + o(1).
\]

(3.19)

By expanding and collecting all terms that are of higher order than \( O(1) \), the Lagrangian can be written as

\[
\ell_{\text{TRSW}} = \int \frac{1}{2} \eta \left| \nabla \Delta^{-1} \omega_0 \right|^2 + \frac{1}{\epsilon} \eta \left| \hat{\omega} \times \nabla \Delta^{-1} (\omega_0 + \epsilon \omega_1) + \epsilon \nabla \Delta^{-1} D_1 \right| \cdot \mathbf{R} - \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta \theta) \, dx \, dy + o(1).
\]

(3.20)

We now use the fact that the velocity field also decomposes into a thermal geostrophic part and an ageostrophic part and apply the inverse transformation to recover the original fluid variables. This yields

\[
\ell_{\text{L1}} = \int \frac{1}{2} \eta |u_{TG}|^2 + \frac{1}{\epsilon} \eta u \cdot \mathbf{R} - \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta \theta) \, dx \, dy .
\]

(3.21)

The subscript \( TL1 \) refers to the thermal \( L1 \) model, which is an extension of [Sal83] to include horizontal variations in buoyancy and bathymetry. However, at this stage we do not have enough information to execute the variational principle. To obtain the information we need, we will introduce a higher order term which completes the Lagrangian, by allowing us to vary with respect to the full velocity field \( u \), interpreted as a Lagrange multiplier

\[
\ell_{\text{L1}} = \int u \cdot \left( \epsilon u_{TG} + \frac{1}{\epsilon} \eta \mathbf{R} \right) - \frac{1}{2} \eta |u_{TG}|^2 - \frac{1}{2\epsilon^2} b(\eta^2 - 2\eta \theta) \, dx \, dy + O(\epsilon).
\]

(3.22)

Equivalently, one can truncate the Lagrangian for TRSW rewritten in the balanced frame (3.9) at \( O(1) \) to obtain this Lagrangian. This emphasises the fact that a component of the velocity field, when it can be expressed in terms of the other variables in the problem, can be used to change the reference frame. At this point, several important questions arise. How does one take variations of this Lagrangian? Can the equation for the Lagrange multiplier \( u \) be found? To answer these questions, we use the methods of [AH96].

### 3.4 The Allen-Holm approach

The [AH96] approach is based on the following observation. Given a hyperregular Lagrangian \( \ell \), one can obtain the corresponding Hamiltonian \( h \) via the Legendre transform

\[
h(m, \eta) = \int \left( m - \frac{\delta \ell}{\delta u} \right) \cdot u + \mathcal{E}(u, \eta, b, \nabla \eta, \nabla b, \text{etc.}),
\]

(3.23)

where \( \mathcal{E} \) can be interpreted as the energy density. The momentum density \( m \) is given in terms of the other fluid variables by the condition \( \delta h / \delta u = 0 \), where \( h \) is the Hamiltonian defined by the Legendre transform in (3.23). In the Legendre transform, the fluid velocity \( u \) appears as a Lagrange multiplier which enforces the relation of \( m \) to the other fluid variables as a dynamically preserved constraint. This definition is usually taken for granted, but in what follows, we shall model the momentum density as a \textit{prescribed} function of the other fluid variables. This means that we will define

\[
m = \overline{m}(\eta, b, \nabla \eta, \nabla b, \text{etc.}) =: \overline{m}[\eta, b].
\]

(3.24)
In this type of modelling, it is necessary to have the explicit enforcement of the momentum definition (3.24), both as a constraint as well as a means of determining the fluid velocity for the model by using Lagrange multipliers. We rearrange the Lagrangian in (3.13) using $\mathbf{m}$ as the momentum density, defined by

$$
\mathbf{m} := \frac{\delta \mathcal{L}_{\text{TRSW}}}{\delta \mathbf{u}} = \eta \mathbf{u} + \frac{1}{Ro} \eta \mathbf{R}.
$$

The Lagrangian in (3.13) can then be written as

$$
\ell_{\text{TRSW}} = \int_D \mathbf{m} \cdot \mathbf{u} - \frac{1}{2} \eta |\mathbf{u}|^2 - \frac{1}{2 \frac{Fr^2}{Ro}} b(\eta^2 - 2\eta) \, dx \, dy.
$$

The role of $\mathbf{m}$ is a transformation of the reference frame. In the original TRSW Lagrangian (2.16), $\mathbf{R}$ puts the motion in a rotating frame. In (3.26), $\mathbf{m}$ changes the frame to the one that is moving with the fluid velocity as well as the rotation due to the vector potential $\mathbf{R}$. Changes of reference frame are common in geophysical fluid dynamics, see for instance [Hol20], where a wave–mean flow decomposition leads to a change of frame. We can substitute the decomposition of the velocity field into geostrophic and ageostrophic components (3.8) with $\mathbf{u}_{\text{TG}}$ defined in (3.7) into the Lagrangian (3.26) and then obtain without approximation the following Lagrangian, which is linear in the velocity $\mathbf{u}$,

$$
\ell_{\text{TRSW}} = \int_D \mathbf{u} \cdot \left( \eta \mathbf{u}_{\text{TG}} + Ro \frac{1}{\eta} \mathbf{u}_A + \frac{1}{Ro} \eta \mathbf{R} \right) - \frac{1}{2} \eta |\mathbf{u}_{\text{TG}} + Ro \mathbf{u}_A|^2 - \frac{1}{2 \frac{Fr^2}{Ro}} b(\eta^2 - 2\eta) \, dx \, dy.
$$

In line with the ordering scheme $\mathcal{O}(\epsilon) = \mathcal{O}(Ro) = \mathcal{O}(Fr)$, we formulate the Lagrangian in terms of a single parameter $\epsilon$. When $\epsilon \ll 1$, one could simply drop the $\mathbf{u}_A$ terms in the Lagrangian to obtain

$$
\ell_{TL1} = \int_D \mathbf{u} \cdot \left( \eta \mathbf{u}_{\text{TG}} + \frac{1}{\epsilon} \mathbf{R} \right) - \frac{1}{2} \eta |\mathbf{u}_{\text{TG}}|^2 - \frac{1}{2 \epsilon^2} b(\eta^2 - 2\eta) \, dx \, dy.
$$

Note that this is the Lagrangian obtained in (3.22) by using the Helmholtz decomposition to decompose $\mathbf{u}$ into vorticity and divergence.

**Remark 3.3.** Since we will use (3.7) as our definition for $\mathbf{u}_{\text{TG}}$, we should keep in mind that according to strict asymptotics the potential energy term should also be expanded using the same assumptions that led to (3.7). By keeping the Lagrangian in the form (3.28) we have included higher order terms, but not all of them, since we have truncated the kinetic energy. In terms of strict asymptotics this means that we do not have a balance among terms in the Lagrangian. A benefit of not expanding the potential energy at this stage is that the variational derivatives can be taken in the usual way and are thus closer to the variational derivatives of the TRSW system.

The approximate Lagrangian (3.28) is also linear in the velocity $\mathbf{u}$, since it has the form

$$
\ell_{TL1} = \int \mathbf{m}[\eta, b] \cdot \mathbf{u} - \mathcal{E}[\eta, b] \, dx \, dy,
$$

with

$$
\mathbf{m}[\eta, b] = \frac{\delta \ell_{TL1}}{\delta \mathbf{u}} = \frac{1}{\epsilon} \eta \mathbf{R} + \eta \mathbf{u}_{\text{TG}} \quad \text{and} \quad \mathcal{E}[\eta, b] = \frac{1}{2} \eta |\mathbf{u}_{\text{TG}}|^2 + \frac{1}{2 \epsilon^2} b(\eta^2 - 2\eta).
$$

Note that $\mathbf{u}_{\text{TG}}$ can be expressed in terms of $\eta$ and $b$. At this stage, in [AH96], the next step after having obtained the Lagrangian in the form above would have been to take the Legendre transformation and obtain the Hamiltonian. Then, by requiring the first variation of the Hamiltonian to vanish, one would obtain the equations of motion. In [HMR98] it was shown, however, that one can obtain the same equations of motion by applying the variational principle on the Lagrangian side, by means of the Euler–Poincaré theorem. The first variation of the Lagrangian is given by

$$
\delta \ell_{TL1} = \int_D \left( \eta \mathbf{u}_{\text{TG}} + \frac{1}{\epsilon} \eta \mathbf{R} \right) \cdot \delta \mathbf{u} + \left( \eta (\mathbf{u} - \mathbf{u}_{\text{TG}}) \right) \cdot \delta \mathbf{u}_{\text{TG}}
$$

$$
+ \left( \mathbf{u}_{\text{TG}} \cdot \mathbf{u} + \frac{1}{\epsilon} \mathbf{u} \cdot \mathbf{R} - \frac{1}{2} |\mathbf{u}_{\text{TG}}|^2 - \frac{1}{\epsilon^2} b(\eta^2 - 2\eta) \right) \delta \eta
$$

$$
+ \left( \frac{1}{2 \epsilon^2} (\eta^2 - 2\eta) \right) \delta b \, dx \, dy.
$$
From the definition of \( u_{TG} \) in \((3.7)\), we now substitute
\[
\delta u_{TG} = \frac{1}{\epsilon} \hat{z} \times \nabla \delta \eta + \frac{1}{2 \epsilon} \hat{z} \times \nabla \delta b. \tag{3.32}
\]
Integration by parts in \((3.31)\) then yields
\[
\delta \ell_{TL1} = \int_D \left( \eta u_{TG} + \frac{1}{\epsilon} \eta R \right) \cdot \delta u + \left( u_{TG} \cdot u + \frac{1}{\epsilon} R \cdot u - \frac{1}{2} |u_{TG}|^2 - \frac{1}{\epsilon} b c - \frac{1}{\epsilon} \hat{z} \cdot \nabla \left( \eta (u - u_{TG}) \right) \right) \delta \eta
+ \left( -\frac{1}{2 \epsilon^2} \eta \delta b \right) \cdot \nabla \cdot \left( \eta (u - u_{TG}) \right) \delta \eta \tag{3.33}
\]
\[-\oint_{\partial D} \hat{z} \times \left( \eta (u - u_{TG}) \right) \left( \delta \eta + \frac{1}{2} \delta b \right) \cdot dx.
\]
The integral over the boundary vanishes provided that the ageostrophic velocity field has no tangential component on the boundary. By means of the Euler–Poincaré theorem, we find the equations of motion given in a form first proposed by Eliassen \cite{Eli49} as
\[
\frac{\partial}{\partial t} u_{TG} - u \times \left( \nabla \times \left( u_{TG} + \frac{1}{\epsilon} R \right) \right) + \nabla \left( \frac{1}{\epsilon} b c + \frac{1}{\epsilon} u_{TG} |^2 + \frac{1}{\epsilon} B \right) - \left( \frac{1}{2 \epsilon^2} (\zeta - h) + \frac{1}{2 \epsilon \eta} \right) \nabla b = 0, \tag{3.34}
\]
\[
\frac{\partial}{\partial t} \eta + \nabla \cdot (\eta u) = 0, \quad \frac{\partial}{\partial t} b + u \cdot \nabla b = 0.
\]
The notation \( B \) in the first of these equations is defined by
\[
B := \hat{z} \cdot \nabla \left( \eta (u - u_{TG}) \right). \tag{3.35}
\]
The function \( B \) keeps track of effects that are generated by higher order vorticity terms, since \( B \) includes the curl of \( u - u_{TG} = O(\epsilon) \). These higher order vorticity terms will contribute in the Kelvin circulation theorem \(3.4\) arising from equations \((3.34)\). Equations \((3.34)\) extend Salmon’s \( L1 \) model \cite{Sal83} to include horizontal buoyancy variations and bottom topography. The boundary conditions carry over from thermal rotating shallow water and are given by
\[
\hat{n} \cdot u = 0 \quad \text{and} \quad \hat{n} \times \nabla b = 0 \quad \text{on the boundary} \ \partial D. \tag{3.36}
\]
Having been derived from the Euler–Poincaré variational principle \cite{HMR98}, the deterministic TL1 equations in \((3.34)\) satisfy the following Kelvin–Noether circulation theorem.

**Theorem 3.4** (Kelvin theorem for the deterministic TL1 model). The deterministic TL1 equations \((3.34)\) imply the following Kelvin circulation law
\[
\frac{d}{dt} \oint_{\partial S} \left( u_{TG} + \frac{1}{\epsilon} R \right) \cdot dx = \oint_{\partial S} \left( \frac{1}{2 \epsilon^2} (\zeta - h) + \frac{1}{2 \epsilon \eta} B \right) \nabla b \cdot dx \tag{3.37}
\]
\[
= \oint_{\partial S} \hat{z} \cdot \nabla \left( \frac{1}{2 \epsilon^2} (\zeta - h) + \frac{1}{2 \epsilon \eta} B \right) \times \nabla b \, dx \, dy.
\]

**Proof.** This result follows from the Kelvin–Noether theorem \(2.6\) for Euler–Poincaré fluid equations. \(\square\)

**Remark 3.5.** The Kelvin circulation theorem \(3.4\) for the deterministic TL1 model \((3.34)\) implies that the misalignment of the horizontal gradients of the free surface elevation and the bathymetry with the horizontal gradient of the buoyancy generates circulation. This result is similar to the corresponding Kelvin circulation theorem \(2.8\) for the deterministic thermal rotating shallow water (TRSW) model. An additional contribution to the generation of circulation relative to theorem \(2.8\) is made by the misalignment of the gradient of the quantity \((B/\eta)\) defined in \((3.35)\) with the gradient of the buoyancy. This additional contribution is due to misalignment of the horizontal gradients of the ageostrophic vorticity and the buoyancy.

The potential vorticity \( q \) for TL1 is defined as
\[
q = \frac{1}{\eta} \left( \omega + \frac{1}{\epsilon} f \right) = \frac{1}{\epsilon \eta} \left( \Delta \left( \frac{1}{2} h + f \right) \right). \tag{3.38}
\]
Note that the potential vorticity \( q \) in (3.38) contains a term in the Coriolis parameter which is order \( O(\epsilon^{-1}) \). This feature will become important in the asymptotic expansion of \( q \) later, in deriving the thermal QG model at the beginning of section 5. In the presence of buoyancy, the evolution of potential vorticity along Lagrangian fluid trajectories is not conserved. Instead, potential vorticity is generated, as indicated in the circulation theorem (3.37) via misalignment of gradients in

\[
\eta \frac{\partial}{\partial t} q + \mathbf{u} \cdot \nabla q = \frac{1}{\epsilon} \mathbf{z} \cdot \nabla \left( \frac{1}{2\epsilon^2}(\zeta - h) + \frac{1}{2\epsilon^2} \eta B \right) \times \nabla b.
\]  

(3.39)

Although the potential vorticity is not conserved along Lagrangian fluid trajectories, the TL1 equations do preserve energy, as well as an infinity of integral conservation laws involving buoyancy and potential vorticity.

**Conservation laws for deterministic TL1.** The deterministic TL1 equations (3.34) conserve the energy

\[
\mathcal{E}_{TL1}(\mathbf{u}_{TG}, \eta, b) = \frac{1}{2} \int_D \eta |\mathbf{u}_{TG}|^2 + \frac{\eta}{\epsilon^2}(\eta - 2h) \, dx \, dy.
\]  

(3.40)

Equations (3.34) also conserve an infinity of integral conservation laws, determined by two arbitrary differentiable functions of buoyancy \( \Phi(b) \) and \( \Psi(b) \) as

\[
C_{\omega, \Psi} = \int_D \eta \Phi(b) + \epsilon \Psi(b) \, dx \, dy = \int_D \left( \Phi(b) + \epsilon \Psi(b) \right) \eta \, dx \, dy,
\]  

(3.41)

where \( \omega \) and \( q \) are defined in equation (3.38). Notice that this family of integral conservation laws for the TL1 equations has the same form as the family of integral conserved quantities for the TRSW equations, defined in (2.22). The proof that \( C_{\omega, \Psi} \) is conserved in time, for any choice of differentiable \( \Phi \) and \( \Psi \), follows from a direct calculation involving the boundary conditions (3.36). The integral conserved quantities in equations (5.13) and (3.41) are associated with smooth transformations which leave invariant the Eulerian fluid quantities in the Lagrangian. As with the TRSW equations, upon introducing stochasticity via the Euler–Poincaré theorem 2.10, the latter conservation laws persist. However, energy conservation does not persist because the stochastic Lagrangian depends explicitly on time through the Brownian motion. In order to use the TL1 equations (3.34) as a predictive model, one needs to be able to determine the Lagrange multiplier \( \mathbf{u} \) from the other variables in the model. This will be our next task.

### 3.4.1 Determining the Lagrange multiplier

By operating with \( \hat{\mathbf{z}} \times \) on the momentum equation in (3.34) and using the definition of \( \mathbf{u}_{TG} \) in (3.7), we obtain

\[
- \frac{1}{\epsilon} \frac{\partial}{\partial t} \mathbf{z} \times \nabla \left( \eta + \frac{1}{2} b \right) + \eta \mathbf{u} = \mathbf{z} \times \nabla \left( \frac{1}{\epsilon^2} b \zeta + \frac{1}{2} |\mathbf{u}_{TG}|^2 + \frac{1}{\epsilon} B \right) - \left( \frac{1}{2\epsilon^2}(\zeta - h) + \frac{1}{2\epsilon^2} B \right) \mathbf{z} \times \nabla b.
\]  

(3.42)

The first term above follows from using the definition of \( \mathbf{u}_{TG} \), by noting that the bathymetry has no time derivative. This allows us to rewrite \( \mathbf{u}_{TG} \) in terms of \( \eta \) and \( b \), rather than \( \zeta_1 \) and \( b_1 \). Taking the time derivative through the gradient in (3.42) allows us to use the continuity equation and the advection equation to obtain an elliptic equation. Substituting the definition for \( \dot{B} \) in (3.35) then leads to the following equation which determines the Lagrange multiplier \( \mathbf{u} \),

\[
- \frac{1}{\epsilon} \mathbf{z} \times \nabla \left( \eta \mathbf{u} + \frac{1}{2} \mathbf{u} \cdot \nabla b \right) + \mathbf{z} \times \nabla \left( \frac{1}{\epsilon^2} b \zeta + \frac{1}{2} |\mathbf{u}_{TG}|^2 + \frac{1}{\epsilon^2} \mathbf{z} \cdot \nabla \times \left( \eta (\mathbf{u} - \mathbf{u}_{TG}) \right) \right) - \left( \frac{1}{2\epsilon^2}(\zeta - h) + \frac{1}{2\epsilon^2} \eta \mathbf{u} = 0. \right.
\]  

(3.43)

Before going to the general case, let us consider the case in which the horizontal gradient of buoyancy vanishes.

**No horizontal buoyancy gradients.** In this case, equation (3.43) reduces to

\[
\frac{1}{\epsilon} \nabla \cdot (\nabla \cdot \mathbf{u}) + \nabla (\nabla \cdot \mathbf{u}) + \eta \mathbf{u} = 0.
\]  

(3.44)

This is the diagnostic partial differential equation used to determine \( \mathbf{u} \) in [Sal83] when variations in bathymetry are absent and it is identical to equation (3.16) in [AH96]. After applying the identity

\[
\Delta \mathbf{u} = \nabla (\nabla \cdot \mathbf{u})
\]  

(3.45)
in equation (3.44), we can rewrite the diagnostic equation (3.43) in simpler form. Here we have used the perpendicular "⊥" notation for brevity, see remark 2.3. The Laplacian identity (3.45) implies that the equation which determines \( \mathbf{u} \) is a linear non-autonomous elliptic partial differential equation (PDE), given by

\[
\frac{1}{\epsilon} \Delta (\eta \mathbf{u}) + q\eta \mathbf{u} = -\nabla \perp \left( \frac{1}{\epsilon^2} \zeta + \frac{1}{2} |\mathbf{u}_{TG}|^2 - \frac{1}{\epsilon} \nabla \perp \cdot \eta \mathbf{u}_{TG} \right).
\]

(3.46)

Note that the coefficient \( q\eta \) in (3.46) is the total vorticity, since \( q\eta \mathbf{u} = \frac{1}{\epsilon} (f + \Delta \zeta) \mathbf{u} \). The solution behaviour of the elliptic equation (3.46) for the quantity \( \eta \mathbf{u} \) depends on the sign of the potential vorticity, \( \eta \), in the following three cases

1. \( \eta > 0 \). The equation for \( \eta \mathbf{u} \) is a weighted, inhomogeneous Helmholtz equation.
2. \( \eta < 0 \). The equation for \( \eta \mathbf{u} \) is a weighted, screened Poisson equation.
3. \( \eta = 0 \). The equation for \( \eta \mathbf{u} \) is a weighted Poisson equation.

In the present situation, there are no horizontal buoyancy gradients. Consequently, the potential vorticity \( \eta \) is preserved along Lagrangian particle trajectories and does not change sign during the calculation. This means that such changes in the solution behaviour of (3.46) do not occur. Thus, the ‘equator’, where \( f \) changes sign, acts as a boundary between the ‘northern and southern hemispheres’, in the absence of horizontal buoyancy gradients. In this case, Lagrangian particles which start in the northern hemisphere stay in the northern hemisphere, because their potential vorticity is conserved in the absence of horizontal buoyancy gradients. The case with horizontal buoyancy gradients is the general case, which we will discuss now.

**General case.** When horizontal gradients of buoyancy are nonzero, equation (3.46) for the determination of the Lagrange multiplier \( \mathbf{u} \) becomes considerably more extensive

\[
\frac{1}{\epsilon} \Delta (\eta \mathbf{u}) + \frac{1}{2\epsilon} \nabla (\mathbf{u} \cdot \nabla b) - \frac{1}{2\epsilon} (\nabla \perp \cdot \mathbf{u}) \nabla \perp b - \frac{1}{2\epsilon \eta} (\mathbf{u} \cdot \nabla \perp \eta) \nabla \perp b + \eta q\mathbf{u} = -\frac{1}{\epsilon^2} \nabla \perp (b\zeta) + \frac{1}{2\epsilon^2} (\zeta - h) \nabla \perp b
\]

\[
- \frac{1}{2} \nabla \perp |\mathbf{u}_{TG}|^2 - \frac{1}{2\epsilon \eta} (\nabla \perp \cdot \eta \mathbf{u}_{TG}) \nabla \perp b + \frac{1}{\epsilon} \nabla \perp \cdot (\eta \mathbf{u}_{TG})
\]

(3.47)

The coefficient for the \( \mathbf{u} \) term now has an additional contribution from the perpendicular gradient of the buoyancy, which changes the conditions for the type of PDE. The zeroth order terms in the presence of horizontal buoyancy gradients indicate that the equator is no longer a stationary boundary between the northern and southern hemispheres. Indeed, the perpendicular gradients of buoyancy in combination with perpendicular gradients of the depth have removed the ‘equatorial boundary’. Likewise, when the horizontal gradient of buoyancy is included, the potential vorticity is not conserved along Lagrangian particle trajectories. Moreover, the effects of the first order terms at this point remain unexamined. At this point, we shall defer further discussion of these elliptic equations to section 5 and leave the discussion of the interpretation of the effects of horizontal buoyancy gradients on the solution behaviour of the elliptic equation (3.47) for the quantity \( \eta \mathbf{u} \) for the TL1 model as an open problem.

Before substituting the asymptotic expansions and truncating the equations to achieve thermal geostrophic balance in terms of strict asymptotics in section 5, we will first derive the stochastic thermal L1 equations.

## 4 The Eliassen approximation of stochastic TRSW

The equation sets for the deterministic and stochastic TRSW models in section 2 and the deterministic TL1 model in the previous section have all been derived in the variational framework of the Euler–Poincaré theorem introduced in section 2.2. The corresponding Kelvin circulation laws for each of these theories follows from their Kelvin–Noether theorem 2.6, proved in section 2.3. Let us now derive the stochastic version of the TL1 equations and their corresponding Kelvin circulation law by following the SALT formulation in the Euler–Poincaré variational framework. To do so, we first investigate the balance relation in the presence of stochasticity.
4.1 Stochastic thermal geostrophic balance

To obtain the stochastic thermal geostrophic balance, we start from the TRSW equations with SALT, given in (2.30) by

\[ du + (d\chi_t \cdot \nabla)u + \sum_k (\nabla \xi_k) \cdot u \, dW^k_t = -\frac{1}{Fr^2} \nabla (b\zeta) \, dt + \frac{1}{2Fr^2} (\zeta - h) \nabla b \, dt \]
\[ - \frac{1}{Ro} f \times d\chi_t - \frac{1}{Ro} \sum_k (\xi_k \cdot R) \circ dW^k_t , \]

(4.1)

with boundary conditions in (2.31). We recall the assumptions that led to deterministic thermal geostrophic balance. We assume that the ratio between the dimensionless numbers in the problem is

\[ \frac{\epsilon}{Ro Fr^2} \sim O(1). \]

(4.2)

This condition is satisfied, for example, when \( O(\epsilon) = O(Ro) = O(Fr) \). This asymptotic regime allows us to continue with a single small parameter, \( \epsilon \). So, we formulate (3.3) as

\[ \zeta(x,t) = \epsilon \zeta_1(x,t), \]
\[ b(x,t) = 1 + \epsilon b_1(x,t), \]
\[ f(x) = 1 + \epsilon f_1(x), \]
\[ h(x) = 1 + \epsilon h_1(x), \]
\[ R(x) = R_0(x) + \epsilon R_1(x), \]

(4.3)

where the additional relations \( \hat{z} \cdot \nabla \times R_0 = 1 \) and \( \hat{z} \cdot \nabla \times R_1 = f_1 \) hold. Upon substituting the asymptotic expansions (4.3) into the stochastic TRSW equations (4.1) and collecting all terms of \( O(\epsilon^{-1}) \), we find

\[ \left( \nabla \zeta_1 + \frac{1}{2} \nabla b_1 \right) + \hat{z} \times u \right) dt + \sum_k \left( \hat{z} \times \xi_k + \nabla (\xi_k \cdot R_0) \right) \circ dW^k_t = 0 . \]

(4.4)

The drift part of this stochastic partial differential equation is the deterministic thermal geostrophic balance (3.6) and the diffusion part provides us with a relation between the noise amplitude \( \xi_k \) and the vector potential for the Coriolis parameter,

\[ u = \hat{z} \times \nabla \left( \zeta_1 + \frac{1}{2} b_1 \right) + O(\epsilon), \]
\[ \xi_k = \hat{z} \times \nabla (\xi_k \cdot R_0) + O(\epsilon). \]

(4.5)

Since the Brownian motions are assumed to be independent, (4.5) needs to be satisfied for each \( k \). We can identify the thermal geostrophic balance velocity field as \( u_{TG} = \hat{z} \times \nabla (\zeta_1 + \frac{1}{2} b_1) \) and expand the velocity field \( u \) as in the deterministic case

\[ u = u_{TG} + O(\epsilon), \]

(4.6)

and following the same reasoning, the \( \xi_k \) can be expanded as

\[ \xi_k = \xi_{TGk} + O(\epsilon). \]

(4.7)

We can now investigate the stochastic thermal L1 model.

4.2 Stochastic TL1

The equations governing the stochastic TL1 model are obtained in the SALT formulation by applying the stochastic Euler–Poincaré theorem 2.10 to the TL1 Lagrangian, given by (3.28). This incorporates the deter-
ministic geostrophic balance. We find the stochastic version of the TL1 equations (3.34), given by

\[
\begin{align*}
\text{d}u_{TG} & - \text{d}\chi_t \times \left( \nabla \times \left( u_{TG} + \frac{1}{\epsilon} R \right) \right) + \nabla \left( \frac{1}{\epsilon^2} \Delta \zeta + \frac{1}{2} |u_{TG}|^2 + \frac{1}{\epsilon} B \right) \text{d}t \\
& + \nabla \left( \sum_k \zeta_k \cdot \left( u_{TG} + \frac{1}{\epsilon} R \right) \right) \circ \text{d}W^k_t - \left( \frac{1}{2\epsilon^2} (\zeta - h) + \frac{1}{2\epsilon \eta} B \right) \nabla b \text{d}t = 0. 
\end{align*}
\]

(4.8)

Here, the function \( B \) is defined by

\[
B := \hat{z} \cdot \nabla \times \left( \eta (u - u_{TG}) \right). 
\]

(4.9)

The boundary conditions are given by

\[
\hat{n} \cdot u = 0 \quad \text{and} \quad \hat{n} \cdot \zeta_k = 0 \quad \text{and} \quad \hat{n} \times \nabla b = 0 \quad \text{on the boundary } \partial D. 
\]

(4.10)

The Kelvin–Noether theorem 2.6 for the stochastic TL1 model is given by the following theorem.

**Theorem 4.1** (Kelvin theorem for the stochastic TL1 model). The stochastic TL1 equations (4.8) imply the following Kelvin circulation law

\[
\begin{align*}
\text{d} \int_{c(\text{d}X_t)} \left( u_{TG} + \frac{1}{\epsilon} R \right) \cdot \text{d}x & = \int_{c(\text{d}X_t)} \left( \frac{1}{2\epsilon^2} (\zeta - h) + \frac{1}{2\epsilon \eta} B \right) \nabla b \cdot \text{d}x \text{ d}t \\
& = \int_{\partial S = c(\text{d}X_t)} \hat{z} \cdot \nabla \left( \frac{1}{2\epsilon^2} (\zeta - h) + \frac{1}{2\epsilon \eta} B \right) \times \nabla b \text{ d}x \text{ d}y \text{ d}t. 
\end{align*}
\]

(4.11)

**Proof.** The proof follows the pattern of the standard Kelvin–Noether theorem 2.6, modulo an application of the Kunita–Hö–Wentzell theorem which provides the chain rule for the Lie derivatives of differential forms by stochastic vector fields, as proved in [dLHLT20]. The loop \( c(\text{d}X_t) \) does not explicitly require the evaluation of a stochastic integral, as it is the push-forward of a stationary loop by the flow which is generated by the vector field \( \text{d}X_t \), see remark 2.7.

\[\square\]

The stochastic TL1 equations do not conserve energy due to their explicit dependence on time via the noise. From the Kelvin circulation theorem associated to the stochastic TL1 equations, an evolution equation for potential vorticity can be derived. This equation shows that potential vorticity is not conserved along Lagrangian particle trajectories, but is generated by the effect also present on the right hand side in the Kelvin circulation theorem 4.11. Recall that the potential vorticity \( q \) is defined by

\[
q = \frac{1}{\eta} \left( \omega + \frac{1}{\epsilon} f \right) = \frac{1}{\epsilon \eta} \left( \Delta \left( \frac{1}{2} h \right) + f \right). 
\]

(4.12)

The evolution equation for \( q \) is given by

\[
\text{d}q + \text{d}\chi_t \cdot \nabla q = \frac{1}{\eta} \hat{z} \cdot \nabla \left( \frac{1}{2\epsilon^2} (\zeta - h) + \frac{1}{2\epsilon \eta} B \right) \times \nabla b \text{ d}t. 
\]

(4.13)

Even though potential vorticity is not a Lagrangian invariant, the stochastic TL1 equations have an infinite family of integral conservation laws, given by

\[
C_{\Phi, \Psi} = \int_D \eta \Phi(b) + \eta q \Psi(b) \text{ d}x \text{ d}y. 
\]

(4.14)

The proof for these conservation laws is a direct calculation that uses the boundary conditions (4.10). In order to use (4.8) as a predictive model, one must be able to determine \( u \) from the other variables in the model. We can proceed as in the deterministic case and derive an elliptic equation for \( u \).
4.2.1 Determining the Lagrange multiplier

By operating with $\hat{z} \times$ on the momentum equation in (4.8) and using the definition of $u_{TG}$ (3.7), we obtain

$$
\frac{1}{\epsilon} \nabla \left( \frac{1}{2} b^2 \right) + \eta q \frac{d}{dt} \mathbf{X}_t = \hat{z} \times \nabla \left( \frac{1}{\epsilon^2} b^2 + \frac{1}{2} |u_{TG}|^2 + \frac{1}{\epsilon} B \right) \mathbf{X}_t + \nabla \left( \sum_k \xi_k \cdot \left( u_{TG} + \frac{1}{\epsilon} R \right) \right) \circ dW_t^k
$$

We continue by taking the stochastic differential through the gradient and substitute the continuity equation for the buoyancy from (4.8). By using the definition for $B$, we can then use the vector calculus identity $\Delta u = \nabla \times \nabla \cdot u + \nabla \cdot \nabla \cdot u$. This leads to two linear non-autonomous elliptic partial differential equations, one for the drift part and one for the diffusion part. The elliptic equation for the drift part is given by

$$
\frac{1}{\epsilon} \Delta (\eta u) + \frac{1}{2\epsilon} \nabla (u \cdot \nabla b) - \frac{1}{2\epsilon^2} (\nabla^\perp \cdot (\nabla^\perp \eta) \nabla^\perp b + \eta q u) = -\frac{1}{\epsilon^2} \nabla^\perp (b^2) + \frac{1}{2\epsilon^2} (\zeta - h) \nabla^\perp b
$$

and the elliptic equation for the diffusion part is given by

$$
\frac{1}{\epsilon} \nabla (\nabla \cdot \xi_k) + \frac{1}{2\epsilon} \nabla (\xi_k \cdot \nabla b) + \nabla^\perp \left( \xi_k \cdot \left( u_{TG} + \frac{1}{\epsilon} R \right) \right) + \eta q \xi_k = 0.
$$

We will further investigate these elliptic equations in the context of thermal quasi-geostrophy (TQG), which is obtained upon introducing the asymptotic expansions (4.3) and truncating at $O(1)$. We will visit the deterministic case first and then proceed to the stochastic case.

5 Thermal QG model

In sections 3 and 4, we made an approximation to the kinetic energy by assuming that the velocity field can be decomposed into a thermal geostrophic part and a higher order part. This led to the thermal L1 equations, given by (3.34) in the deterministic case, and by (4.8) in the stochastic case. These in turn can be approximated further to yield the motion equation for thermal quasi-geostrophy (TQG), which will be the subject of this section. We will start with the deterministic case first and then proceed to the stochastic case.

5.1 Deterministic TQG model

To obtain the thermal quasi-geostrophic model, one could expand the TL1 equations (3.34) and find that the continuity equation features the divergence of $u$, which can then be solved for by substitution. This approach has the disadvantage that equations are expanded before substitution, which loses accuracy. Instead we follow the derivation for the elliptic equation which determines the Lagrange multiplier $u$. By operating with $\hat{z} \times$ on the TL1 momentum equation in (3.34) and using the definition of $u_{TG}$, we obtain

$$
\frac{1}{\epsilon} \frac{\partial}{\partial t} \nabla \left( \frac{1}{2} b^2 \right) + \eta q \frac{d}{dt} \mathbf{X}_t = \hat{z} \times \left( \nabla \left( \frac{1}{\epsilon^2} b^2 + \frac{1}{2} |u_{TG}|^2 + \frac{1}{\epsilon} B \right) \right) - \frac{1}{\epsilon^2} (\zeta - h) \nabla^\perp b.
$$

We now take the divergence of (5.1) and substitute the continuity equation for the depth, which yields

$$
\frac{1}{\epsilon} \frac{\partial}{\partial t} \Delta (\eta + \frac{1}{2} b^2) - q \frac{\partial}{\partial t} \eta + \eta u \cdot \nabla q = J \left( \frac{1}{\epsilon^2} (\zeta - h) + \frac{1}{2\epsilon^2} B b \right),
$$

so the potential energy term $\eta q \frac{d}{dt}$ also appears in the vorticity equation. Equivalently, one can take the two dimensional curl, or the perpendicular divergence, to arrive directly at (5.2). In a moment, we will expand these equations in the asymptotic regime introduced in (3.3) and truncate at $O(1)$ to obtain the thermal quasi-geostrophic equations (TQG). In the asymptotic expansions it will be helpful to note that the potential vorticity, defined in (3.38), contains a term that is of order $O(\epsilon^{-1})$. 

21
Asymptotic expansion to the deterministic TQG regime. By means of the asymptotic expansions (3.3) which were used to derive an expression for \( \mu_{TQG} \), one can expand equation (5.2) and collect terms of the same order. Let us focus on the right hand side of (5.2) first. The Jacobian operator is a differential operator, which means that only derivatives of the buoyancy are present. Thus, in line with the asymptotic expansions in (3.3), we have \( \nabla b = \epsilon \nabla b_1 \). Hence, we may factor out an overall \( \epsilon^{-1} \). Then \( \zeta - h \) appears multiplying \( O(\epsilon^{-1}) \) and \( B \) appears multiplying \( O(1) \). However, since \( B \) contains \( u - \mu_{TQG} = O(\epsilon) \), it will actually contribute only at \( O(\epsilon) \).

Introducing the asymptotic expansions for \( \zeta, b, f \) and \( b \) and truncating at \( O(1) \) thus leads to the following motion equation for thermal quasi-geostrophy (TQG) [WD13, Zei18]

\[
\frac{\partial}{\partial t} \left( \Delta \left( \zeta_1 + \frac{1}{2} b_1 \right) - \zeta_1 \right) + \mu_{TQG} \cdot \nabla \left( \Delta \left( \zeta_1 + \frac{1}{2} b_1 \right) + f_1 \right) = \frac{1}{2} \hat{z} \times \nabla (\zeta_1 - b_1) \cdot \nabla b_1. \tag{5.3}
\]

Because the depth equation in (3.34) was substituted into equation (5.2), the potential energy term \( \partial \zeta_1 / \partial t \) remains in the vorticity equation (5.3). In the asymptotic expansion, the deterministic buoyancy equation keeps its form, as

\[
\frac{\partial}{\partial t} b_1 + \mu_{TQG} \cdot \nabla b_1 = 0, \tag{5.4}
\]

and the boundary conditions at this order become

\[
\hat{n} \cdot \mu_{TQG} = 0 \quad \text{and} \quad \hat{n} \times \nabla b_1 = 0 \quad \text{on the boundary } \partial D. \tag{5.5}
\]

Equations (5.3) and (5.4) together with the boundary conditions (5.5) form a closed model which approximates the TL1 model (3.34). This completes the present derivation of the thermal quasi-geostrophic (TQG) model, cf., [WD13, Zei18].

The vorticity equation (5.3) for TQG can be rewritten in terms of the stream function \( \psi := \zeta_1 + \frac{1}{2} b_1 \) and the Jacobian operator \( J(\psi, a) = \hat{z} \cdot \nabla \psi \times \nabla a = \mu_{TQG} \cdot \nabla a \) for a function \( a(\mathbf{x}) \) to obtain the more compact form,

\[
\frac{\partial}{\partial t} (\Delta \psi - \psi + b_1) + J(\psi, \Delta \psi + f_1) = -\frac{1}{2} J(h_1, b_1). \tag{5.6}
\]

Here, we have added and subtracted \( \frac{1}{2} b_1 \) in the time derivative in vorticity equation (5.3). Another \( \frac{1}{2} b_1 \) contribution comes from the forcing term on the right hand side, upon replacing the free surface elevation \( \zeta_1 \) by the stream function \( \psi \) and then using the buoyancy equation (5.4), written now as

\[
\frac{\partial}{\partial t} b_1 + J(\psi, b_1) = 0. \tag{5.7}
\]

The boundary conditions are

\[
\hat{n} \times \nabla \psi = 0 \quad \text{and} \quad \hat{n} \times \nabla b_1 = 0 \quad \text{on the boundary } \partial D. \tag{5.8}
\]

When the bathymetry is flat and the Coriolis parameter is constant, equation (5.6) reduces to the TQG equation found in [WD13, Zei18]. Since the Jacobian operator is zero when the arguments are functionally related, it is possible to write the deterministic TQG equation in terms of a type of potential vorticity variable, which we will call \( q \). This notation forms a close link between QG without buoyancy and TQG, with

\[
q := \Delta \psi - \psi + f_1. \tag{5.9}
\]

The definition of \( q \) in (5.9) allows us to formulate the vorticity equation in (5.6) as

\[
\frac{\partial}{\partial t} (q + b_1) + J(\psi, q) = -\frac{1}{2} J(h_1, b_1), \tag{5.10}
\]

The formulation in terms of \( q \) as in (5.10) is particularly useful in showing that these equations conserve energy, but also to note that the TQG equations can be related to Rayleigh-Bénard convection.

Remark 5.1. We can write (5.10) in such a way that all terms that depend on the buoyancy variations appear on the right hand side,

\[
\frac{\partial}{\partial t} q + J(\psi, q) = \frac{1}{2} J(\psi, b_1) + \frac{1}{2} J(\zeta_1 - h_1, b_1),
\]

\[
\frac{\partial}{\partial t} b_1 + J(\psi, b_1) = 0. \tag{5.11}
\]
This notation reveals that the equations (5.11) have a resemblance to ideal Rayleigh-Bénard convection. The Rayleigh-Bénard convection problem in a vertical $xz$–plane, formulated in terms of vorticity and stream function, is given by

$$\frac{\partial}{\partial t} q + J(\psi, q) = \alpha g T_z,$$
$$\frac{\partial}{\partial t} T + J(\psi, T) = 0. \tag{5.12}$$

Here $q = \Delta \psi$ is the vorticity, $T$ is the temperature, $\psi$ is the stream function, $g$ is gravity and $\alpha$ is the thermal expansion coefficient. For the typical Rayleigh-Bénard convection problem, in the vertical direction there are two solid boundaries and in the horizontal direction, one either uses periodic boundary conditions or solid boundaries. The bottom boundary is being heated and the top boundary is cooled, in such a way that the temperature difference is constant. Similar boundary conditions can be established for the thermal quasi-geostrophic equations. The main differences between the two models is that the forcing terms in TQG involve derivatives in every direction, whereas Rayleigh-Bénard convection only involves derivatives of the temperature in the vertical $z$-direction. On the other hand, TQG is obtained as a model based on thermal geostrophic balance, whereas the Rayleigh-Bénard model does not impose any balance.

Conservation laws for deterministic TQG. The deterministic TQG system (5.6) and (5.7) conserves the energy

$$E_{TQG}(q, b_1) = \frac{1}{2} \int_D (q - f_1)((\Delta - 1)^{-1}(q - f_1) + h_1 b_1) \, dx \, dy. \tag{5.13}$$

The proof that the TQG equations conserve energy is a direct calculation that requires the boundary conditions (5.8). The deterministic TQG equations also conserve an infinity of integral conservation laws, determined by two arbitrary differentiable functions of buoyancy $\Phi(b_1)$ and $\Psi(b_1)$ as

$$C_{\Phi, \Psi} = \int_D \Phi(b_1) + q \Psi(b_1) \, dx \, dy. \tag{5.14}$$

The proof for the family of integral conservation laws in (5.14) is a direct calculation that requires the boundary conditions (5.8). The integral conservation laws for the TQG equations has the same form as the family of integral conserved quantities for the TRSW equations, defined in (2.22), and likewise the integral conserved quantities (3.41) for the TL1 equations with the corresponding potential vorticity variable for TL1 in (3.38). As with the TRSW equations, introducing stochasticity via the Euler–Poincaré theorem preserves the conservation laws in (5.14). However, introducing stochasticity does not preserve energy, because the stochastic Lagrangian does depend explicitly on time through the Brownian motion.

Elliptic equation. Upon substituting the asymptotic expansions (3.3) and collecting the leading order terms, the elliptic equation for TL1 (3.47) reads

$$\frac{1}{\epsilon} \Delta u_{TG} + \frac{1}{\epsilon} u_{TG} = -\frac{1}{\epsilon} \nabla^\perp \zeta_1 + \frac{1}{2\epsilon} \nabla^\perp b_1 + \frac{1}{\epsilon} \nabla^\perp \nabla^\perp \cdot u_{TG} + O(1). \tag{5.15}$$

By means of the identity $\Delta u_{TG} = \nabla \nabla \cdot u_{TG} + \nabla^\perp \nabla^\perp \cdot u_{TG}$ and using the fact that $u_{TG}$ is divergence free, we obtain at leading order the definition for $u_{TG}$. At the next order, the elliptic equation will provide an expression for $u$ that is consistent with the asymptotic regime.

5.2 TQG with stochastic advection by Lie transport (SALT)

In a moment, we will obtain the SALT version of thermal quasi-geostrophy by following the reasoning in the deterministic case which brought us to (5.2).

Asymptotic expansion to the stochastic TQG regime. The stochastic TQG equations can be obtained from the same asymptotic expansion procedure for the stochastic TL1 equations in (4.8) as we took earlier to derive the deterministic TQG equations (5.3) and (5.4) together with the boundary conditions (5.5) from the asymptotic expansion of the deterministic TL1 equations in (3.34). This is possible, because none of the manipulations we performed on the TL1 equations (3.34) in that procedure involved taking time derivatives, which would not have been allowed in the presence of stochasticity. Thus, substituting the asymptotic expansions
in (3.3) into the stochastic TL1 equations of motion in (4.8) and truncating at $\mathcal{O}(1)$ yields the stochastic thermal quasi-geostrophic (TQG) equation of motion after following the same sequence of manipulations as in section 5.1. Upon defining the stochastic vector field

$$\mathbf{d} \mathbf{x}_t := \mathbf{u}_{\text{TQG}}(\mathbf{x}, t) \, dt + \sum_k \xi_k(\mathbf{x}) \circ dW_t^k,$$

where $\mathbf{u}_{\text{TQG}} = \mathbf{\hat{z}} \times \nabla (\zeta + \frac{1}{2} b_1)$, the stochastic TQG motion equation is obtained as

$$d \left( \nabla \left( \zeta + \frac{1}{2} b_1 \right) - \zeta \right) + d \mathbf{x}_t \cdot \nabla \left( \nabla \left( \zeta + \frac{1}{2} b_1 \right) + f_1 \right) = \frac{1}{2} \mathbf{\hat{z}} \times \nabla (\zeta - h_1) \cdot \nabla b_1 \, dt. \quad (5.17)$$

As before in obtaining (5.3), the depth equation in (3.34) has been used to write equation (5.17) in this form, which has introduced the potential energy term $-\Delta \zeta$. The stochastic buoyancy equation obtained from applying the asymptotic expansion assumptions in (3.3) is given by

$$db_t + d \mathbf{x}_t \cdot \nabla b_1 = 0, \quad (5.18)$$

and the corresponding boundary conditions become,

$$\mathbf{n} \cdot \mathbf{u}_{\text{TQG}} = 0 \quad \text{and} \quad \mathbf{n} \cdot \xi_k = 0 \quad \text{and} \quad \mathbf{n} \times \nabla b_1 = 0 \quad \text{on the boundary } \partial \mathcal{D}. \quad (5.19)$$

Since $\nabla \cdot \mathbf{u}_{\text{TQG}} = 0$, we also require that $\nabla \cdot \xi_k = 0$. The divergence–free condition on the $\xi_k$ appears naturally when we expand the elliptic equation (4.17) for stochastic TL1 in the asymptotic regime (4.3). The TQG equations (5.17) and (5.18) together with the boundary conditions in (5.19) form a closed model that approximates the stochastic TL1 model (4.8).

The deterministic TQG model, given by equations (5.3)-(5.5) and the stochastic TQG model, given by equations (5.17)-(5.19) will be the main subjects of discussion in the next section. The stochastic vorticity equation (5.17) for TQG can be rewritten using scalar potentials for the divergence free thermal geostrophic velocity field $\mathbf{u}_{\text{TQG}}$ and the divergence free spatial vector fields $\xi_k$,

$$d \mathbf{x}_t := \mathbf{u}_{\text{TQG}}(\mathbf{x}, t) \, dt + \sum_k \xi_k(\mathbf{x}) \circ dW_t^k = \mathbf{\hat{z}} \times \nabla \left( \psi dt + \sum_k \vartheta_k(\mathbf{x}) \circ dW_t^k \right), \quad (5.20)$$

where $\psi = \zeta + \frac{1}{2} b_1$ and $\vartheta_k = \xi_k \cdot \mathbf{R}_0$. Now we can use the Jacobian operator $J$ to obtain a more compact form of (5.17) given by

$$d \left( \Delta \psi + \psi + \frac{1}{2} b_1 \right) + J \left( \psi dt + \sum_k \vartheta_k \circ dW_t, \Delta \psi + f_1 \right) = \frac{1}{2} J (\zeta - h_1, b_1) \, dt. \quad (5.21)$$

Here, $f_1(\mathbf{x})$ and $h_1(\mathbf{x})$ are, respectively, the variations at order $\mathcal{O}(\epsilon)$ of the Coriolis parameter and the bottom topography (or, bathymetry). Both $f_1(\mathbf{x})$ and $h_1(\mathbf{x})$ are prescribed functions of space, and they are constant in time. The stochastic advection equation for the buoyancy $b_1(\mathbf{x}, t)$ represents scalar tracer transport; namely,

$$db_t + J \left( \psi dt + \sum_k \vartheta_k \circ dW_t, b_1 \right) = 0. \quad (5.22)$$

The boundary conditions in this formulation are

$$\mathbf{n} \times \nabla \psi = 0 \quad \text{and} \quad \mathbf{n} \times \nabla \vartheta_k = 0 \quad \text{and} \quad \mathbf{n} \times \nabla b_1 = 0 \quad \text{on the boundary } \partial \mathcal{D}. \quad (5.23)$$

The presence of stochasticity introduces explicit time dependence and implies that these equations do not conserve energy. Also, the stochastic vorticity equation (5.21) cannot be formulated completely in terms of the potential vorticity variable $q$, defined in (5.9). Such a formulation was possible in the deterministic case and was given by (5.10). However, the following partial formulation in terms of $q$ is useful in the calculation showing that these equations do still have an uncountable infinity of integral conserved quantities,

$$d \left( \psi dt + \sum_k \vartheta_k \circ dW_t, q + \psi - \frac{1}{2} b_1 \right) = \frac{1}{2} J (\psi - h_1, b_1) \, dt. \quad (5.24)$$
Because of the stochastic term in (5.24), the family of integral conserved quantities for stochastic TQG are not the same as in the deterministic case. This is because the equations (5.21) cannot be formulated only in terms of \( q := \Delta \psi - \psi + f_1 \). The family of integral conserved quantities which persists for stochastic TQG is given by

\[
C_\Phi = \int_D \Phi(b_1) \, dx \, dy.
\]  

(5.25)

The proof is a direct calculation that involves the boundary conditions (5.23). If the functions \( \vartheta_k \) are either constant, a function of the stream function or a function of buoyancy, then we recover the family of conservation laws which we saw in the deterministic case (5.14).

### Elliptic equation.

In the stochastic case, the elliptic equation for the Lagrange multiplier \( u \) splits up into a drift part (4.16) and a diffusion part (4.17). The drift part implies the definition for \( u_{TG} \) after expanding with (4.3), as we saw in the deterministic situation. The diffusion part, after expanding with (4.3), at leading order reads

\[
\frac{1}{\epsilon} \nabla \nabla \cdot \xi_{TG_k} + \frac{1}{\epsilon} \nabla^\perp (\xi_{TG_k} \cdot R_0) + \frac{1}{\epsilon} \xi_{TG_k} + O(1) = 0.
\]  

(5.26)

From the thermal geostrophic balance in stochastic TRSW, we have in (4.5) a definition for \( \xi_{TG_k} \) which is divergence free. The first term in (5.26) drops out and we find the defining relation for \( \xi_{TG_k} \). At the next order, the elliptic equation will provide an expression for \( \xi_k \) that is consistent with the asymptotic regime.

### 5.3 Numerical TQG example

We implemented the TQG equations (5.11) using finite element methods (FEM) for the spatial variables. The FEM algorithm we used is an adaptation of the algorithm formulated in [BBvdV06], and was implemented using Firedrake\(^2\), see [RHM+17]. In particular, we approximate the vorticity and buoyancy fields in first order discrete Galerkin finite element space, and approximate the stream function in first order continuous Galerkin finite element space. For the time variable, we used an optimal third order strong stability preserving Runge-Kutta method, see [Got05, CCH+19].

Figure 3 shows a snapshot at a certain time taken from a high resolution numerical run of the TQG equations. In this numerical example, we used the following boundary and initial conditions. The domain is \([0, 2\pi]^2\) discretised at a resolution of 512\(^2\). The boundary conditions are periodic in the vertical direction and walls in the horizontal direction. The parameters and initial conditions are

\[
\begin{align*}
    b_1(x, y) &= \cos x + \frac{1}{2} \cos 2x + \frac{1}{3} \cos 3x, \\
    f_1(x, y) &= 0, \quad \text{\((f\text{-plane})\)} \\
    b_1(x, y, 0) &= -\frac{1}{1 + \exp(-x + \pi)} + \frac{1}{2}, \\
    q(x, y, 0) &= -\exp(-5(y - \pi)^2), \\
    \zeta_1(x, y, 0) &= \psi(x, y, 0) - \frac{1}{2} b_1(x, y, 0).
\end{align*}
\]  

(5.27)

The stream function \( \psi = \zeta_1 + b_1/2 \) is calculated from the potential vorticity \( q \) by means of the elliptic problem given in (5.9).

\(^2\)http://www.firedrakeproject.org/index.html
Figure 3: These images from an evolutionary computational simulation of deterministic TQG illustrate the result of the creation of a buoyancy front from an initial state and its subsequent emergent instabilities which produce a cascade of horizontal circulations to smaller scales. In the top left, the buoyancy profile is shown and in the bottom left, the sea surface height is shown. In the top right, the stream function is shown and in the bottom right, the vorticity is shown. The domain is periodic at the upper and lower boundaries, while the stream functions are constant and equal on the lateral boundaries. The bathymetry varies in the lateral direction as a sum of cosines with wavenumbers $k = 1, 2,$ and $3$. The initial condition had a Gaussian-profile strip of vorticity along the lateral mid-line and the buoyancy began with a $k = 1$ sine profile in the lateral direction. The figure shows that a buoyancy front had developed and then generated a cascade of smaller mushroom-like dipole circulations and trains of Kelvin-Helmholtz roll-ups via interaction between the buoyancy and bathymetry gradients, followed by shear instability. Compare this configuration with the chlorophyll tracers shown in Figure 1.
6 Conclusion and outlook

Our aim in this paper has been to prepare the mathematical framework for our impending investigation of Stochastic Transport in Upper Ocean Dynamics (STUOD) by using the stochastic data assimilation algorithms developed and applied previously to determine the eigenvectors $\xi_i(x)$ in the cases of the stochastic Euler fluid equation and the 2-layer stochastic QG model in [CCH+18, CCH+19]. This framework has been established by deriving a sequence of realistic 2D models of Upper Ocean Dynamics with buoyancy effects by using nested asymptotic expansions with a shared stochastic variational structure. The process of developing these sequential derivations has revealed several open mathematical problems at each level of approximation for these new nonlinear stochastic partial differential equations, as listed below.

(i) An extensive computational simulation study will be needed for classifying the solution behaviour of these new stochastic TRSW and TQG equations. This computational study has been left as a future step, after having established its efficacy in section 5.3.

(ii) These computational simulations will be required in the calibration of the eigenvectors $\xi_k(x)$ for the noise in equation (5.20) and their subsequent use in the new framework for data calibration, uncertainty quantification and data assimilation using particle filters following the SALT algorithm developed in [CCH+18, CCH+19]. This future simulation study will prepare these models for applications in the analysis of observed oceanic cyclogenesis as seen in Figure 1.

(iii) Investigation of the numerous potential effects of the horizontal buoyancy gradients appearing in the elliptic equation for the thermal L1 Lagrange multiplier $u$ (3.47) has been left as an open mathematical problem for further analysis and computational simulation.

(iv) The issue of well-posedness of these new nonlinear stochastic partial differential equations (2.34) for TRSW and (5.24) for TQG has also been left for future mathematical investigation.

(v) Finally, we recall that our derivation of the stochastic barotropic TQG balanced model in (5.24) has neglected the potentially important effects of baroclinic instabilities which tend to re-stratify the fluid. In particular, a future study with baroclinic TQG would extend the QG analysis of baroclinic instability of the currents around the Lofoten Basin given in [Isa15] to include thermal effects. For an in-depth discussion of baroclinic effects in comparison to balanced models, see [CFFFK16].

Acknowledgments

We are grateful for constructive suggestions offered in discussions with C. J. Cotter, B. Chapron, D. Crisan, S.R. Ephrati, B. Fox-Kemper, O. Lang, J.M. Leahy, J. C. McWilliams, E. Mémin, O. Street and S. Takao. We are also grateful to the European Space Agency for access to the Sentinel 3B satellite data shown in Figure 1. During this work, DDH and WP were partially supported by ERC Synergy Grant 856408 - STUOD (Stochastic Transport in Upper Ocean Dynamics). EL was supported by EPSRC grant [grant number EP/L016613/1]. EL is also grateful for the warm hospitality shown to him at the Imperial College London EPSRC Centre for Doctoral Training in the Mathematics of Planet Earth mpecdt.org.

References

[AH96] John S Allen and Darryl D Holm. Extended-geostrophic Hamiltonian models for rotating shallow water motion. *Physica D: Nonlinear Phenomena*, 98(2-4):229–248, 1996.

[AM78] Ralph Abraham and Jerrold E Marsden. *Foundations of mechanics*, volume 36. Benjamin/Cummings Publishing Company Reading, Massachusetts, 1978.

[BBK+18] Igor Bashmachnikov, Tatyana Belonenko, Pavel Kuibin, Denis Volkov, and Victor Foux. Pattern of vertical velocity in the Lofoten vortex (the Norwegian Sea). *Ocean Dynamics*, 68(12):1711–1725, 2018.

[BBvdV06] Erik Bernsen, Onno Bokhove, and Jaap J. W. van der Vegt. A (dis)continuous finite element model for generalized 2D vorticity dynamics. *Journal of Computational Physics*, 211(2):719–747, 2006.
[BCC+18] Jochen Bröcker, Ben Calderhead, Davoud Cheraghi, Colin Cotter, Dan Crisan, Darryl Holm, Tobias Kuna, Beatrice Pelloni, Ted Shepherd, and Hilary Weller. *Mathematics of Planet Earth: A Primer*. World Scientific, 2018.

[BKP+20] Igor L. Bashmachnikov, Igor E Kozlov, Larisa A Petenko, Nataliia I Glock, and Claudia Wekerle. Eddies in the North Greenland Sea and Fram Strait from satellite altimetry, SAR and high-resolution model data. *Journal of Geophysical Research: Oceans*, page e2019JC015832, 2020.

[BSB+17] IL Bashmachnikov, MA Sokolovskiy, TV Belonenko, DL Volkov, PE Isachsen, and X Carton. On the vertical structure and stability of the Lofoten vortex in the Norwegian Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 128:1–27, 2017.

[CCH+18] Colin Cotter, Dan Crisan, Darryl D Holm, Wei Pan, and Igor Shevchenko. Modelling uncertainty using circulation-preserving stochastic transport noise in a 2-layer quasi-geostrophic model. *arXiv preprint arXiv:1802.05711*, 2018.

[CCH+19] Colin Cotter, Dan Crisan, Darryl D Holm, Wei Pan, and Igor Shevchenko. Numerically modeling stochastic Lie transport in fluid dynamics. *Multiscale Modeling & Simulation*, 17(1):192–232, 2019.

[CFFFK16] Jörn Callies, Glenn Flierl, Raffaele Ferrari, and Baylor Fox-Kemper. The role of mixed-layer instabilities in submesoscale turbulence. *Journal of Fluid Mechanics*, 788:5–41, 2016.

[CHLN20] Dan Crisan, Darryl D Holm, James-Michael Leahy, and Torstein Nilssen. Variational principles for fluid dynamics on rough paths. *arXiv preprint arXiv:2004.07829*, 2020.

[CHP10] Colin J Cotter, Darryl D Holm, and James R Percival. The square root depth wave equations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2124):3621–3633, 2010.

[dLHLT20] Aythami Bethencourt de Leon, Darryl D Holm, Erwin Luesink, and So Takao. Implications of Kunita–Itô–Wentzell formula for $k$-forms in stochastic fluid dynamics. *Journal of Nonlinear Science*, pages 1–34, 2020.

[Eli49] Arnt Eliassen. *The quasi-static equations of motion with pressure as independent variable*. Grøndahl & sons boktr., I kommisjon hos Cammermeyers boghandel, 1949.

[FB20] Aleksandr M Fedorov and Tatyana V Belonenko. Interaction of mesoscale vortices in the Lofoten basin based on the GLORYS database. *Russian Journal of Earth Sciences*, 20, 2020.

[FSL18] Boris N Filyushkin, Mikhail A Sokolovskiy, and Konstantin V Lebedev. Evolution of an intrathermocline lens over the Lofoten basin. In *The Ocean in Motion*, pages 333–347. Springer, 2018.

[Got05] Sigal Gottlieb. On high order strong stability preserving Runge–Kutta and multi step time discretizations. *Journal of Scientific Computing*, 25(1):105–128, 2005.

[HL19] Darryl D Holm and Erwin Luesink. Stochastic wave-current interaction in stratified shallow water dynamics. *arXiv preprint arXiv:1910.10627*, 2019.

[HMR98] Darryl D Holm, Jerrold E Marsden, and Tudor S Ratiu. The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81, 1998.

[HMRW85] Darryl D Holm, Jerrold E Marsden, Tudor Ratiu, and Alan Weinstein. Nonlinear stability of fluid and plasma equilibria. *Physics reports*, 123(1-2):1–116, 1985.

[Hol15] Darryl D Holm. Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2176):20140963, 2015.

[Hol20] Darryl D Holm. Variational formulation of stochastic wave-current interaction (swci). *arXiv preprint arXiv:2002.04291*, 2020.

[HSS09] Darryl D Holm, Tanya Schmah, and Cristina Stoica. *Geometric mechanics and symmetry: from finite to infinite dimensions*, volume 12. Oxford University Press, 2009.

[Isa15] Pål Erik Isachsen. Baroclinic instability and the mesoscale eddy field around the Lofoten basin. *Journal of Geophysical Research: Oceans*, 120(4):2884–2903, 2015.

[Ito84] K. Ito. *An Introduction to Probability Theory*. Cambridge University Press, 1984.

[McW07] James C McWilliams. Irreducible imprecision in atmospheric and oceanic simulations. *Proceedings of the National Academy of Sciences*, 104(21):8709–8713, 2007.
[McW19] James C McWilliams. A survey of submesoscale currents. *Geoscience Letters*, 6(1):1–15, 2019.

[MNC+09] Nikolai Maximenko, Peter Niiler, Luca Centurioni, Marie-Helene Rio, Oleg Melnichenko, Don Chambers, Victor Zlotnicki, and Boris Galperin. Mean dynamic topography of the ocean derived from satellite and drifting buoy data using three different techniques. *Journal of Atmospheric and Oceanic Technology*, 26(9):1910–1919, 2009.

[RHM+17] Florian Rathgeber, David A. Ham, Lawrence Mitchell, Michael Lange, Fabio Luporini, Andrew T. T. Mcrae, Gheorghe-Teodor Bercea, Graham R. Markall, and Paul H. J. Kelly. Firedrake: Automating the Finite Element Method by Composing Abstractions. *ACM Transactions on Mathematical Software*, 43(3):1–27, January 2017.

[Rip95] P Ripa. Low frequency approximation of a vertically averaged ocean model with thermodynamics. *Revista Mexicana de Física*, 42(1):117–135, 1995.

[Sal83] Rick Salmon. Practical use of Hamilton’s principle. *Journal of Fluid Mechanics*, 132:431–444, 1983.

[VKL15] Denis L Volkov, Arseny A Kubryakov, and Rick Lumpkin. Formation and variability of the Lofoten basin vortex in a high-resolution ocean model. *Deep Sea Research Part I: Oceanographic Research Papers*, 105:142–157, 2015.

[WD13] Emma S Warneford and Paul J Dellar. The quasi-geostrophic theory of the thermal shallow water equations. *Journal of Fluid Mechanics*, 723:374–403, 2013.

[Zei18] Vladimir Zeitlin. *Geophysical fluid dynamics: understanding (almost) everything with rotating shallow water models*. Oxford University Press, 2018.