Multi-Scale Methods for Geophysical Flows

Christian L. E. Franzke, Marcel Oliver, Jens D. M. Rademacher and Gualtiero Badin

Abstract Geophysical flows comprise a broad range of spatial and temporal scales, from the planetary to meso-scale and microscopic turbulence regimes. The relation of scales and flow phenomena is essential in order to validate and improve current numerical weather and climate prediction models. While regime separation is often possible on a formal level via multi-scale analysis, the systematic exploration, structure preservation, and mathematical details remain challenging. This chapter provides an entry to the literature and reviews fundamental notions as background for the later chapters in this collection and as a departure point for original research in the field.

1 Introduction

The climate system varies on a multitude of temporal and spatial scales which interact nonlinearly with each other. It is common that phenomena with small spatial scales also vary fast while phenomena with large spatial scales vary more slowly. For instance, small scale turbulent eddies have spatial scales from millimeters up to a meter and exist between a few seconds to a few minutes. In the atmosphere, meso-

Christian L. E. Franzke
Meteorological Institute and Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany, e-mail: christian.franzke@uni-hamburg.de

Marcel Oliver
Mathematical Sciences, Jacobs University, Bremen, Germany e-mail: oliver@member.ams.org

Jens D. M. Rademacher
Department of Mathematics, University of Bremen, Bremen, Germany e-mail: jdmr@uni-bremen.de

Gualtiero Badin
Institute of Oceanography and Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany, e-mail: gualtiero.badin@uni-hamburg.de

scale phenomena like convection and gravity waves have length scales between a few 100 meters and about 20 km and time scales between hours and days. Synoptic weather systems have lifetimes of a few days and spatial scales of up to 2000 km while planetary-scale teleconnection patterns can extent throughout the hemisphere and have time scales from a week to decades.

At the most fundamental level, scale separation in geophysical flows is described by the concept of *balance*. When the Rossby number (the ratio between internal forces and Coriolis forces) or the Froude number (the ratio of inertial flow velocity to gravity wave speed) is small, there exists a slow or *balanced* component which evolves nonlinearly and interacts only weakly with the high frequency components. The fast motions can often be approximately characterized as the high frequency linear waves in the linearized equations of motion. We remark that the term "waves" is sometimes used in a lose sense for describing spatio-temporal oscillations diagnosed by Fourier wave modes, in particular in experimental studies.

A precise characterization of balance is a perennial theme in geophysical fluid dynamics; a recent review can be found in McIntyre (2015). Balance can be described in two ways: kinematically via *balance relations* which define almost-invariant objects through a phase space constraint, and dynamically as *balance models* which are closed sets of equations representing the slow dynamics on a balanced manifold, thus approximating the flow of the full system under balanced initial conditions. However, the notion of balance is intrinsically approximate: Emergence of imbalance from balanced initial conditions is generic, though often exponentially small, cf. Vanneste (2013). Temam and Wirosoetisno (2010) argue that even in the viscous case there is no exact invariant balance manifold. A recent discussion can be found in Whitehead and Wingate (2014).

Balance in the equatorial band of latitudes is a more subtle concept, as the Coriolis parameter changes sign at the equator causing a singularity in straightforward small-Rossby number expansions. There are suitable equatorial scalings and balance assumptions to formally circumvent the problem, usually at the price of losing some (linear) waves (McIntyre, 2015; Verkley and van der Velde, 2010; Theiss and Mohebalhojeh, 2009). Recent work by Chan and Shepherd (2013) shows that it is possible to capture equatorial Rossby and Kelvin waves in a full hierarchy of equatorial balance models.

While the different scale regimes can have different physical mechanisms driving them, they are all described by the same general set of equations of motion. *Multi-scale asymptotics* can be used as a systematic means of deriving new sets of equations which only describe the flows at certain temporal and spatial scales, and their interactions with other regimes; see, for example, Majda and Klein (2003) or Klein (2010). These new sets of equations then describe the underlying dynamics of the respective flow regimes and, thus, provide a better understanding of the flow dynamics and the dominant balance conditions.

Recent studies elucidated the interactions between the planetary and synoptic scale regimes (Dolaptchiev and Klein, 2013) and between the planetary and mesoscale regimes (Shaw and Shepherd, 2009). The latter study also discusses how these multi-scale models can be used to systematically derive energy consistent subgrid-scale parameterizations.

An unavoidable ingredient in comprehensive models of climate and global atmosphere and ocean dynamics is the parameterizations of unresolved degrees of freedom. The numerical aspects in the context of geostrophic turbulence are discussed in **[insert reference to M3 chapter in this volume]**. Of particular interest to us is the use of stochastic modeling techniques for developing such parameterizations, in particular the connection of stochastic subgrid parameterizations with traditional deterministic multi-scale asymptotics. Theoretical work on the elimination of fast scales is often based on averaging techniques, which naturally apply to stochastic modeling as well. Mathematical introductions to this direction of model reduction are given by Givon et al. (2004); Pavliotis and Stuart (2008) while Franzke et al. (2015) and Gottwald et al. (2017) provide introductions to stochastic climate modeling. The parametrization problem is also an issue of data collection and assimilation, which we do not discuss here.

Part of the motivation for pursuing stochastic modeling approaches derives from the observation that most numerical weather predictions are under-dispersive, i.e. the observed weather lies too often outside of the forecast ensemble. To increase the ensemble spread, stochastic parameterizations have been introduced. The European Centre for Medium Range Weather Forecasts (ECMWF) is leading the introduction of stochastic parameterizations into numerical weather and seasonal climate prediction models. They currently use two different schemes operationally, the Stochastically Perturbed Parameterization Tendencies Scheme (SPPT) and the Stochastic Kinetic-Energy Backscatter Scheme (SKEBS), described in Palmer et al. (2009).

SPPT is based on the notion that, especially with increasing numerical resolution, the equilibrium assumption no longer holds and the subgrid-scale state should be sampled rather than represented by the equilibrium mean. Consequently, SPPT multiplies the accumulated physical tendencies at each grid-point and time step with a random pattern that has spatial and temporal correlations. SKEBS aims at representing model uncertainty arising from unresolved and unrepresented subgrid-scale processes by introducing random perturbations to the streamfunction. SKEBS is based on the rationale that a small fraction of the numerically dissipated energy will be re-injected into the resolved scales due to nonlinear triad interactions and acts effectively as a systematic forcing. Current deterministic climate models, however, neglect this energy pathway. The use of stochastic parameterizations in climate models is less well established due to the required re-tuning of all deterministic parameterizations. The current use of stochastic parameterizations is rather *ad hoc* (Franzke et al., 2015) and does not consider energy and momentum consistency. For instance, SPPT has no rigorous mathematical justification; it is based on trial and error and empirical evidence that it increases the spread of ensemble forecasts. However, this does not necessarily make the forecasts more precise. SKEBS has some mathematical justification. Theory suggests that most of the "subgrid-scale" energy should be re-injected close to the truncation scale, i.e., at scales where numerical dissipation acts. However, in practice, the re-injected energy gets damped very quickly with no improvement in forecast skill. The usual remedy is to re-inject the energy into all scales at a rate empirically determined by trying to match a known energy spectrum. While there is growing evidence showing the benefits of stochastic parameterizations, there is also a need for more rigorous mathematical approaches, particularly in the light of the observation that errors introduced by inconsistent treatment of the interactions between resolved and unresolved scales can be very significant (Shaw and Shepherd, 2009; Burkhardt and Becker, 2006; Becker, 2003).

This chapter, first, intends to give an overview on current progress and open issues in multi-scale modeling of geophysical flows. We aim at some broadness, but our focus is clearly biased by our own research interests in geometric and structure preserving methods, dynamical systems techniques, and stochastic modeling.

Second, and no less important, a large part of the chapter is devoted to laying down foundational concepts, starting with the equations of motion, nondimensionalization, the introduction of the classical limits in geophysical fluid dynamics, variational and Hamiltonian methods, dissipation, and stochastic model reduction. While much of this is standard textbook material, covered in great detail, for example, in the classical books by Gill (1982), Kamenkovich et al. (1986), and Pedlosky (1987), or in the more recent books by Salmon (1998), Vallis (2006), and Olbers et al. (2012), our intent is to introduce the foundations most relevant to the questions raised here in concise and unified notation. In our presentation, we take the rotating Boussinesq equations as the single parent model from which all other models arise, aim at a clear statement of simplifying assumptions, and give some consideration to the full Coriolis term and to equatorial scalings.

The remainder of this chapter is structured as follows. Section 2 introduces the rotating Boussinesq equations as the governing equations of geophysical flow, discusses imbalance variables, nondimensionalization, and several simplified models derived in different scaling limits: the hydrostatic approximation in form of the primitive equations, the quasi-geostrophic equations in several forms, and the shallow water equations. Section 3 discusses the variational principle which gives rise to the equations of motion as well as Poisson and Nambu formulations of the dynamics. Section 4 provides a high-level overview of the role of dissipation and its relation to turbulence and to dynamical systems methods for studying linear and nonlinear waves. Section 5 introduces systematic approaches for the stochastic modeling of fast motions. The chapter closes with a brief outline of questions we are addressing in our current research.

2 The governing equations

2.1 Rotating Boussinesq equations

As the starting point for this exposition, we consider the rotating Boussinesq equations in the tangent plane approximation. These equations simplify the full equations for global atmospheric or oceanic dynamics in two fundamental ways.

First, the Boussinesq approximation is based on the observation that density fluctuations are so small that their impact on inertia is negligible while their impact on buoyancy remains significant. This assumption is typically very well satisfied in the ocean, so that most ocean general circulation models assume the Boussinesq approximation. For atmospheric flows, the Boussinesq approximation is more tenuous, but under the so-called anelastic approximation, the assumption of constant background pressure, and the use of vertical pressure coordinates, the equations governing large-scale atmospheric flows can also be written in the same form. One main consequence of the Boussinesq approximation is the absence of acoustic waves.

Second, the tangent plane approximation is based on the observation that the local metric is nearly Cartesian. Thus, we work in Cartesian coordinates on a tangent plane while still allowing variations in the Coriolis parameter with latitude so that equatorial scalings can be explored. A tangent plane model is clearly not suitable for building global models, but is more convenient for studying scale interaction from small to synoptic scale motions.

We shall also restrict ourselves to the simplest possible equation of state where changes in density depend linearly on changes in temperature and neglect source terms. With these provisions, our governing equations read

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

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

$$D_t \rho = \kappa \Delta \rho \,, \tag{1c}$$

where $D_t = \partial_t + \boldsymbol{u} \cdot \boldsymbol{\nabla}$ denotes the material derivative, $\boldsymbol{u}$ is the three-dimensional fluid velocity field,

$$\boldsymbol{\Omega} = |\boldsymbol{\Omega}| \begin{pmatrix} 0 \\ \cos\vartheta \\ \sin\vartheta \end{pmatrix} \tag{2}$$

the angular velocity vector describing the rotation of the earth at latitude ϑ, ρ_0 a constant reference density, p the departure from hydrostatic pressure, g the constant of gravity, ρ the departure from the constant reference density, $\boldsymbol{k}$ the unit vector in z (vertical) direction, ν the coefficient of viscosity, and κ is proportional to the coefficient of thermal diffusion.

A derivation and a discussion of the underlying assumptions can be found in any classical textbook on geophysical fluid dynamics, e.g. Gill (1982), Pedlosky (1987), or Vallis (2006).

The dissipative terms on the right hand side describe, at this point, molecular viscosity and diffusion processes. In typical large-scale circulation problems, they act at scales far beyond what can be resolved in any practical numerical simulation. Thus, they will be replaced, explicitly or implicitly via a stable numerical scheme by eddy diffusion and/or numerical diffusion terms. Such terms may not be harmonic or isotropic, an issue to which we will return in Section 4.1 below.

We shall consider (1) on a domain Ω with a free upper boundary at $z = h(\boldsymbol{x}_{\mathrm{h}}, t)$ and rigid bottom at $z = -b(\boldsymbol{x}_{\mathrm{h}})$. We shall not discuss lateral boundary conditions here. For highly idealized studies, for example on the phenomenology of geostrophic turbulence, periodic lateral boundary conditions are frequently used. At the bottom boundary, we impose the impermeability condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$, which may be written

$$\boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} b = w \,, \tag{3a}$$

plus viscous boundary conditions or bottom drag parameterizations if applicable. Here and below, $\boldsymbol{n}$ denotes the outward unit normal vector and the subscript "h" refers to the horizontal components. The free surface is described by the condition $\sigma(\boldsymbol{x}, t) \equiv h(\boldsymbol{x}_{\mathrm{h}}, t) - z = 0$ and is subject to the kinematic boundary condition $D_t \sigma = 0$, i.e.,

$$\partial_t h + \boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} h = w \,. \tag{3b}$$

As a second condition at the free surface, we have the dynamic boundary condition

$$p = p_{\mathrm{s}} \,, \tag{3c}$$

i.e., pressure equals a specified external pressure p_{s} at the surface. In the viscous case, the free surface boundary conditions are augmented by (wind) stress conditions. Integrating the incompressibility constraint (1b) in z and using the two kinematic boundary conditions (3a) and (3b), we obtain the *free surface equation*

$$\partial_t h + \boldsymbol{\nabla}_{\mathrm{h}} \cdot \int_{-b}^{h} \boldsymbol{u}_{\mathrm{h}} \, \mathrm{d}z = 0 \,. \tag{4}$$

Propagation in time then uses the horizontal momentum equations, the free surface equation, and the advection of buoyancy (1c) as prognostic quantities while the vertical velocity w is reconstructed from the incompressibility constraint; see, e.g., Klingbeil and Burchard (2013).

The free surface equation can be approximated in various ways. Replacing the upper limit of integration in (4) by $z = 0$ while still keeping the time derivative $\partial_t h$ gives a linearization of the free surface equation which can be formulated in the spirit of Chorin's projection method where the free surface update is obtained by solving an elliptic equation. This removes time-step restrictions due to fast surface waves; surface waves are damped when the time step becomes too large.

A more drastic approximation is a *rigid lid* upper boundary, where all equations are posed on a fixed domain bounded above at $z = 0$. Consequently, the dynamic boundary condition (3c) must be dropped. The time evolution of all fields can be computed by an incompressible solver. An approximation of the elevation of the free surface may then be diagnostically obtained by solving the equivalent hydrostatic pressure relation

$$p_{\mathrm{s}} = g \rho_0 h \tag{5}$$

for h. The rigid lid approximation removes all surface gravity waves, which is often a reasonable approximation in the ocean.

We finally remark that the inviscid Boussinesq equations conserve the *total energy*

$$H = \int_\Omega \frac{1}{2} |\boldsymbol{u}|^2 + \frac{g}{\rho_0} \rho z \, \mathrm{d}\boldsymbol{x} \tag{6}$$

and materially conserve *potential vorticity*

$$q = (2\boldsymbol{\Omega} + \boldsymbol{\omega}) \cdot \boldsymbol{\nabla} \rho \,. \tag{7}$$

In other words, q satisfies the advection equation

$$D_t q = 0 \,. \tag{8}$$

Both conservation laws can be verified by direct computation. However, in Section 3.1 we will show that they emerge elegantly from symmetries in the underlying variational principle.

2.2 Imbalance variables

Large-scale geophysical flow, at least in the sub-equatorial regime, is to a substantial part determined by potential vorticity alone (McIntyre and Norton, 2000). It is therefore natural to use potential vorticity as one of the prognostic variables and to augment the set of prognostic variables by so-called *imbalance variables* suitably chosen such that they remain small so long as the flow is nearly balanced. In the context of the f-plane shallow water equations, these additional variables are divergence and ageostrophic vorticity; they underlie the discussion of spontaneous emission of gravity waves by Ford et al. (2000), also see the discussion in **[Badin et al., L2 chapter, this volume]**. Mohebalhojeh and Dritschel (2001) report numerical advantages when simulating shallow water using this set of variables, and Dritschel et al. (2016) find that the differences between several variational balance models can only be understood when looking at balance relations formulated in terms of such imbalance variables. Imbalance variables for the primitive equations are introduced, for example, by McIntyre and Norton (2000).

For the f-plane Boussinesq system in the traditional approximation, there are different possible choices. Dritschel and Viúdez (2003) use the horizontal components of

$$\boldsymbol{A} = \boldsymbol{\nabla} \times \boldsymbol{u} - \frac{g}{\rho_0 f} \boldsymbol{\nabla} \rho \,, \tag{9}$$

which leads to a nonlinear inversion problem for recovering the vector potential of $\boldsymbol{A}$. Alternatively, it is also possible to use ageostrophic vorticity and divergence (Vanneste, 2013). However, as noted above, the traditional and the hydrostatic approximations are not independent. Thus, we shall derive (for purposes of exposition working at the linear level only) a version of the nontraditional f-plane Boussinesq system in imbalance variables.

To begin, we note that linear inertia-gravity waves require rotation and strong stratification, which can be expressed by assuming a decomposition of the density of the form (33) with constant background profile, so that

$$D_t \rho - \frac{N^2 \rho_0}{g} \, w = 0 \tag{10}$$

and therefore, by incompressibility,

$$\partial_z D_t \rho + \frac{N^2 \rho_0}{g} \, \boldsymbol{\nabla}_{\mathrm{h}} \cdot \boldsymbol{u}_{\mathrm{h}} = 0 \, . \tag{11}$$

Further, taking the full divergence of the inviscid Boussinesq momentum equation (1a), invoking incompressibility, differentiating in time, and using (11), we find

$$\boldsymbol{\nabla} \cdot (2\boldsymbol{\Omega} \times \partial_t \boldsymbol{u}) + \frac{1}{\rho_0} \, \partial_t \Delta p - N^2 \, \boldsymbol{\nabla}_{\mathrm{h}} \cdot \boldsymbol{u}_{\mathrm{h}} = \mathrm{NL} \, , \tag{12}$$

where NL is used to denote any number of nonlinear terms, possibly different from one line to the next. Next, taking the horizontal divergence of (1a) yields

$$\partial_t \boldsymbol{\nabla}_{\mathrm{h}} \cdot \boldsymbol{u}_{\mathrm{h}} + \boldsymbol{\nabla}_{\mathrm{h}} \cdot (2\boldsymbol{\Omega} \times \boldsymbol{u})_{\mathrm{h}} + \frac{1}{\rho_0} \, \Delta_{\mathrm{h}} p = \mathrm{NL} \, . \tag{13}$$

This motivates taking the *horizontal divergence*

$$\delta = \boldsymbol{\nabla}_{\mathrm{h}} \cdot \boldsymbol{u}_h \tag{14}$$

and the *acceleration divergence* or *ageostrophic vorticity*

$$\gamma = -\boldsymbol{\nabla}_{\mathrm{h}} \cdot (2\boldsymbol{\Omega} \times \boldsymbol{u})_{\mathrm{h}} - \frac{1}{\rho_0} \, \Delta_{\mathrm{h}} p \tag{15}$$

as imbalance variables, so that $\partial_t \delta = \gamma + \mathrm{NL}$. Then, using (13) to eliminate $\partial_t \Delta p$, we compute

$$\begin{aligned} \partial_t \Delta \gamma &= -\Delta \boldsymbol{\nabla}_{\mathrm{h}} \cdot (2\boldsymbol{\Omega} \times \partial_t \boldsymbol{u})_{\mathrm{h}} + \Delta_{\mathrm{h}} \boldsymbol{\nabla} \cdot (2\boldsymbol{\Omega} \times \partial_t \boldsymbol{u}) - N^2 \, \Delta_{\mathrm{h}} \delta + \mathrm{NL} \\ &= -\partial_{zz} \boldsymbol{\nabla}_{\mathrm{h}} \cdot (2\boldsymbol{\Omega} \times \partial_t \boldsymbol{u})_{\mathrm{h}} + \Delta_{\mathrm{h}} (2\boldsymbol{\Omega}_{\mathrm{h}}^{\perp} \cdot \partial_{tz} \boldsymbol{u}_{\mathrm{h}}) - N^2 \, \Delta_{\mathrm{h}} \delta + \mathrm{NL} \, . \end{aligned} \tag{16}$$

Writing $(2\boldsymbol{\Omega} \times \boldsymbol{u})_{\mathrm{h}} = f \, \boldsymbol{u}_{\mathrm{h}}^{\perp} - w \, 2\boldsymbol{\Omega}_{\mathrm{h}}^{\perp}$ in the first term on the right, substituting the horizontal momentum equation for all instances of $\partial_t \boldsymbol{u}_{\mathrm{h}}$, invoking incompressibility to replace all instances of $\partial_z w$ by $-\delta$, eliminating $\Delta_{\mathrm{h}} p$ via the definition of $\gamma = \partial_t \delta + \mathrm{NL}$, and collecting terms, we find

$$\partial_t \Delta\gamma = -f^2\,\partial_{zz}\delta - f\,\partial_z(2\mathbf{\Omega}_\mathrm{h}\cdot\nabla_\mathrm{h}\delta) + (2\mathbf{\Omega}_\mathrm{h}^\perp\cdot\nabla_\mathrm{h})^2\delta \\ + f\,\partial_z(2\mathbf{\Omega}_\mathrm{h}\cdot\nabla_\mathrm{h}^\perp\nabla_\mathrm{h}^\perp\cdot u_\mathrm{h} - 2\mathbf{\Omega}_\mathrm{h}\cdot\Delta_\mathrm{h}\boldsymbol{u}_\mathrm{h}) - (N^2 + |2\mathbf{\Omega}_\mathrm{h}|^2)\,\Delta_\mathrm{h}\delta + \mathrm{NL}\,. \quad (17)$$

Finally, applying the vector identity $\nabla_\mathrm{h}\nabla_\mathrm{h} + \nabla_\mathrm{h}^\perp\nabla_\mathrm{h}^\perp = I\,\Delta_\mathrm{h}$ twice and collecting terms, we can write the system in the form of a nonlinearly perturbed wave equation, where

$$\partial_t\delta - \gamma = \mathrm{NL}\,, \quad (18a)$$

$$\partial_t\Delta\gamma + (2\mathbf{\Omega}\cdot\nabla)^2\delta + N^2\,\Delta_\mathrm{h}\delta = \mathrm{NL}\,. \quad (18b)$$

Without the nonlinear terms, this system is equivalent to the equation considered by Gerkema and Shrira (2005a) to study linear waves on the nontraditional f-plane; for related work on the nontraditional β-plane, see Gerkema and Shrira (2005b), and Kasahara and Gary (2010) and Stewart and Dellar (2012). The shallow water equations with full Coriolis parameter on the sphere are considered by Tort et al. (2014) and linear stability of nonlinear jet-type solutions is investigated by Tort et al. (2016). We finally note that when the traditional approximation is made, (18) directly reduces to the system studied, e.g., by Vanneste (2013).

When the nonlinear terms in (18) are retained, these equations together with the advection equation for potential vorticity (7) form a closed system for the evolution of the f-plane Boussinesq equations because $(\boldsymbol{u}, p, \rho)$ can be recovered from (q, δ, γ) by inverting nonlinear elliptic equations. These involve boundary conditions, and for vertically bounded domains with $w = 0$ at $z = 0, H$ require solving $D_t\rho = 0$ at the top and bottom boundaries. The two dimensional fields $b(\boldsymbol{x}_\mathrm{h}, 0, t)$ and $b(\boldsymbol{x}_h, H, t)$ are therefore additional degrees of freedom. Further complications arise for horizontally bounded domains.

2.3 Mid-latitude scalings

Let us now consider possible scaling regimes for (1) on a tangent plane at mid-latitude ϑ_0. In this regime, the flow is expected to be horizontally isotropic, but vertical scales will generally differ. We thus split (1a) into the horizontal and vertical component equations and non-dimensionalize by introducing typical horizontal velocity U, typical vertical velocity W, typical Coriolis parameter $f_0 = 2|\mathbf{\Omega}|\,\cos\vartheta_0$, typical horizontal length scale L, typical vertical length scale H, typical time scale T, typical pressure scale P, and typical density scale Γ, inserting $\boldsymbol{u} = U\hat{\boldsymbol{u}}$, $w = W\hat{w}$, $2\mathbf{\Omega} = f_0\hat{\mathbf{\Omega}}$, etc., into the Boussinesq equations (1), and finally dropping the hats, we obtain

$$\frac{U}{T}\,\partial_t \boldsymbol{u}_\mathrm{h} + \frac{U^2}{L}\,\boldsymbol{u}_\mathrm{h}\cdot\boldsymbol{\nabla}_\mathrm{h}\boldsymbol{u}_\mathrm{h} + \frac{UW}{H}\,w\,\partial_z\boldsymbol{u}_\mathrm{h} + f_0 U\Omega_z\,\boldsymbol{u}_\mathrm{h}^\perp - f_0 W\boldsymbol{\Omega}_\mathrm{h}^\perp\,w = -\frac{P}{L\rho_0}\,\boldsymbol{\nabla}_\mathrm{h}p\,, \tag{19a}$$

$$\frac{W}{T}\,\partial_t w + \frac{UW}{L}\,\boldsymbol{u}_\mathrm{h}\cdot\boldsymbol{\nabla}_\mathrm{h}w + \frac{W^2}{H}\,w\,\partial_z w + f_0 U\boldsymbol{\Omega}_\mathrm{h}^\perp\cdot\boldsymbol{u}_\mathrm{h} = -\frac{P}{H\rho_0}\,\partial_z p - \frac{g\Gamma}{\rho_0}\,\rho\,, \tag{19b}$$

$$\frac{U}{L}\,\boldsymbol{\nabla}_\mathrm{h}\cdot\boldsymbol{u}_\mathrm{h} + \frac{W}{H}\,\partial_z w = 0\,, \tag{19c}$$

$$\frac{\Gamma}{T}\,\partial_t\rho + \frac{U\Gamma}{L}\boldsymbol{u}_\mathrm{h}\cdot\boldsymbol{\nabla}_\mathrm{h}\rho + \frac{WN^2\rho_0}{g}\,w\,\partial_z\rho = 0\,. \tag{19d}$$

Here and below, N denotes the typical Brunt–Väisälä or buoyancy frequency defined as

$$N^2 = -\frac{g}{\rho_0}\left[\frac{\partial\rho}{\partial z}\right], \tag{20}$$

where we write $[\partial\rho/\partial z]$ to denote the typical vertical density gradient.

We look at the problem on a horizontally advective time scale, i.e., on a time scale in which fluid parcels travel a horizontal distance of order one. This fixes

$$\frac{1}{T} \sim \frac{U}{L}\,. \tag{21}$$

Our goal is now to estimate the vertical velocity scale W. Introducing the *aspect ratio*

$$\alpha = \frac{H}{L}\,, \tag{22}$$

we may obtain a first simple estimate, $W \lesssim \alpha U$, directly from the incompressibility condition. However, the vertical velocity gradient does not necessarily participate in the dominant balance in the incompressibility relation. Indeed, we shall see that rotation as well as stratification may provide sharper scaling bounds on W.

Turning to the "thermodynamic equation" (19d), we obtain the scaling bound

$$W \sim \lambda\,\frac{g\Gamma}{\rho_0}\,\frac{U}{N^2L} = \lambda\,\frac{g\Gamma}{\rho_0}\,\frac{L}{U}\,\alpha^2\,\mathrm{Fr}^2 \tag{23}$$

for some $\lambda \lesssim 1$ with *Froude number*

$$\mathrm{Fr} = \frac{U}{NH}\,. \tag{24}$$

Assuming that the dominant balance in the horizontal momentum equation (19a) is between horizontal Coriolis force and horizontal pressure gradient and that the dominant balance in the vertical momentum equation (19b) is between buoyancy and vertical pressure gradient, we have

$$f_0 LU\rho_0 \sim P \sim g\Gamma H\,. \tag{25}$$

Introducing the *Rossby number*

$$\mathrm{Ro} = \frac{U}{f_0 L} \tag{26}$$

and writing $f \equiv \Omega_z$ for the nondimensionalized rotation rate, we can rewrite the momentum equations as

$$\mathrm{Ro}\,(\partial_t \boldsymbol{u}_\mathrm{h} + \boldsymbol{u}_\mathrm{h} \cdot \boldsymbol{\nabla}_\mathrm{h} \boldsymbol{u}_\mathrm{h}) + \lambda\,\mathrm{Fr}^2\, w\,\partial_z \boldsymbol{u}_\mathrm{h} + f\,\boldsymbol{u}_\mathrm{h}^\perp - \lambda\alpha\,\frac{\mathrm{Fr}^2}{\mathrm{Ro}}\,\boldsymbol{\Omega}_\mathrm{h}^\perp\, w = -\boldsymbol{\nabla}_\mathrm{h} p\,, \tag{27a}$$

$$\lambda\,\alpha^2\,\mathrm{Fr}^2\,(\partial_t w + \boldsymbol{u}_\mathrm{h} \cdot \boldsymbol{\nabla}_\mathrm{h} w) + \lambda^2\,\alpha^2\,\frac{\mathrm{Fr}^4}{\mathrm{Ro}}\, w\,\partial_z w - \alpha\,\boldsymbol{\Omega}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h} = -\partial_z p - \rho\,. \tag{27b}$$

Note that, so far, we have not assumed smallness of any parameter beyond fixing the dominant balance in the momentum equations. Let us now, in addition, assume that $\lambda \sim 1$, i.e., that the estimate on W implied by the thermodynamic equation is sharp, and further assume that we are on a 3D advective time scale, i.e., that

$$\mathrm{Ro} = \mathrm{Fr}^2\,. \tag{28}$$

Then

$$\mathrm{Ro}\,(\partial_t \boldsymbol{u}_\mathrm{h} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u}_\mathrm{h}) + f\,\boldsymbol{u}_\mathrm{h}^\perp - \alpha\,\boldsymbol{\Omega}_\mathrm{h}^\perp\, w = -\boldsymbol{\nabla}_\mathrm{h} p\,, \tag{29a}$$

$$\alpha^2\,\mathrm{Ro}\,(\partial_t w + \boldsymbol{u} \cdot \boldsymbol{\nabla} w) - \alpha\,\boldsymbol{\Omega}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h} = -\partial_z p - \rho\,. \tag{29b}$$

2.4 Hydrostatic approximation

Then hydrostatic balance is based on the scaling

$$\mathrm{Ro} \gg \alpha^2\,\mathrm{Fr}^2 \tag{30}$$

or $\alpha \ll 1$. Thus, in 3D advective scaling, hydrostatic balance is purely a small aspect ratio assumption, and if it is made, then the vertical contributions to the Coriolis force are small as well. We altogether obtain the hydrostatic *primitive equations*

$$\mathrm{Ro}\,(\partial_t \boldsymbol{u}_\mathrm{h} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u}_\mathrm{h}) + f\,\boldsymbol{u}_\mathrm{h}^\perp = -\boldsymbol{\nabla}_\mathrm{h} p\,, \tag{31a}$$

$$\partial_z p = -\rho\,, \tag{31b}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0\,, \tag{31c}$$

$$D_t \rho = 0\,. \tag{31d}$$

In this context, we remark that it has long been known that the hydrostatic approximation requires the "traditional approximation" where the vertical contributions to the Coriolis force are neglected to ensure that the Coriolis force is energy-neutral; see, for example, the discussion in White (2002) and Klein (2010).

While there have been early studies of the global circulation using balance models as discussed further below, contemporary OGCMs are based on the primitive equations. For the study of small scale phenomena, the use of non-hydrostatic models is advancing (Fringer, 2009). In the atmosphere, where non-hydrostatic effects are more pronounced, non-hydrostatic models are routinely used (Saito et al., 2007).

2.5 The quasi-geostrophic approximation on the β-plane

We now turn to the quasi-geostrophic approximation on a mid-latitude β-plane. It is a more severe approximation than hydrostaticity. We modify the assumptions made in Section 2.3 in three respects. First, we look at the scaling

$$\mathrm{Fr} \sim \mathrm{Ro}\,, \tag{32}$$

thereby breaking the 3D-advective scaling. Advection in the vertical is thus coming in at higher order than advection in the horizontal direction. Second, we assume that pressure variations are small relative to a static vertical stratification profile. We write

$$\rho(\boldsymbol{x}, t) = \bar{\rho}(z) + \rho'(\boldsymbol{x}, t)\,; \tag{33}$$

it is then convenient to introduce the perturbation pressure ψ satisfying

$$\partial_z \psi = \partial_z p + \bar{\rho}\,. \tag{34}$$

Third, we make the β-plane approximation with the assumption that the change of the Coriolis parameter in y is $O(\mathrm{Ro})$. In nondimensionalized variables, this assumption reads

$$f = 1 + \mathrm{Ro}\,\beta y\,, \tag{35}$$

where $\mathrm{Ro}\,\beta$ is the meridional gradient of the Coriolis parameter. Inserting (33) into the inviscid thermodynamic equation (1c) and dropping the prime, we nondimensionalize as follows:

$$\frac{U\Gamma}{L}\,(\partial_t \rho + \boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} \rho) + \frac{W\Gamma}{H}\, w\, \partial_z \rho + \frac{W N^2 \rho_0}{g}\, w\, \partial_z \bar{\rho} = 0\,. \tag{36}$$

Keeping the last term in the dominant balance leads, once again, to a scaling relation of the form (23) with $\lambda = 1$, so that, in non-dimensional variables,

$$\partial_t \rho + \boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} \rho + \mathrm{Ro}\, w\, \partial_z \rho + w\, \partial_z \bar{\rho} = 0\,. \tag{37}$$

Similarly, the momentum equations in non-dimensional variables, under the same dominant balance assumptions as in Section 2.3 above, read

$$\mathrm{Ro}\,(\partial_t \boldsymbol{u}_{\mathrm{h}} + \boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} \boldsymbol{u}_{\mathrm{h}} + \mathrm{Ro}\, w\, \partial_z \boldsymbol{u}_{\mathrm{h}}) + f\, \boldsymbol{u}_{\mathrm{h}}^{\perp} - \alpha\, \mathrm{Ro}\, \boldsymbol{\Omega}_{\mathrm{h}}^{\perp}\, w = -\boldsymbol{\nabla}_{\mathrm{h}} \psi\,, \tag{38a}$$

$$\alpha^2\, \mathrm{Ro}^2\, (\partial_t w + \boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} w + \mathrm{Ro}\, w\, \partial_z w) + \alpha\, \boldsymbol{\Omega}_{\mathrm{h}}^{\perp} \cdot \boldsymbol{u}_{\mathrm{h}} = -\partial_z \psi - \rho\,, \tag{38b}$$

$$\boldsymbol{\nabla}_{\mathrm{h}} \cdot \boldsymbol{u}_{\mathrm{h}} + \mathrm{Ro}\, \partial_z w = 0\,. \tag{38c}$$

The quasi-geostrophic equations are obtained by the leading order of this system in the formal limit $\mathrm{Ro} \to 0$ and $\alpha \to 0$. The leading order vertical momentum equation is, once again, the hydrostatic balance relation

$$\partial_z \psi = -\rho \,. \tag{39}$$

To determine the leading order balance in the horizontal momentum equation (38a), we need to separate the divergence from the curl component. For the divergence, the dominant contribution is geostrophic balance, i.e.

$$\boldsymbol{u}_\mathrm{h} = \boldsymbol{\nabla}_\mathrm{h}^\perp \psi \,. \tag{40}$$

(Note that there is no f in this relation as the deviation from constant Coriolis parameter comes in at $O(\mathrm{Ro})$ only.) The leading order of the curl of (38a) could be obtained by direct computation, but it is easier to work from the expression for the Boussinesq potential vorticity which, in dimensional variables, is given by (7). In non-dimensional variables, the potential vorticity reads

$$\begin{aligned} q &= \Big(f_0\, \boldsymbol{\Omega}_\mathrm{h} + \frac{U}{H}\, \partial_z \boldsymbol{u}_\mathrm{h}^\perp - \frac{W}{L}\, \boldsymbol{\nabla}_\mathrm{h}^\perp w\Big) \cdot \frac{\Gamma}{L}\, \boldsymbol{\nabla}_\mathrm{h} \rho \\ &\quad + \Big(f_0\, f + \frac{U}{L}\, \boldsymbol{\nabla}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h}\Big)\Big(\frac{\Gamma}{H}\, \partial_z \rho + \frac{N^2 \rho_0}{g}\, \partial_z \bar{\rho}\Big) \\ &= \frac{U\Gamma}{HL}\, \bigg(\Big(\frac{\alpha}{\mathrm{Ro}}\, \boldsymbol{\Omega}_\mathrm{h} + \partial_z \boldsymbol{u}_\mathrm{h}^\perp - \alpha^2\, \mathrm{Ro}\, \boldsymbol{\nabla}_\mathrm{h}^\perp w\Big) \cdot \boldsymbol{\nabla}_\mathrm{h} \rho \\ &\quad + \Big(\frac{1}{\mathrm{Ro}} + \beta y + \boldsymbol{\nabla}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h}\Big)\Big(\partial_z \rho + \frac{1}{\mathrm{Ro}}\, \partial_z \bar{\rho}\Big)\bigg) \end{aligned} \tag{41}$$

and satisfies the advection equation

$$(\partial_t + \boldsymbol{u}_\mathrm{h} \cdot \boldsymbol{\nabla}_\mathrm{h} + \mathrm{Ro}\, w\, \partial_z) q = 0 \,. \tag{42}$$

Since $\bar{\rho}$ does not depend on time or on the horizontal position, the leading order terms appear at $O(\mathrm{Ro}^{-1})$ and read

$$D_t^\mathrm{h} (\partial_z \rho + \boldsymbol{\nabla}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h}\, \partial_z \bar{\rho} + \beta y\, \partial_z \bar{\rho}) + w\, \partial_{zz} \bar{\rho} = 0 \,, \tag{43}$$

where $D_t^\mathrm{h} = \partial_t + \boldsymbol{u}_\mathrm{h} \cdot \boldsymbol{\nabla}_\mathrm{h}$ denotes the horizontal material derivative. Eliminating w with the leading order terms of (37), dividing through by $\partial_z \bar{\rho}$ and simplifying, we obtain the quasi-geostrophic potential vorticity equation,

$$D_t^\mathrm{h} \Big(\partial_z \frac{\rho}{\partial_z \bar{\rho}} + \boldsymbol{\nabla}_\mathrm{h}^\perp \cdot \boldsymbol{u}_\mathrm{h} + \beta y\Big) = 0 \,. \tag{44}$$

Using hydrostatic balance, the leading order of (38b) and the curl of geostrophic balance (40), the quasi-geostrophic equations can be written in the well known closed form

$$D_t^{\mathrm{h}} \Big(\Delta_{\mathrm{h}} \psi - \partial_z \frac{\partial_z \psi}{\partial_z \bar{\rho}} + \beta y \Big) = 0 \,, \tag{45a}$$

$$\boldsymbol{u}_{\mathrm{h}} = \boldsymbol{\nabla}_{\mathrm{h}}^{\perp} \psi \,. \tag{45b}$$

The advected quantity in (45a) is the quasi-geostrophic potential vorticity. To recover the stream function ψ from the potential vorticity, we need to solve a second order equation. It is elliptic provided that $\partial_z \bar{\rho} < 0$, i.e., the fluid is stably stratified. Moreover, we need boundary conditions. At the lateral boundaries, the no-flux condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$ constrains only the tangential derivative of ψ. A consistent lateral boundary condition is $\psi = 0$ on a simply-connected domain; for the multiply-connected case, see McWilliams (1977). In more idealized situations, channel geometries or, on the f-plane, periodic boundary conditions are frequently used.

At the top and bottom boundaries, we use Neumann boundary conditions. Due to hydrostatic balance, $\partial_z \psi = -\rho$, where ρ satisfies the leading order of (37),

$$D_t^{\mathrm{h}} \rho + w \, \partial_z \bar{\rho} = 0 \,. \tag{46}$$

At the top boundary, it is common to assume rigid lid conditions, i.e., $w = 0$. For a correction to a rigid lid condition, see, e.g., Olbers et al. (2012). At the bottom boundary, the impermeability condition reads $\boldsymbol{u}_{\mathrm{h}} \cdot \boldsymbol{\nabla}_{\mathrm{h}} b + w = 0$, with $b(\boldsymbol{x}_{\mathrm{h}})$ denoting the equilibrium depth.

To re-dimensionalize the quasi-geostrophic equations, it is convenient to express ψ in units of a horizontal stream function, not in pressure units. Moreover, we introduce the z-dependent Brunt–Väisälä frequency

$$N^2(z) = -\frac{g}{\rho_0} \, \frac{\mathrm{d}\rho}{\mathrm{d}z} \,. \tag{47}$$

Then we can write (45) in terms of the *quasi-geostrophic potential vorticity* q as

$$\partial_t q + \boldsymbol{\nabla}_{\mathrm{h}}^{\perp} \psi \cdot \boldsymbol{\nabla} q = 0 \,, \tag{48a}$$

$$q = f + \Delta_{\mathrm{h}} \psi + f_0^2 \, \partial_z \frac{\partial_z \psi}{N^2(z)} \,. \tag{48b}$$

Linearized about the trivial solution, these equations allow propagation of internal waves in the horizontal plane. Their phase velocities c_n are related to the eigenvalues $\lambda_n = c_n^{-2}$ of the vertical structure operator, the last term in (48b). Of particular importance is the speed c_1 of the fastest wave. The associated horizontal length scale is called the *(first) internal* or *baroclinic Rossby radius of deformation*,

$$L_{\mathrm{d}} = \frac{c_1}{f_0} = \frac{NH}{\pi f_0} \tag{49}$$

where the second equality holds in the case of uniform stratification. It is the scale at which buoyancy forces and Coriolis forces are equally important; it is also approximately the scale of strongest conversion of potential into kinetic energy via baroclinic instability; see, e.g., Vallis (2006) for details. In the ocean, L_d varies from less that 10 km at high latitudes up to 200 km in the tropics.

As the quasi-geostrophic equations are typically used for proof-of-concept studies, two simplifications are particularly useful. First, if the equations are posed in flat horizontal layer and density and velocity are independent of z, we obtain the two-dimensional or barotropic quasi-geostrophic equation

$$\partial_t q + \nabla^\perp \psi \cdot \nabla q = 0\,, \tag{50a}$$

$$q = f + \Delta\psi\,, \tag{50b}$$

where we drop the subscript "h" on the operators as all fields are fully two-dimensional. This form of the barotropic quasi-geostrophic equation implicitly carries a rigid lid assumption. It is possible to derive the equation without this assumption, in which case the expression for potential vorticity acquires an extra term; see (65) and the surrounding discussion.

Second, we can derive a simple model for a baroclinic rotating flow by assuming that the fluid moves in two uniform layers with constant depths H_1 and H_2, respectively, with layer 1 assumed on top of layer 2. We suppose that the potential vorticities q_i and stream functions ψ_i, where $i \in \{1, 2\}$, are taken at the layer centers. The density, on the other hand, is taken at the layer boundaries. Using finite differences, $\partial_z \psi \approx 2(\psi_1 - \psi_2)/(H_1 + H_2)$ at the layer interface. At the top and bottom interfaces, boundary condition (46) with $w = 0$ applies and the hydrostatic relation implies that $\partial_z \psi$ is advected. Altogether, using finite differences across each layer for the outer z-derivative in (48b), we obtain the quasi-geostrophic two-layer equations

$$(\partial_t + \nabla^\perp \psi_1 \cdot \nabla)\Big(f + \Delta\psi_1 + \frac{2 f_0^2}{N^2 H_1 (H_1 + H_2)} (\psi_2 - \psi_1)\Big)\,, \tag{51a}$$

$$(\partial_t + \nabla^\perp \psi_2 \cdot \nabla)\Big(f + \Delta\psi_2 + \frac{2 f_0^2}{N^2 H_2 (H_1 + H_2)} (\psi_1 - \psi_2)\Big)\,. \tag{51b}$$

When the two layers have equal depth $H_1 = H_2 = H/2$, we can write the two-layer quasi-geostrophic system in the symmetric form

$$\partial_t q_i + \nabla^\perp \psi_i \cdot \nabla q_i = 0\,, \tag{52a}$$

$$q_i = f + \Delta\psi_i + (-1)^i\, k_d^2\, (\psi_1 - \psi_2)/2\,, \tag{52b}$$

with $k_d = L_d^{-1}$, where the (single) internal Rossby radius is now given by

$$L_d = \frac{NH}{\sqrt{8} f_0}\,. \tag{53}$$

A detailed exposition of more general multi-level and multi-layer models can be found in Pedlosky (1987) or Kamenkovich et al. (1986). For a variational perspective on quasi-geostrophic theory, see Holm and Zeitlin (1998) and Bokhove et al. (1998).

2.6 Rotating shallow water equations

We shall finally introduce the rotating shallow water equations which describe the horizontal motion in a thin layer of an incompressible fluid with a free surface and constant density. The shallow water equations are often used as a simple test bed in situations where baroclinic effects are negligible or to be excluded.

When the density is constant, the Boussinesq equations (1) reduce to the three-dimensional Euler equations with rotation and gravitational forces acting in the vertical. Following the notation used above, they read

$$D_t \boldsymbol{u} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\frac{1}{\rho_0} \boldsymbol{\nabla} p - g\boldsymbol{k}\,, \tag{54a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0\,. \tag{54b}$$

We now make the following scaling assumptions. First, as before, we impose an advective time scale (21). Second, we assume that the aspect ratio α, defined in (22), is small. Without stratification, we only have the divergence condition (54b) to constrain the vertical velocity, so that $W \sim \alpha U$. Third, we suppose that

$$\mathrm{Ro} \lesssim \mathrm{Bu} = \frac{gH}{f_0^2\, L^2}\,, \tag{55}$$

where the dimensionless parameter Bu is known as the Burger number, which can also be interpreted as the square ratio of the Rossby radius of deformation to the horizontal length scale L. We remark that condition (55) is consistent with both the semi-geostrophic and the quasi-geostrophic scaling regime of shallow water theory, see the discussion in Section 2.7 below.

Under these assumptions, the formal limit $\alpha \to 0$ imposes, again simultaneously, the hydrostatic and the traditional approximation, so that, using a rescaled pressure but otherwise dimensional variables,

$$D_t \boldsymbol{u}_\mathrm{h} + f\, \boldsymbol{u}_\mathrm{h}^\perp = -\boldsymbol{\nabla}_\mathrm{h} p\,, \tag{56a}$$

$$0 = \partial_z p - g\,, \tag{56b}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0\,, \tag{56c}$$

The hydrostatic equation (56b) implies that the pressure is fully determined by the hydrostatic pressure. To be definite, let us suppose that $z = 0$ describes the equilibrium free surface, $h(\boldsymbol{x}_\mathrm{h}, t)$ the departure of the free surface from equilibrium, and $b(\boldsymbol{x}_\mathrm{h})$ the equilibrium depth of the layer. If the pressure at the free surface is zero, integration of (56b) yields

$$p(\boldsymbol{x}_\mathrm{h}, t) = g\,(h(\boldsymbol{x}_\mathrm{h}, t) - z)\,. \tag{57}$$

In particular, the horizontal pressure gradient is entirely independent of z. Thus, *assuming* that the velocity field is initially independent of z, it will remain so for all times, and the material advection operator D_t in (56a) reduces to horizontal material advection D_t^{h}. Finally, we can eliminate w from the divergence condition (56c) by vertical integration under zero normal flow boundary conditions, yielding a continuity equation for the total layer depth. Dropping the h-subscript, as all quantities are now fully two-dimensional, the resulting equations, known as the *rotating shallow water equations*, read

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + f\, \boldsymbol{u}^{\perp} + g\, \boldsymbol{\nabla} h = 0\,, \tag{58a}$$

$$\partial_t h + \boldsymbol{\nabla} \cdot ((h + b)\boldsymbol{u}) = 0\,. \tag{58b}$$

These equations can also be derived from a variational principle, see, e.g., Salmon (1998). The generalization to stratified rotating shallow water equations is based on assuming that the fluid consists of layers with constant density that are separated by interfaces. As for the above single layer rotating shallow water equations, one assumes that the amplitude of the deformation of each interface to be much less than the layer depths. For a variational derivation of the multi-layer rotating shallow water equations, see, e.g., Salmon (1982) and Stewart and Dellar (2010).

The rotating shallow water equations conserve the energy

$$H = \frac{1}{2} \int_{\Omega} h\, |\boldsymbol{u}|^2 + g\, h^2 \, \mathrm{d}\boldsymbol{x}\,, \tag{59}$$

and materially conserve the potential vorticity

$$q = \frac{f + \zeta}{h + b} = \frac{\zeta_{\mathrm{a}}}{h + b} \tag{60}$$

where

$$\zeta = \boldsymbol{\nabla}^{\perp} \cdot \boldsymbol{u} \tag{61}$$

is the *relative vorticity* and $\zeta_{\mathrm{a}} = f + \zeta$ the *absolute vorticity*. Both conservation laws can be shown to arise as Noetherian conservation laws as outlined in Section 3.1.

2.7 Geostrophic scalings

The Boussinesq equations with a free surface upper boundary also support surface gravity waves. Surface waves are retained in the shallow water approximation and textbook linear wave theory shows that they their maximal phase velocity is bounded by $c_{\mathrm{e}} = \sqrt{gH}$. This speed defines, as does (49) for the internal modes, the *external* or *barotropic Rossby radius of deformation*

$$L_{\mathrm{e}} = \frac{c_{\mathrm{e}}}{f_0} = \frac{\sqrt{gH}}{f_0} \tag{62}$$

as the horizontal scale at which free surface effects and rotation are of equal importance. In the deep ocean, for example, $L_{\mathrm{e}} \approx 2000\,\mathrm{km}$, thus it is often appropriate to consider a rigid lid which imposes $L_{\mathrm{e}} = \infty$ and focus on the first internal Rossby radius L_{d}. In the atmosphere, the two radii are much closer with $L_{\mathrm{e}} \approx 2000\,\mathrm{km}$ and $L_{\mathrm{d}} \approx 800\,\mathrm{km}$.

Geostrophic scalings are introduced to capture the regime when rotational forces are dominant. This regime is universally characterized by a small Rossby number (26), but the limit can be approached in different ways distinguished by the scaling of the *Burger number*

$$\mathrm{Bu} = \frac{L_{\mathrm{Ro}}^2}{L^2} = \frac{\mathrm{Ro}^2}{\mathrm{Fr}^2}\,, \tag{63}$$

where, depending on context, L_{Ro} denotes the internal or external Rossby radius of deformation. In the internal wave picture, the Froude number is given by (24). For surface gravity waves in the shallow water framework, (63) can be seen as implicitly defining the shallow water Froude number which, however, is not needed as an independent scaling parameter.

Qualitatively, there are three different regimes. When the Burger number is small, the flow is rotation dominated. The *semi-geostrophic* limit, discussed further below, is representative of this regime. At $\mathrm{Bu} = O(1)$, buoyancy and rotation are both important. The *quasi-geostrophic* limit is representative of this regime. For large Burger numbers, the effects of stratification are dominant; we will not consider this situation further. See, e.g., Babin et al. (2002) for a detailed exposition of the different geostrophic scaling regimes.

We already took a detailed look at quasi-geostrophy in Section 2.5. There, we took the rotating Boussinesq equations as the starting point and scaled with $\mathrm{Bu} = O(1)$ where the Burger number was linked to the internal Rossby radius. We specialized to single-layer and double-layer models in a second step. However, it is also possible to derive a single-layer quasi-geostrophic equation from the rotating shallow water equations. In this case, buoyancy comes from the free surface elevation, so the Burger number must be based on the external Rossby radius. In order to maintain geostrophic balance at leading order, we must assume that variations in the surface elevation are small, more precisely, are $O(\mathrm{Ro})$. It is then possible to derive a quasi-geostrophic potential vorticity by linearizing the shallow water potential vorticity (60) as follows. Taking a constant layer depth $b = H$ with $h \ll H$ and $\zeta \ll f$, we can write

$$q = \frac{f+\zeta}{H}\,\frac{1}{1+h/H} \approx \frac{f+\zeta}{H}\left(1-\frac{h}{H}\right) \approx \frac{1}{H}\left(f - f\,\frac{h}{H} + \zeta\right). \tag{64}$$

Dropping the constant prefactor, recalling that $\zeta = \Delta\psi$, and using leading order geostrophic balance in the form $f\psi \approx gh$, we obtain a new quasi-geostrophic potential vorticity

$$q = f - L_{\mathrm{e}}^{-2}\,\psi + \Delta\psi\,. \tag{65}$$

This expression contains an extra *stretching term* $L_e^{-2}\,\psi$ not present in (50b) which is a remnant of free surface effects. It is required on scales larger than the external Rossby radius, even though the use of the quasi-geostrophic approximation on such scales becomes questionable as changes in the Coriolis parameter are no longer small.

Let us now turn our attention to the situation when the Burger number is small. In this case, the most important distinguished limit has $\mathrm{Bu} = \mathrm{Ro}$. It is referred to as the *semi-geostrophic* limit or the *frontal geostrophic* regime in the context of frontal geostrophic adjustment (Zeitlin et al., 2003). The study of this limit goes back to the geostrophic momentum approximation (Eliassen, 1948). The resulting semi-geostrophic equations were rewritten by Hoskins and solved via an ingenious change of coordinates (Hoskins, 1975; Cullen and Purser, 1984). They continue to attract interest due to their connection to optimal transport theory and the resulting possibility to make mathematical sense of generalized frontal-type solutions (Benamou and Brenier, 1998; Cullen, 2008) and for the turbulence emerging from them (Ragone and Badin, 2016). The geostrophic momentum approximation and Hoskins' transformation inspired Salmon (1983, 1985) to make corresponding approximations directly to Hamilton's principle so as to preserve geometrical structure and automatically preserve conservation laws. Salmon's approach is generalized in Oliver (2006); corresponding results for stratified flow where the primitive equations serve as the parent model are due to Salmon (1996) and Oliver and Vasylkevych (2016). Numerical studies suggest that Salmon's so-called L_1 model is particularly robust and accurate (Allen et al., 2002; Dritschel et al., 2016). A direct numerical comparison of Hoskins' semi-geostrophic equations with its generalized solution structure and the L_1 family of models which possess global classical solutions has not yet been done.

Both semi-geostrophic and quasi-geostrophic models cannot support inertia-gravity waves. Rossby waves are the only possible linear wave solutions; see, e.g., Vallis (2006) for quasi-geostrophic Rossby wave theory. Thus, even though quasi-geostrophic models formally allow larger Burger numbers, their derivation imposes a smallness assumption on buoyancy perturbations, so that it is *a priori* not clear whether they are more accurate than any of the semi-geostrophic models for nearly balanced flow even in the $\mathrm{Bu} = O(1)$ regime. Despite some preliminary attempts aimed at exploring frontal scale turbulence (Badin, 2014), to our knowledge a systematic numerical study has not yet been done.

Mathematically rigorous justifications of the quasi-geostrophic splitting into fast and slow modes have started with the work of Embid and Majda (1996) and Babin et al. (1996, 1997). These ideas are extended and reviewed, for example, by Babin et al. (2002), Majda (2003), Saint-Raymond (2010), and Cheng and Mahalov (2013). The book by Majda (2003), in particular, contains an elementary exposition of averaging over fast waves which can be seen as a deterministic precursor of the stochastic averaging in Section 5.3 below, and Dutrifoy et al. (2009) consider a model for equatorial balance.

A full mathematical justification for any of the semi-geostrophic limits remains, to the best of our knowledge, open.

2.8 Equatorial scalings

Near the equator, the mid-latitude scalings discussed above break down as the rotation vector $\mathbf{\Omega}$ aligns with the horizontal. In the equatorial β-plane approximation, $f_0 = 0$ and the vertical component of the Coriolis force takes the form $f = \beta y$. Correspondingly, the mid-latitude notion of geostrophic balance breaks down. Thus, near the equator the fluid ceases to be approximately constrained to quasi-twodimensional motion. Moreover, at least in the atmosphere, different physics—moist processes and deep convection—become dominant features of the observed dynamics. Whether some notion of balance in the absence of moist processes persists up to and across the equator is currently not well understood.

One instance where these difficulties manifest themselves is the weakening of scale separation between different types of linear waves as compared to the mid-latitudes. For example, linearizing the equatorial rotating shallow water equations about a steady state with constant height field yields the dispersion curves for eigenmodes plotted in Figure 1. In addition to the slow Rossby and Kelvin waves also Yanai waves, which are mixed Rossby-gravity waves, are present. Here, only the Rossby waves reflect the non-trivial Coriolis effect of the β-plane approximation and are geostrophically balanced. However, Kelvin waves are important factors in low frequency variations in the tropics.

As in the mid-latitudes, it is natural to seek a scaling which is able to filter the faster waves in some asymptotic limit. The mid-latitude definition of the Rossby number, where $\mathrm{Ro} = U/(fL)$, becomes singular at the equator and cannot be used as a scaling parameter. Noting that the characteristic length scales may be strongly anisotropic near the equator, we may proceed as follows. Let c denote the characteristic wave

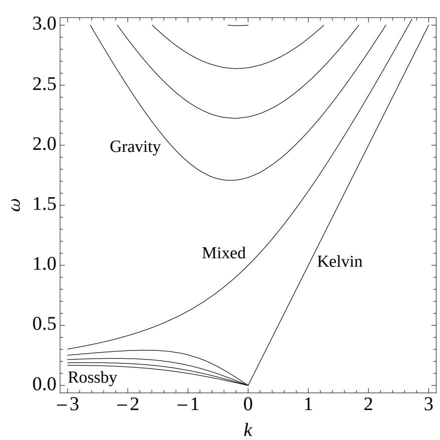

Fig. 1 Dispersion curves of linear equatorial waves of the rotating shallow water equations. Here k is the wavenumber, on the scale $\sqrt{\beta/c_e}$, and ω the frequency, on the scale $\sqrt{\beta c_e}$ with β the meridional gradient of the Coriolis parameter and $c_e = \sqrt{gH}$ the gravity wave speed for mean fluid thickness H.

speed. For example, in the shallow water approximation, $c = c_\mathrm{e} = \sqrt{gH}$ while for stratified quasi-geostrophic or Boussinesq flow, $c = c_1 = NH/\pi$ as introduced in the previous sections. The meridional length scale L_y at which rotation is comparable to wave propagation satisfies $f L_y = c$. Moreover, at this distance from the equator, the vertical Coriolis parameter satisfies $f = \beta L_y$. Solving for the meridional length scale, we obtain

$$L_y = \sqrt{\frac{c}{\beta}}, \tag{66}$$

which can be regarded as an *equatorial Rossby radius of deformation*. In combination with a typical zonal scale L_x, we may define an *equatorial Rossby number* as

$$\mathrm{Ro}_\mathrm{eq} = \frac{U}{f L_x} = \frac{U}{\sqrt{c\beta}\, L_x} \,. \tag{67}$$

Its ratio with the Froude number $\mathrm{Fr} = U/c$ leads to the *equatorial Burger number* as the horizontal aspect ratio

$$\mathrm{Bu} = \frac{L_y^2}{L_x^2} = \frac{\mathrm{Ro}_\mathrm{eq}^2}{\mathrm{Fr}^2} \,. \tag{68}$$

It is typically small; $\varepsilon = \sqrt{\mathrm{Bu}}$ has been used as the main expansion parameter, for example, by Majda (2003) and Chan and Shepherd (2013).

Assuming, in addition, a low vertical aspect ratio in the sense that $H^2 \ll L_x\, L_y$, one obtains hydrostatic balance as for the mid-latitude scaling in Section 2.4 (Chan and Shepherd, 2013). In order to obtain a hierarchy of balanced models from an ε-expansion, they assume $\mathrm{Fr} = 1$ and apply the slaving method of Warn et al. (1995). Balanced models in the tropics which do not assume the traditional approximation were derived by Julien et al. (2006).

3 Variational principles and Hamiltonian mechanics

The inviscid Boussinesq equations and all of the simplified models discussed in Section 2 are infinite dimensional Hamiltonian systems. In this section, we will review several formally equivalent points of view: Hamilton's variational principle in which the equations of motion arise as stationary points of an action functional, the Poisson formulation of Hamiltonian fluid mechanics, and, closely related, the Nambu formulation.

The derivation of each of the models from Section 2 can be made systematic through the use of Hamilton's variational principle. This approach has multiple advantages: first of all, it allows to formulate a geometrical setting of the dynamics: from the study of the Lagrangian density, one can study the continuous symmetries of the system and, by Noether's theorem, derive associated conservation laws; further, one can apply different dynamical approximations directly to the Lagrangian density. When the approximations respect the continuous symmetries, the approximated systems will possess analogous conserved quantities. Fluid Lagrangians, in particular, have a special symmetry, the particle relabeling symmetry, which states that the exchange of fluid labels does not affect the distribution of mass. The associated conserved quantity to this symmetry is the potential vorticity of the fluid. This result shows that potential vorticity does not just appear from a skillful manipulation of the equations of motion, but is instead a signature of a more fundamental property of the system.

While Hamilton's principle has been used to derive numerical methods for partial differential equations (**?**), structure preserving numerical approximations for fluid equations more readily arise by discretizing Poisson or Nambu brackets. While we will not go into the numerical aspects here, we explain the Poisson and Nambu formulations using the rotating shallow water equations as an example.

3.1 Variational Principles

The inviscid form of the Boussinesq equations (1) can be derived from a variational principle as follows. For simplicity, we consider the case of a rigid lid upper boundary. Let $\mathfrak{g}$ denote the Lie algebra of vector fields on Ω satisfying the incompressibility condition (1b) on the domain Ω with impermeability conditions on all boundaries. Let $\boldsymbol{\eta}$ denote the flow of a time dependent vector field $\boldsymbol{u} \in \mathfrak{g}$, i.e.,

$$\dot{\boldsymbol{\eta}}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{\eta}(\boldsymbol{a}, t), t) \qquad \text{with} \qquad \boldsymbol{\eta}(\boldsymbol{a}, 0) = \boldsymbol{a} \,. \tag{69}$$

Here and in the following, the letter $\boldsymbol{a}$ is used for Lagrangian label coordinates, while $\boldsymbol{x} = \boldsymbol{\eta}(\boldsymbol{a}, t)$ denotes the corresponding Eulerian position at time t. As $\boldsymbol{u}$ is divergence free, $\boldsymbol{\eta}$ is volume preserving. In the following, we shall write $\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}$ for short. Correspondingly, advection of density equation, (1c) with $\kappa = 0$, is equivalent to

$$\rho \circ \boldsymbol{\eta} = \rho_{\text{in}} \,, \tag{70}$$

where ρ_{in} is the given initial distribution of the density.

Throughout the chapter, we use the letters L and ℓ (with appropriate subscripts as necessary) to distinguish Lagrangians expressed in Lagrangian and Eulerian quantities, respectively. With this notation in place, the Boussinesq Lagrangian reads

$$L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_{\text{in}}) = \int_\Omega \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \frac{1}{2} |\dot{\boldsymbol{\eta}}|^2 - \frac{g}{\rho_0} \rho_{\text{in}} \, \eta_3 \, \mathrm{d}\boldsymbol{a} \,, \tag{71}$$

where $\boldsymbol{\nabla} \times \boldsymbol{R} = 2\boldsymbol{\Omega}$. We note that L can be expressed in terms of purely Eulerian quantities as

$$L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_{\text{in}}) = \int_\Omega \boldsymbol{R} \cdot \boldsymbol{u} + \frac{1}{2} |\boldsymbol{u}|^2 - \frac{g}{\rho_0} \rho z \, \mathrm{d}\boldsymbol{x} \equiv \ell(\boldsymbol{u}, \rho) \,. \tag{72}$$

The first term in the Lagrangian is the Coriolis term. It only contributes to the symplectic form, but does not feature in the energy. The second and third terms are the difference of kinetic and potential energies, as for simple mechanical systems.

We observe that L is invariant under compositions of the flow map with arbitrary volume and domain preserving maps. This is known as the *particle relabeling symmetry*. For such Lagrangians, the Euler–Poincaré theorem for continua (Holm et al. 1998 or Holm et al. 2009, Theorem 17.8) asserts that the following are equivalent.

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

$$\delta \int_{t_1}^{t_2} L(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}; \rho_{\text{in}})\, \mathrm{d}t = 0 \tag{73}$$

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

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

$$\delta \int_{t_1}^{t_2} \ell(\boldsymbol{u}, \rho)\, \mathrm{d}t = 0\,, \tag{74}$$

where the variations $\delta\boldsymbol{u}$ and $\delta\rho$ are subject to the Lin constraints

$$\delta\boldsymbol{u} = \dot{\boldsymbol{w}} + \nabla\boldsymbol{w}\,\boldsymbol{u} - \nabla\boldsymbol{u}\,\boldsymbol{w} = \dot{\boldsymbol{w}} + [\boldsymbol{u}, \boldsymbol{w}]\,, \tag{75a}$$

$$\delta\rho + \boldsymbol{w} \cdot \nabla\rho = 0\,, \tag{75b}$$

with $\boldsymbol{w}$ as in (i).

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

$$\int_\Omega (\partial_t + \mathcal{L}_{\boldsymbol{u}})\boldsymbol{m} \cdot \boldsymbol{w} + \frac{\delta\ell}{\delta\rho}\, \mathcal{L}_{\boldsymbol{w}}\, \rho\, \mathrm{d}\boldsymbol{x} = 0 \tag{76}$$

for every $\boldsymbol{w} \in \mathfrak{g}$, where $\mathcal{L}$ denotes the Lie derivative and $\boldsymbol{m}$ is the momentum one-form

$$\boldsymbol{m} = \frac{\delta\ell}{\delta\boldsymbol{u}}\,. \tag{77}$$

In the language of vector fields in a region of $\mathbb{R}^3$, the Euler–Poincaré equation (76) reads

$$\int_\Omega \Big(\partial_t\boldsymbol{m} + (\nabla\times\boldsymbol{m})\times\boldsymbol{u} + \nabla(\boldsymbol{m}\cdot\boldsymbol{u}) + \frac{\delta\ell}{\delta\rho}\,\nabla\rho\Big) \cdot \boldsymbol{w}\, \mathrm{d}\boldsymbol{x} = 0 \tag{78}$$

for every $\boldsymbol{w} \in \mathfrak{g}$. Due to the Hodge decomposition, the term in parentheses must be a gradient, i.e.,

$$\partial_t\boldsymbol{m} + (\nabla\times\boldsymbol{m})\times\boldsymbol{u} + \frac{\delta\ell}{\delta\rho}\,\nabla\rho = \nabla\phi\,. \tag{79}$$

Noting that

$$\frac{\delta \ell}{\delta \boldsymbol{u}} = \boldsymbol{R} + \boldsymbol{u} \qquad \text{and} \qquad \frac{\delta \ell}{\delta \rho} = -\frac{gz}{\rho_0}\,, \tag{80}$$

re-defining the potential

$$\phi = -\frac{p}{\rho_0} - \frac{g}{\rho_0}\, z\rho - \frac{1}{2}\, |\boldsymbol{u}|^2\,, \tag{81}$$

and using the vector identity $(\boldsymbol{\nabla} \times \boldsymbol{u}) \times \boldsymbol{u} = \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \frac{1}{2}\, \boldsymbol{\nabla} |\boldsymbol{u}|^2$, we can write the Euler–Poincaré equation for L as

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\frac{1}{\rho_0}\, \boldsymbol{\nabla} p - \frac{g\rho}{\rho_0}\, \boldsymbol{k}\,. \tag{82}$$

Thus, we recovered the inviscid form of the momentum equation (1a).

We remark that the traditional approach to variational derivation of the primitive equations treats the geopotential as a Lagrange multiplier responsible for enforcing the incompressibility constraint. Here, we build the constraint into the definition of the configuration space. The gradient of the geopotential then appears naturally due to the fact that the L^2 pairing with divergence free vector fields determines a vector field only up to a gradient. Both approaches, of course, lead to identical equations of motion.

As the Boussinesq Lagrangian (71) is invariant under time translation, the model possesses a conserved energy of the form

$$H = \int_\Omega \frac{\delta \ell}{\delta \boldsymbol{u}} \cdot \boldsymbol{u}\, \mathrm{d}\boldsymbol{x} - \ell(\boldsymbol{u}, \rho) = \int_\Omega \frac{1}{2}\, |\boldsymbol{u}|^2 + \frac{g}{\rho_0}\, \rho z\, \mathrm{d}\boldsymbol{x}\,. \tag{83}$$

The symmetry of the balance model Lagrangian under particle relabeling leads to a material conservation law for the potential vorticity of the balance model. It can be derived geometrically as follows. First note that the abstract Euler–Poincaré equation (76) can be written

$$(\partial_t + \mathcal{L}_{\boldsymbol{u}})\boldsymbol{m} + \frac{\delta \ell}{\delta \rho}\, \mathrm{d}\rho = \mathrm{d}\phi\,, \tag{84}$$

where d denotes the exterior derivative. Taking the exterior (wedge) product between the exterior derivative of (84) and $\mathrm{d}\rho$, we find that

$$\begin{aligned} 0 &= \mathrm{d}\big((\partial_t + \mathcal{L}_{\boldsymbol{u}})\boldsymbol{m} + \frac{\delta \ell}{\delta \rho}\, \mathrm{d}\rho - \mathrm{d}\phi\big) \wedge \mathrm{d}\rho \\ &= (\partial_t + \mathcal{L}_{\boldsymbol{u}})(\mathrm{d}\boldsymbol{m} \wedge \mathrm{d}\rho) - \mathrm{d}\boldsymbol{m} \wedge \mathrm{d}(\partial_t + \mathcal{L}_{\boldsymbol{u}})\rho \\ &= (\partial_t + \mathcal{L}_{\boldsymbol{u}})(\mathrm{d}\boldsymbol{m} \wedge \mathrm{d}\rho)\,, \end{aligned} \tag{85}$$

where we used the commutativity of Lie and exterior derivatives in the second equality and the advection of ρ in the third equality. In three dimensions, we can identify this conservation law with material advection of the scalar quantity

$$q = *(\mathrm{d}\boldsymbol{m} \wedge \mathrm{d}\rho)\,, \tag{86}$$

where $*$ denotes the Hodge dual operator. Indeed, writing $\mu = \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}z$ to denote the canonical volume form, we have $\mathrm{d}\boldsymbol{m} \wedge \mathrm{d}\rho = q\mu$, so that

$$0 = (\partial_t + \mathcal{L}_{\boldsymbol{u}})(q\mu) = \mu\,(\partial_t + \mathcal{L}_{\boldsymbol{u}})q + q\,\mathcal{L}_{\boldsymbol{u}}\,\mu\,. \tag{87}$$

Since the flow is volume preserving and μ is non-degenerate, this proves that $\partial_t q + \mathcal{L}_{\boldsymbol{u}}\, q = 0$, i.e., q is conserved on fluid particles.

Going back to the language of vector calculus, using expression (80) for the momentum and writing $\boldsymbol{\omega} = \boldsymbol{\nabla} \times \boldsymbol{u}$ for the relative vorticity, we recover the familiar expression for the *Ertel potential vorticity*,

$$q = (\boldsymbol{\nabla} \times \boldsymbol{m}) \cdot \boldsymbol{\nabla}\rho = (2\boldsymbol{\Omega} + \boldsymbol{\omega}) \cdot \boldsymbol{\nabla}\rho\,. \tag{88}$$

We remark that the derivation above corresponds to taking the inner product of the curl of the Euler–Poincaré equations (79) with $\boldsymbol{\nabla}\rho$ and manipulating correspondingly; the advantage of the abstract approach is that commuting exterior and Lie derivative in traditional notation is not linked to any intrinsic operation, thus requires tedious verification.

3.2 Variational model reduction

The variational principle which underlies the equations of motion can be used to derive simplified models. All approximations are done at the level of the Lagrangian; the simplified equations of motion then arise from the approximated Lagrangian in a second step. When the approximations respect the symmetries of the Lagrangian, the associated conservation laws, in our setting the conservation of energy and of potential vorticity, will be preserved as well.

The use of the Hamilton principle to derive balance models was pioneered by Salmon (1985, 1996). Salmon's method was generalized by Oliver (2006) where the approximate balance manifold is interpreted as a Dirac constraint induced by a truncated change of coordinates. Structure preservation, however, does not imply well-posedness of the initial value problem, as is evident in the numerical study of Dritschel et al. (2016) which indicates a strong preference for Salmon's so-called L_1-model. Çalık et al. (2013) prove global well-posedness for a large class of variational balance models, including the L_1-model. Gottwald and Oliver (2014) give a justification of the general method to arbitrary order of approximation in a finite dimensional model context, and Oliver and Vasylkevych (2013, 2016) show that the variational approach is flexible enough to cover spatially varying Coriolis functions and stratified flow, respectively.

The use of language from geometry and of the associated variational principles is particularly advantageous when working in general curvilinear coordinates (Tort and Dubos, 2014) which is crucial when working with global models but not as pertinent to the more theoretical questions we raise here and will not be discussed further.

In a different line of research, Koide and Kodama (2012) and references cited therein demonstrate that a stochastic version of the Hamilton principle can give rise to various dissipative partial differential equations, including the Navier–Stokes equations and higher order corrections to harmonic diffusion. A different approach is taken by Holm (2015) who derives fully stochastic fluid equations from a variational principle and discusses their conservation law structure.

3.3 Poisson formulation

All of the models above can also be cast in a Poisson or a Nambu formulation. We illustrate this using the rotating shallow water equations as an example.

The rotating shallow water equations can be cast in a noncanonical Poisson formulation (see, e.g., Shepherd 1990). Suppose F is an arbitrary functional of $\boldsymbol{u}$ and h. Then

$$\dot{F} = \{F, H\} \tag{89}$$

with shallow water Hamiltonian H given by (59) and Poisson bracket

$$\{F, H\} = \int_\Omega q\, F_{\boldsymbol{u}}^\perp \cdot H_{\boldsymbol{u}} - F_{\boldsymbol{u}} \cdot \boldsymbol{\nabla} H_h + H_{\boldsymbol{u}} \cdot \boldsymbol{\nabla} F_h \, \mathrm{d}\boldsymbol{x} \tag{90}$$

where we use subscripts to express functional derivatives. A Poisson bracket is skew-symmetric, satisfies the Leibniz rule

$$\{f, gh\} = \{f, g\}\, h + \{f, h\}\, g \tag{91}$$

and the Jacobi identity

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0\,. \tag{92}$$

The shallow water equations (58) are recovered by substituting point evaluations of h and $h\boldsymbol{u}$ for the functional F. The Poisson formulation can also be written in the form

$$\dot{F} = \int_\Omega F_{\boldsymbol{\xi}}^T\, \mathbb{J}\, H_{\boldsymbol{\xi}} \, \mathrm{d}\boldsymbol{x}\,, \tag{93}$$

where $\boldsymbol{\xi} = (\boldsymbol{u}, h)$ and $\mathbb{J}$ is the noncanonical Poisson operator

$$\mathbb{J} = -\begin{pmatrix} qJ & \boldsymbol{\nabla} \\ \boldsymbol{\nabla}^T & 0 \end{pmatrix}\,. \tag{94}$$

Here J denotes the canonical 2×2 symplectic matrix $J\boldsymbol{w} \equiv \boldsymbol{w}^\perp$ for any $\boldsymbol{w} \in \mathbb{R}^2$ and we think of $\boldsymbol{\nabla}$ as a column operator.

From (94) it is possible to find the Casimir invariants of the system, i.e., functionals C satisfying

$$\{F, C\} = 0 \tag{95}$$

for every functional F. Expressing the bracket in terms of the Poisson operator defined via (93) and (94), we see that the Casimirs are precisely the functionals which belong to the kernel of $\mathbb{J}$, i.e.,

$$\mathbb{J}\, C_{\boldsymbol{\xi}} = 0\,. \tag{96}$$

When $\mathbb{J}$ is invertible, the Casimirs are just the constant functionals with respect to $\boldsymbol{\xi}$. For the noncanonical Poisson brackets of fluid dynamics, however, $\mathbb{J}$ typically has a nontrivial kernel, i.e., nontrivial Casimirs. Here, for the rotating shallow water equations, it can be shown that there is a class of Casimirs of the form

$$C = \int_\Omega h\, \gamma(q)\, \mathrm{d}\boldsymbol{x}\,, \tag{97}$$

where γ is an arbitrary function of potential vorticity, not to be confused with the ageostrophic vorticity. Indeed, by the chain rule,

$$C_{\boldsymbol{u}} = h\, \gamma'(q)\, q_{\boldsymbol{u}} = -\boldsymbol{\nabla}^\perp \gamma'(q)\,, \tag{98a}$$

$$C_h = \gamma(q) - q\, \gamma'(q)\,, \tag{98b}$$

so that, using (94),

$$\mathbb{J}\, C_{\boldsymbol{\xi}} = \begin{pmatrix} -q\, \boldsymbol{\nabla}\, \gamma'(q) - \boldsymbol{\nabla}\, \gamma(q) + \boldsymbol{\nabla}(q\, \gamma'(q)) \\ \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}^\perp \gamma'(q) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\,. \tag{99}$$

Hence, the Casimir condition (97) is satisfied. The conservation law for mass correspond to $\gamma = 1$, while $\gamma(q) = q$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega h\, q\, \mathrm{d}\boldsymbol{x} = \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \zeta_{\mathrm{a}}\, \mathrm{d}\boldsymbol{x} = 0\,, \tag{100}$$

where the first equality is due to (60). By Stokes' theorem, (100) corresponds to the usual conservation of circulation, which implies conservation of potential vorticity (60). More generally, the choice $\gamma(q) = q^n$ corresponds to the conservation of the nth order enstrophy.

We remark that the Poisson formulation above can be cast into a mixed Eulerian–Lagrangian form that can be seen as a precursor to numerical particle or particle-mesh schemes (Bokhove and Oliver, 2006).

3.4 Nambu formulation

Even if the Hamiltonian formulation of the equations of motion for fluid flows is useful to address questions regarding the geometry of the system, including the study of continuous symmetries and associated conservation laws through Noether's theorem, a systematic method for the derivation of the noncanonical Poisson brackets is still generally lacking and the derivation often relies on guesswork. An alternative formulation of dynamics was proposed by Nambu (1973), who suggested an extension of Hamiltonian dynamics which is based on Liouville's theorem and, differently from classical Hamiltonian mechanics, makes use of several conserved quantities, i.e. the Casimirs of the system, which can be considered as additional Hamiltonians and define manifolds whose intersection determine the trajectory in state space. In analogy to Poisson brackets in Hamiltonian mechanics, the resulting dynamics is determined by the Nambu bracket. A Nambu bracket is an n-linear map acting on smooth functions on a manifold that is completely antisymmetric, satisfies the Leibniz rule

$$\{f_1, \ldots, f_{n-1}, gh\} = \{f_1, \ldots, f_{n-1}, g\}\, h + \{f_1, \ldots, f_{n-1}, h\}\, g \tag{101}$$

and the Jacobi identity

$$\begin{aligned} &\{\{f_1, \ldots, f_{n-1}, g_1\}, g_2, \ldots, g_n\} + \{g_1, \{f_1, \ldots, f_{n-1}, g_2\}, g_3, \ldots, g_n\} + \ldots \\ &\quad + \{g_1, \ldots, g_{n-1}, \{f_1, \ldots, f_{n-1}, g_n\}\} = \{f_1, \ldots, f_{n-1}, \{g_1, \ldots, g_n\}\}\,. \end{aligned} \tag{102}$$

A general theory of Nambu–Poisson structures was outlined by Takhtajan (1994).

For the rotating shallow water equations it is convenient to pass from $(\boldsymbol{u}, h)$ to (ζ, δ, h). Using the chain rule for functional derivatives,

$$F_{\boldsymbol{u}} = -\nabla^\perp F_\zeta - \nabla F_\delta \tag{103}$$

while F_h remains unchanged. With (103), the Poisson bracket for the rotating shallow-water equations (90) takes the form of the sum of two Poisson brackets and a Nambu bracket (Salmon, 2005, 2007), i.e.

$$\dot{F} = \{F, H\}_{\delta\delta} + \{F, H\}_{\zeta\zeta} + \{F, H\}_{\zeta\delta h} \tag{104}$$

with

$$\{F, H\}_{\delta\delta} = \int_\Omega q\, \nabla^\perp F_\delta \cdot \nabla H_\delta \,\mathrm{d}\boldsymbol{x}\,, \tag{105a}$$

$$\{F, H\}_{\zeta\zeta} = \int_\Omega q\, \nabla^\perp F_\zeta \cdot \nabla H_\zeta \,\mathrm{d}\boldsymbol{x}\,, \tag{105b}$$

and

$$\{F,H\}_{\zeta\delta h} = \int_\Omega q\,(\nabla F_\delta\cdot\nabla H_\zeta - \nabla H_\delta\cdot\nabla F_\zeta) + \nabla F_\delta\cdot\nabla H_h - \nabla H_\delta\cdot\nabla F_h\,\mathrm{d}\boldsymbol{x}\,. \tag{105c}$$

The subscripts indicate the variables entering the functional derivatives in the brackets. Each of these brackets is antisymmetric and has Casimir functionals given by the kernel of its associated Poisson operator, i.e., satisfy (97). Once again, of particular interest are the n moments

$$Z_n = \frac{1}{2+n}\int_\Omega h\,q^{n+2}\,\mathrm{d}\boldsymbol{x} = \frac{1}{2+n}\int_\Omega \frac{\zeta_{\mathrm{a}}^{n+2}}{h^{n+1}}\,\mathrm{d}\boldsymbol{x}\,, \tag{106}$$

where $n = 0$ yields the usual enstrophy. Using (106), the Poisson brackets in (104) can be rewritten as

$$\begin{aligned}\{F,H\}_{\delta\delta} &= \{F,H,Z_n\}_{\delta\delta\zeta}\\ &= \frac{1}{3+2n}\int_\Omega \frac{1}{q^n}\,[J(F_\mu,H_\mu)\,(Z_n)_\zeta + \mathrm{cyc}(F,H,Z_n)]\,\mathrm{d}\boldsymbol{x}\,, \end{aligned}\tag{107a}$$

$$\begin{aligned}\{F,H\}_{\zeta\zeta} &= \{F,H,Z_n\}_{\zeta\zeta\zeta}\\ &= \frac{1}{3+2n}\int_\Omega \frac{1}{q^n}\,[J(F_\zeta,H_\zeta)\,(Z_n)_\zeta + \mathrm{cyc}(F,H,Z_n)]\,\mathrm{d}\boldsymbol{x}\,, \end{aligned}\tag{107b}$$

and

$$\begin{aligned}\{F,H\}_{\zeta\delta h} &= \{F,H,Z_n\}_{\zeta\delta h}\\ &= -\frac{1}{1+n}\int_\Omega \frac{1}{q^n}\Big(\frac{\partial_x F_\mu\,\partial_x H_\zeta - \partial_x H_\mu\,\partial_x F_\zeta}{\partial_x q}\,\partial_x (Z_n)_h + \mathrm{cyc}(F,H,Z_n)\\ &\quad + \frac{\partial_y F_\mu\,\partial_y H_\zeta - \partial_y H_\mu\,\partial_y F_\zeta}{\partial_y q}\,\partial_y (Z_n)_h + \mathrm{cyc}(F,H,Z_n)\Big)\,\mathrm{d}\boldsymbol{x}\,, \end{aligned}\tag{107c}$$

where "cyc" denotes expressions derived from the previous term by all cyclic permutations of the indicated symbols.

The sum of the three brackets comprises the complete Nambu bracket for the shallow water equations. This formulation has the advantage that it is easily discretized. The resulting numerical schemes have excellent conservation properties (Salmon, 2005). Note that in general, each Poisson bracket corresponds to an infinite number of distinct Nambu brackets, depending on the set of Casimirs which are used (Chatterjee, 1996). This non-uniqueness of the Nambu formulation implies the existence of different conservative numerical schemes for the same set of equations, allowing substantial flexibility for the problem under consideration, for example for the elimination of the singularities emerging from the terms $(\partial_x q)^{-1}$ and $(\partial_y q)^{-1}$ in the third bracket through a rewriting of the brackets before discretization.

For the Nambu brackets for other geophysical fluid equations, including Rayleigh–Bénard convection, see Blender and Badin (2015).

4 Dissipation, turbulence, and nonlinear waves

4.1 Viscosity and dissipation

So far, we have looked at a variety of models for geophysical flow without dissipation. However, consistent modeling of viscous dissipation and boundary friction is essential for ensuring proper distribution of energy across scales. Including frictional forces is a non-trivial problem, in particular when looking at effective models for large scale motion and for numerical simulation where physical dissipation ranges are typically far smaller than what can be numerically resolved.

At the microscopic level, the molecular kinematic viscosity for geophysical fluid flow is modeled as a Newtonian viscosity. Likewise, buoyancy is diffusive with a harmonic diffusion operator, as indicated in our initial formulation of the Boussinesq equations (1). While the viscosity coefficients are extremely small, their presence is important on a mathematical level due to the regularizing or smoothing effect. Moreover, small bulk viscosities and frictional layers, though thin, influence the large scale flow even in the limit of vanishing viscosity, which gives good reasons to take viscous effects into account even for laminar flow. An example is uniformly enhanced Lagrangian drift from an oscillatory layer (Julien and Knobloch, 2007). An early overview on the linear stability of stratified flow with viscosity can be found in Emanuel (1979, 1984).

Crucially, enstrophy or energy cascades in geostrophic or fully developed three-dimensional turbulence require an energy sink at the small scales. Thus, energetically consistent models require dissipation usually referred to as "turbulent viscosity" or "eddy viscosity." Finally, dissipation is also implicitly or explicitly required to ensure the stability of numerical schemes.

The simplest *ad hoc* approach adds dissipation to the horizontal and vertical velocities with different viscosity coefficients (Pedlosky, 1987). This results in the following form of the viscous rotating Boussinesq equations

$$D_t \boldsymbol{u}_{\mathrm{h}} + (2\boldsymbol{\Omega} \times \boldsymbol{u})_{\mathrm{h}} = -\frac{1}{\rho_0} \boldsymbol{\nabla}_{\mathrm{h}} p + \nu_{\mathrm{h}} \Delta \boldsymbol{u}_{\mathrm{h}}, \tag{108a}$$

$$D_t w + \boldsymbol{\Omega}_{\mathrm{h}}^{\perp} \cdot \boldsymbol{u}_h = -\frac{1}{\rho_0} \partial_z p - \frac{g\rho}{\rho_0} + \nu_z \Delta w \tag{108b}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0, \tag{108c}$$

$$D_t \rho = \kappa \Delta \rho, \tag{108d}$$

where ν_{h} and ν_z are coefficients of horizontal and vertical eddy viscosity, respectively.

In single-layer simplified models such as the barotropic quasi-geostrophic equation (50), dissipation comes in the form of bottom drag as the main energy sink and viscosity (or hyperviscosity) as the main enstrophy sink. Under the β-plane approximation $f = f_0 + \beta y$, the dynamics can be written in terms of the relative vorticity $\zeta = \Delta\psi$, so that

$$\partial_t \zeta + \nabla^\perp \psi \cdot \nabla \zeta + \beta\, \partial_x \psi = F - \lambda\, \zeta + \nu\, \Delta \zeta\,, \tag{109}$$

where the terms on the right hand side are forcing F, linear (Rayleigh) damping to model bottom friction with parameter λ describing an inverse time scale for the vorticity decay due to bottom drag, and Newtonian viscosity with parameter ν. In a numerical simulation of a forced system, viscosity is typically replaced by hyperviscosity to remove enstrophy without depleting energy across most resolved scales. To model geostrophic turbulence, both types of dissipation are necessary. A more detailed discussion of this topic is provided in **[provide proper reference to M3 chapter in this volume]**.

The relevance of the form of viscosity for the mathematical theory is highlighted by the fact that only very recently the well-posedness of the primitive equations with only horizontal viscosity has been established (Cao et al., 2016). For full viscosity in all directions much more is known, in particular the asymptotic stability of (nearly) geostrophically balanced dynamics was shown in Temam and Wirosoetisno (2010). Inclusion of at least some kind of dissipation is crucial for these results, but is not necessarily physically consistent nor suitable as a model for energy dissipation on these scales, which is the motivation for the study in Olbers and Eden (2013).

As mentioned, there are no universally accepted criteria for the correct form of dissipation, in particular in connection with numerical schemes. However, a noteworthy prominent approach motivated also by the problem of unfeasible small scale resolution in practice is "Large Eddy Simulation" (LES) as discussed, e.g., in Sagaut (2006).

In the (next) simplest form, the subgrid closure and parametrization problem of viscosity yields an effective "filtered" Navier-Stokes model for the velocity vector $\boldsymbol{u}$ in which the turbulent viscosity term reads

$$-\nabla \cdot (\nu_e(\boldsymbol{u})\, S(\boldsymbol{u}))\,, \tag{110}$$

where $S(\boldsymbol{u}) = \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)$ is the rate-of-strain tensor and $\nu_e(\boldsymbol{u})$ the eddy viscosity. For instance, Smagorinsky (1963) considers $\nu_e(\boldsymbol{u}) = C\,|S(\boldsymbol{u})|$ with constant C. Variants that are more faithful in preserving the turbulent kinetic energy spectrum can be found in Schaefer-Rolffs and Becker (2013), Schaefer-Rolffs et al. (2015), and Trias et al. (2015).

On a different level, we note that including viscous terms necessitates additional boundary conditions. This in turn yields viscous so-called Ekman boundary layers, which feed back onto the large-scale motion. We also point out that balancing eddy buoyancy fluxes and viscous effects may necessitate additional scaling anisotropy (Grooms et al., 2011).

Concluding this discussion we emphasize that the presence of dissipation requires to include driving mechanism in order to maintain non-trivial flow. The precise form of the source term is yet another non-trivial modeling issue. Simple idealized configurations are discussed in **[provide proper reference to M3 chapter in this volume]**; for a discussion in a more realistic context, see, for example, Olbers and Eden (2013).

4.2 Nonlinear waves and dynamical systems methods

Wave phenomena at different scales organize the structure of geophysical flows and play a vital role for the transport of energy. From a bifurcation theory viewpoint, linear waves raise expectations for nonlinear wave phenomena. The only nonlinear terms in the standard models as introduced above stem from the material derivative. Geostrophic balance partly turns these into a Poisson bracket nonlinearity, which vanishes for monochromatic plane waves. For this reason, the barotropic quasi-geostrophic equation (50) possesses linear Rossby waves with unconstrained amplitude as exact nonlinear solutions and, more generally, special plane waves can sometimes be exact solutions of nonlinear fluid equations (Majda, 2003; Julien and Knobloch, 2007). However, the barotropic quasi-geostrophic equations are special in this regard. For its nonlinear parent models, waves are not of plane form but nonlinearly selected, even in the unstratified setting of the rotating shallow water equations.

In a different manner, baroclinic instability relates to bifurcations from Rossby waves via amplitude equations. A discussion of this for a two layer model can be found in Pedlosky (1987). On the level of the equatorial shallow water wave equations, the amplitude equation approach to nonlinear phenomena has been exploited by Boyd in the 1980s to identify various modulated Rossby and Kelvin waves, cf. Boyd (1980) and the more recent Boyd (2002).

However, nonlinear waves in geophysical flows received much less attention after the period of seminal progress in the 1980s. In the past decade the subject has regained momentum in particular from a more mathematical viewpoint (Zeitlin et al., 2003; Bouchut et al., 2005; Constantin, 2013; Hsu, 2014); see Khouider et al. (2013) for a recent review. A major motivation for the study of nonlinear waves is their role in balanced flow, energy transport, and wave breaking in the form of shocks for inviscid models (Zeitlin et al., 2003).

The vast majority of research considers idealized planar or cylindrical geometry and either no viscosity or simple molecular dissipation. The effect of viscosity and Ekman layer formation has been mathematically studied in Dalibard and Saint-Raymond (2010) concerning existence and stability of steady states in a certain scaling limit of fast rotation and small layer thickness. Rossby–Haurwitz waves are explicit nonlinear spherical waves which received increasing attention recently (Thuburn and Li, 2000; Callaghan and Forbes, 2006; Ibragimov, 2011; Schubert et al., 2009; Smith and Dritschel, 2006; Boyd and Zhou, 2008). An "intermediate" model accounting for geometric terms from the Mercator projection has been proposed in Bates and Grimshaw (2014) and provides a possible connection between planar and spherical nonlinear wave phenomena.

Another aspect of nonlinear waves is the emergence of shocks in frontogenesis and the adjustment problem. With attention to energy transfer in the geostrophic adjustment problem, this has been studied in Blumen and Wu (1995) and more recently by Reznik (2015), where additional references can be found. In a simplified one-dimensional setting of the shallow water equations, an essentially explicit approach to fronts has been given in Plougonven and Zeitlin (2005), also see **[provide reference to Chapter L2 in this volume]**.

While dynamical systems methods have been originally developed for finite dimensional problems, dissipative or more generally parabolic partial differential equations often have a finite dimensional character, which allows for application of tools from generic dynamical systems theory. For instance, compactness and viscosity allow to infer the existence of an invariant and attracting inertial manifold as proven for the spherical Navier–Stokes equations in Temam and Wang (1993).

For domains with one large or unbounded direction the spatial dynamical systems perspective pushed nonlinear wave theory, in particular regarding stability, for prototypical parabolic equations in the past decades; see, e.g., Kapitula and Promislow (2013), Sandstede (2002), Meyries et al. (2014), and references therein. In the much more prominent inviscid and dispersive case, the presence of conserved quantities gives the mathematical theory a different character, especially concerning stability. We refer to the recent study of Balmforth et al. (2013).

In general, the nonlinear response of a nonlinear parabolic system to linear instabilities can be cast in terms of reduced equations which filter different aspects. On extended domains, the aforementioned modulation equations describe the slow dynamics over large scales near onset; rigorous error estimates have been derived in several contexts (e.g. Schneider, 1994; Doelman et al., 2009). For parabolic systems, the classical center manifold reduction provides a conjugacy of the full dynamical system to a decoupled product of trivial linear flow and a lower dimensional nonlinear part (unique up to exponential error) that is more amenable to analysis, in particular via normal forms (Haragus and Iooss, 2011). More selective filtering of bifurcating solutions can be cast in terms of Lyapunov–Schmidt reduction, a standard tool to prove existence of non-trivial bifurcating solutions in amplitude equations or truncated normal forms on a center manifold (Vanderbauwhede, 2012). Indeed, the full flow of a reduced equation may be an invalid approximation (over infinite time horizons), but selected solutions may still have exact counterparts in the full system. Viscous regularization can also act as a filter to reduce the complexity of bifurcations at onset, such as resonance phenomena and allow to determine nonlinear stability for objects other than single shocks (Beck et al., 2010).

The bifurcation theory viewpoint can be exploited numerically by using continuation algorithms to compute branches of nonlinear solutions, their stability and bifurcations and thus, for these solutions, offers an alternative to direct numerical simulation. Dedicated software packages include AUTO[1], LOCA[2], and pde2path[3].

In the context of geophysical flows, bifurcations naturally occur in compartment models or coupled ocean atmosphere systems, e.g. explored numerically in Dijkstra et al. (2014). A review of bifurcation analyses can be found in Simonnet et al. (2009) and more broadly for nonlinear waves in fluid turbulence in Kawahara et al. (2012). Co-existing branches of equatorial nonlinear waves have been found numerically by

[1] http://indy.cs.concordia.ca/auto/

[2] http://www.cs.sandia.gov/loca/

[3] http://www.staff.uni-oldenburg.de/hannes.uecker/pde2path/

Boyd (2002); Boyd and Zhou (2008). Towards nonlinear waves in layered models, an approach that simplifies nonlinear interactions is to retain only resonance triads. A recent study of this type can be found in Bates and Grimshaw (2014).

5 Stochastic Model Reduction

Recent advances in systematic stochastic climate modeling (Majda et al., 2008; Franzke et al., 2005, 2015; Franzke and Majda, 2006; Monahan and Culina, 2011) provided new insights into the structural form of the terms accounting for the interaction between resolved and unresolved processes. The approach is based on the adiabatic elimination of fast variables (Kurtz, 1973; Papanicolaou, 1976; Gardiner, 2009) and demonstrates the necessity of nonlinear damping and state-dependent noise; previously, only linear damping and additive noise have been considered. State-dependent noise is known to be responsible for noise-induced drifts (Gardiner, 2009), which has the potential to ameliorate some of the know biases. The Mori–Zwanzig formalism (e.g. Zwanzig, 2001; Chorin et al., 2000; Wouters and Lucarini, 2013; Gottwald et al., 2017) predicts the emergence of memory terms in reduced-order models. Memory terms are rarely considered in current parametrization schemes. Thus, there is an urgent need to develop a systematic framework for the interaction between resolved and unresolved processes and their representation in numerical climate models.

In the context of climate science, stochastic models were first proposed by Hasselmann (1976). A major advance came with the development of the stochastic mode reduction strategy, which is described in detail in Majda et al. (2001, 2008) and is applied, for example, in Majda et al. (2001, 2002, 2003), Franzke et al. (2005), Franzke and Majda (2006), and Franzke et al. (2015). Stochastic mode reduction starts from the equations used in a climate model, with an external forcing, a linear operator and a quadratic nonlinear operator. Splitting the state vector into slow and fast components, assuming scale separation, and replacing the quadratic self-interaction of the fast modes by a stochastic process leads to a stochastic differential equation for the slow variables alone by using the stochastic mode reduction procedure (see next subsection). By doing so, structurally new terms arise such as a deterministic cubic term which acts predominantly as nonlinear damping and both additive and multiplicative noise terms. The multiplicative noise and the cubic term stem from the nonlinear interaction between the resolved and unresolved modes. Rigorous justification is possible only in the limit of time scale separation, though in practice the reduced order models perform well even in the case of moderate or no time scale separation (Majda et al., 2008; Dolaptchiev et al., 2013; Stinis, 2006; Franzke et al., 2005; Franzke and Majda, 2006). How to do this in a systematic fashion is the topic of the next subsection.

5.1 Basic setup

To illustrate the basic idea of the stochastic mode reduction strategy, we consider the abstract dynamical system

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \boldsymbol{F} + Lz + B(z, z)\,, \tag{111}$$

where z is the state vector, $\boldsymbol{F}$ denotes the forcing, L a linear and B a quadratic nonlinear operator. For convenience, we assume that F is constant in time; for time dependent forcing, see Franzke (2013). The linear operator L contains, in particular, advection, the effect of topography, and linear damping. The operator B conserves energy and satisfies the Liouville property, i.e., the dynamical system with only B on its right hand side is measure preserving (for details, see Franzke et al. 2005). We note that important simplified models for geophysical flow such as the barotropic vorticity equation or the quasi-geostrophic equations can be studied in this framework; see Franzke et al. (2005) for the former and Franzke and Majda (2006) for the latter.

The state vector $z = (\boldsymbol{x}, \boldsymbol{y})$ is now split into slow modes $\boldsymbol{x}$ and fast modes $\boldsymbol{y}$. This decomposition is typically done using Empirical Orthogonal Function (EOF) analysis (Franzke et al., 2005; Franzke and Majda, 2006) or Principal Interaction Patterns (PIP) (Kwasniok, 2004; Crommelin and Majda, 2004). These patterns now constitute a complete orthonormal basis and can be used as basis functions in the same way as Fourier modes are used for spectral models (Holton and Hakim, 2012). Furthermore, because the leading EOFs typically also decay the slowest (Franzke et al., 2005; Franzke and Majda, 2006) it is sensible to use the leading EOFs as the resolved modes. Now we can rewrite (111) in terms of slow and fast modes:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{F}^x + L^{xx}\boldsymbol{x} + L^{xy}\boldsymbol{y} + B^{xxx}(\boldsymbol{x}, \boldsymbol{x}) + B^{xxy}(\boldsymbol{x}, \boldsymbol{y}) + B^{xyy}(\boldsymbol{y}, \boldsymbol{y})\,, \tag{112a}$$

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{F}^y + L^{yx}\boldsymbol{x} + L^{yy}\boldsymbol{y} + B^{yxx}(\boldsymbol{x}, \boldsymbol{x}) + B^{yxy}(\boldsymbol{x}, \boldsymbol{y}) + B^{yyy}(\boldsymbol{y}, \boldsymbol{y})\,. \tag{112b}$$

We use here the following notation: the first superscript denotes the variable (left hand side time derivative) the corresponding right hand side term acts on (e.g. F^x acts on $\boldsymbol{x}$), the second and third superscripts denote the variables whose actions and interactions induce a tendency.

In order to carry out the stochastic mode reduction strategy, we introduce a small parameter ε which quantifies the time scale separation between the slow modes $\boldsymbol{x}$ and fast modes $\boldsymbol{y}$. The parameter ε can also be interpreted as the ratio of the autocorrelation time scale between the slow and the fast modes. We posit

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{F}^x + L^{xx}\boldsymbol{x} + B^{xxx}(\boldsymbol{x}, \boldsymbol{x}) + \frac{1}{\varepsilon}\left(L^{xy}\boldsymbol{y} + B^{xxy}(\boldsymbol{x}, \boldsymbol{y}) + B^{xyy}(\boldsymbol{y}, \boldsymbol{y})\right), \tag{113a}$$

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \frac{1}{\varepsilon}\left(\boldsymbol{F}^y + L^{yx}\boldsymbol{x} + L^{yy}\boldsymbol{y} + B^{yxx}(\boldsymbol{x}, \boldsymbol{x}) + B^{yxy}(\boldsymbol{x}, \boldsymbol{y})\right) + \frac{1}{\varepsilon^2} B^{yyy}(\boldsymbol{y}, \boldsymbol{y})\,. \tag{113b}$$

While the introduction of the time scale parameter ε in front of some of the tendency terms is somewhat arbitrary and currently mainly based on physical intuition, our current research aims at putting it on a more systematic footing by using multi-scale asymptotics for geophysical flows (Klein, 2010; Shaw and Shepherd, 2009; Dolaptchiev and Klein, 2013).

So far, the equations are fully deterministic. We now make the following *stochastic modeling assumption*: we assume that deterministic nonlinear terms involving $\boldsymbol{x}$ are mixing with sufficiently fast decay of correlation so that the nonlinear term involving only the fast modes $B^{yyy}(\boldsymbol{y},\boldsymbol{y})$ can be represented by a stochastic term (Majda et al., 1999, 2001), i.e., we approximate

$$\frac{1}{\varepsilon^2} B^{yyy}(\boldsymbol{y},\boldsymbol{y})\,\mathrm{d}t \approx \text{Stochastic Process}\,. \tag{114}$$

The intuition is that this term is effectively delta-correlated on the slow time scale. It is illustrative to assume that the stochastic process has the form of an Ornstein–Uhlenbeck (OU) process

$$\frac{1}{\varepsilon^2} B^{yyy}(\boldsymbol{y},\boldsymbol{y})\,\mathrm{d}t \approx -\frac{\Gamma}{\varepsilon^2}\,\boldsymbol{y}\,\mathrm{d}t + \frac{\sigma}{\varepsilon}\,\mathrm{d}\boldsymbol{W}\,, \tag{115}$$

where Γ and σ are positive definite matrices and $\boldsymbol{W}$ is a vector-valued Wiener process. However, the stochastic process does not need to be explicitly specified as shown by Franzke et al. (2005). Inserting (115) into the deterministic model (113), we obtain a system of stochastic differential equations

$$\mathrm{d}\boldsymbol{x} = \Big(\boldsymbol{F}^x + L^{xx}\boldsymbol{x} + B^{xxx}(\boldsymbol{x},\boldsymbol{x}) + \frac{1}{\varepsilon}\,\big(L^{xy}\boldsymbol{y} + B^{xxy}(\boldsymbol{x},\boldsymbol{y}) + B^{xyy}(\boldsymbol{y},\boldsymbol{y})\big)\Big)\,\mathrm{d}t\,, \tag{116a}$$

$$\mathrm{d}\boldsymbol{y} = \Big(\frac{1}{\varepsilon}\,\big(\boldsymbol{F}^y + L^{yx}\boldsymbol{x} + L^{yy}\boldsymbol{y} + B^{yxx}(\boldsymbol{x},\boldsymbol{x}) + B^{yxy}(\boldsymbol{x},\boldsymbol{y})\big) - \frac{\Gamma}{\varepsilon^2}\,\boldsymbol{y}\Big)\,\mathrm{d}t + \frac{\sigma}{\varepsilon}\,\mathrm{d}\boldsymbol{W}\,. \tag{116b}$$

We interpret the SDE in (116b) in the sense of Itô. In general, there are two main ways of evaluating the stochastic integrals featuring in the integral form of an SDE, Itô and Stratonovich. From the mathematical perspective, both interpretations are different but interchangeable as an Itô integral can always be rewritten in terms of a Stratonovich integral and vice versa. In modeling, however, the SDE usually arises from a more complicated process or dynamical system in a certain asymptotic limit. In this case, an Itô differential equation represents the situation where the limit dynamics has vanishing autocorrelation while a Stratonovich interpretation represents the situation when the limit dynamics retains finite autocorrelation; see, e.g., the discussion in Moon and Wettlaufer (2014). In our situation, the modeling assumption already entails the idea that terms which cause finite autocorrelations are kept in the deterministic part of (116b) so that the noise term only represents the asymptotically uncorrelated parts. Thus, the noise term in (116b) should be interpreted in the sense of Itô. We remark that from the technical perspective, an Itô process has the Martingale property which greatly simplifies working with stochastic time integrals. However, functions of Itô processes cannot be differentiated using the normal chain rule as an extra term arises. The stochastic chain rule is called Itô's Lemma and gives rise to the parabolic term in the Fokker–Planck and Kolmogorov backwards equations which we will encounter in the next section; for a background on these concepts, see, e.g., Gardiner (2009).

5.2 Slow dynamics via the Kolmogorov backward equation

In the following, we seek an effective equation for the slow modes of the system (116) in the limit $\varepsilon \to 0$. The method uses adiabatic elimination of fast variables (Kurtz, 1973; Papanicolaou, 1976; Pavliotis and Stuart, 2008) and has been pioneered in the present setting by Majda et al. (1999, 2001). The strategy is the following. Every Itô stochastic differential equation, in particular our system (116), has an associated Kolmogorov backward equation (KBE). (For deterministic systems, ODE or PDE, this equation also exists and is usually referred to as the Liouville equation.) It is a parabolic PDE that describes the backward-in-time evolution of the probability distribution function of the system for a given hit probability. While it is rarely possible to solve KBEs numerically in more than a small number of dimensions, the KBE can be useful as an intermediate step in the derivation of a model that is again practically computable. Here, we will subject the KBE to classical multiple scales asymptotics with respect to the small parameter ε. The derivation is successful if the leading nontrivial contribution to the asymptotic series takes the form of a Kolmogorov backward equation in the slow variable $\boldsymbol{x}$ only, which can be rewritten and simulated as an effective Itô stochastic differential equation.

We start by recalling that for an Itô SDE in the abstract form

$$\mathrm{d}\boldsymbol{x}(t) = \boldsymbol{a}(\boldsymbol{x},t)\,\mathrm{d}t + B(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{W}(t)\,, \tag{117}$$

the corresponding Kolmogorov backward equation reads

$$\frac{\partial p(\boldsymbol{x},t)}{\partial t} = -\sum_{i=1}^{N} a_i(\boldsymbol{x},t)\,\frac{\partial p(\boldsymbol{x},t)}{\partial x_i} - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}(\boldsymbol{x},t)\,\frac{\partial^2 p(\boldsymbol{x},t)}{\partial x_i \partial x_j} \tag{118}$$

with

$$D_{ij}(\boldsymbol{x},t) = \sum_{k=1}^{M} b_{ik}(\boldsymbol{x},t)\,b_{jk}(\boldsymbol{x},t)\,. \tag{119}$$

A derivation and general introduction can be found, for example, in the book by Risken (1996). Applying this notion to our system in the specific form (116), we obtain

$$-\frac{\partial \varrho^\varepsilon}{\partial s} = \frac{1}{\varepsilon^2}\,\mathcal{L}_1\varrho^\varepsilon + \frac{1}{\varepsilon}\,\mathcal{L}_2\varrho^\varepsilon + \mathcal{L}_3\varrho^\varepsilon\,, \tag{120a}$$

where the operators $\mathcal{L}_i$ are given by

$$\mathcal{L}_1 = \sum_j \Big(-\gamma_j \, y_j \, \frac{\partial}{\partial y_j} + \frac{1}{2} \, \sigma_j^2 \, \frac{\partial^2}{\partial y_j^2} \Big) , \tag{121a}$$

$$\mathcal{L}_2 = \sum_{j,k} \Big(L_{jk}^{xy} y_k + \frac{1}{2} \sum_l (2 \, B_{jkl}^{xxy} \, x_k \, y_l + B_{jkl}^{xyy} \, y_k \, y_l) \Big) \frac{\partial}{\partial x_j} + \sum_{j,k} \Big(L_{jk}^{yx} \, x_k + L_{jk}^{yy} \, y_k + \frac{1}{2} \sum_l (B_{jkl}^{xyy} \, x_k \, x_l + 2 \, B_{jkl}^{yyx} \, y_k \, x_l) \Big) \frac{\partial}{\partial y_j} , \tag{121b}$$

$$\mathcal{L}_3 = \sum_j \Big(F_j \sum_k L_{jk}^{xx} \, x_k + \frac{1}{2} \sum_{kl} B_{jkl}^{xxx} \, x_k \, x_l \Big) \frac{\partial}{\partial x_k} . \tag{121c}$$

The adiabatic elimination of fast variables now proceeds by performing an asymptotic expansion of the KBE (120) in powers of ε. We shall sketch only the main steps and refer the reader to the original work by Majda et al. (2001) for full details. To begin with, we expand ϱ^ε as a formal power series, writing

$$\varrho^\varepsilon = \varrho_0 + \varepsilon \, \varrho_1 + \varepsilon^2 \, \varrho_2 + \dots \tag{122}$$

Inserting this expansion into (120) and selecting terms of equal order in ε, we obtain a sequence of equations, the first three of which read

$$\mathcal{L}_1 \varrho_0 = 0 , \tag{123a}$$

$$\mathcal{L}_1 \varrho_1 = -\mathcal{L}_2 \varrho_0 , \tag{123b}$$

$$\mathcal{L}_1 \varrho_2 = -\frac{\partial \varrho_0}{\partial s} - \mathcal{L}_3 \varrho_0 - \mathcal{L}_2 \varrho_1 . \tag{123c}$$

The structure of these equations tells us that we need to find suitable solvability conditions. Equation (123a) implies that ϱ_0 belongs to the null space of $\mathcal{L}_1$, i.e. $\mathbb{P}\varrho_0 = \varrho_0$, where $\mathbb{P}$ is the expectation with respect to the invariant measure of the OU process and $\mathcal{L}_1$ is the generator of the OU process. Hence, (123a) shows that ϱ_0 is independent of the fast variables $\boldsymbol{y}$, thus represents an invariant measure for the fast dynamics in the full SDE. It is easy to see that (123a) and (123b) satisfy the solvability conditions. Now we want to sketch the derivation of dynamic equation for ϱ_0 which can be derived from the solvability condition for equation (123c).

Taking the expectation of (123b), we obtain the following solvability condition

$$\mathbb{P}\mathcal{L}_2 \varrho_0 = \mathbb{P}\mathcal{L}_2 \mathbb{P} \varrho_0 = 0 . \tag{124}$$

If this equation were not be satisfied, the fast modes would induce fast effects on the slow modes. This leads to the solution of (123b):

$$\varrho_1 = -\mathcal{L}_1^{-1} \mathcal{L}_2 \mathbb{P} \varrho_0 . \tag{125}$$

By inserting this expression into (123c) and taking the expectation, we get the solvability condition for ϱ_2:

$$-\frac{\partial \varrho_0}{\partial s} = \mathbb{P}\mathcal{L}_3\mathbb{P}\varrho_0 - \mathbb{P}\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2\mathbb{P}\varrho_0\,. \tag{126}$$

In the limit that $\varepsilon \to 0$, ϱ^ε converges to ϱ_0, the solution of equation (126), see Theorem 4.4 in Majda et al. (2001) and also Kurtz (1973).

The Kolmogorov backward equation (126) can now be transformed back into an effective stochastic differential equation which is computable. The resulting Itô SDE has the form

$$\mathrm{d}\boldsymbol{x} = (\boldsymbol{F} + L\boldsymbol{x} + B(\boldsymbol{x},\boldsymbol{x}) + M(\boldsymbol{x},\boldsymbol{x},\boldsymbol{x}))\,\mathrm{d}t + \sigma_A\,\mathrm{d}\boldsymbol{W}_A + \sigma_W(\boldsymbol{x})\,\mathrm{d}\boldsymbol{W}_M\,, \tag{127}$$

where full expressions for the different terms can be found in Majda et al. (2001). Equation (127) depends only on the slow variables $\boldsymbol{x}$ and approximates the statistics of the slow variables of the full SDE (116). We note that its right hand side contains structurally new terms such as a cubic nonlinearity which generally acts as a damping term but still allows some unstable nonlinear directions (Majda et al., 2009), as well as additive and multiplicative noise terms. The multiplicative noise term arises from the nonlinear interaction between resolved and unresolved modes while the additive noise is the results of nonlinear interactions between unresolved modes and the linear interaction between resolved and unresolved modes, respectively.

5.3 Direct Averaging

It is instructive to illustrate the stochastic mode reduction procedure by working directly on the equations without the use of the KBE for a special case which allows the direct analytic derivation of the effective equations. For this purpose we use the following nonlinear triad stochastic interaction equations (Majda et al., 1999) by using the method of averaging (Kurtz, 1973; Papanicolaou, 1976; Monahan and Culina, 2011):

$$\mathrm{d}x_1(t) = \frac{b_1}{\varepsilon}\,x_2(t)\,y(t)\,\mathrm{d}t \tag{128a}$$

$$\mathrm{d}x_2(t) = \frac{b_2}{\varepsilon}\,x_1(t)\,y(t)\,\mathrm{d}t \tag{128b}$$

$$\mathrm{d}y(t) = \frac{b_3}{\varepsilon}\,x_1(t)\,x_2(t)\,\mathrm{d}t - \frac{\gamma}{\varepsilon^2}\,y(t)\,\mathrm{d}t + \frac{\sigma}{\varepsilon}\,\mathrm{d}W(t) \tag{128c}$$

Now we formally solve the equation for y:

$$y(t) = \mathrm{e}^{-\frac{\gamma t}{\varepsilon^2}}\,y + \frac{b_3}{\varepsilon}\int_0^t \mathrm{e}^{-\frac{\gamma(t-s)}{\varepsilon^2}}\,x_1(s)\,x_2(s)\,\mathrm{d}s + h(t) \tag{129}$$

where

$$h(t) = \frac{\sigma}{\varepsilon} \int_0^t \mathrm{e}^{-\frac{\gamma(t-s)}{\varepsilon^2}} \, \mathrm{d}W(s) \tag{130}$$

We note that (129) contains a time integral which can be interpreted as a memory kernel (Gottwald et al., 2017). The existence of a memory kernel is already predicted in the Mori–Zwanzig formalism (Mori, 1965; Zwanzig, 1973, 2001; Wouters and Lucarini, 2013; Chorin et al., 2000).

However, in the case of time scale separation so that $\varepsilon \to 0$, the above term becomes Markovian (Majda et al., 2001) by performing integration by parts:

$$y(t) \to \varepsilon \frac{b_3}{\gamma} \, x_1(t) \, x_2(t) + h(t) \tag{131}$$

and

$$h(t) \to \varepsilon \frac{\sigma}{\gamma} \, \mathrm{d}W(t) \tag{132}$$

If we now insert equations (131) and (132) into (128a) and (128b), we obtain

$$\mathrm{d}x_1(t) = \frac{b_1 b_3}{\gamma} \, x_1(t) \, x_2^2(t) \, \mathrm{d}t + \frac{b_1 \sigma}{\gamma} \, x_2(t) \circ \mathrm{d}W(t) \,, \tag{133a}$$

$$\mathrm{d}x_2(t) = \frac{b_2 b_3}{\gamma} \, x_1^2(t) \, x_2(t) \, \mathrm{d}t + \frac{b_2 \sigma}{\gamma} \, x_1(t) \circ \mathrm{d}W(t) \,, \tag{133b}$$

which must to be interpreted in the Stratonovich sense as indicated by the $\circ$ in front of the Wiener process. The Stratonovich interpretation arises because the integration performed in (130) introduces non-negligible autocorrelation. However, we can transform the Stratonovich SDE into Itô form, picking up an additional noise-induced drift term:

$$\mathrm{d}x_1(t) = \frac{b_1 b_2 \, \sigma^2}{2\gamma^2} \, x_1(t) \, \mathrm{d}t + \frac{b_1 \, b_3}{\gamma} \, x_1(t) \, x_2^2(t) \, \mathrm{d}t + \frac{b_1 \, \sigma}{\gamma} \, x_2(t) \, \mathrm{d}W(t) \,, \tag{134}$$

$$\mathrm{d}x_2(t) = \frac{b_1 b_2 \, \sigma^2}{2\gamma^2} \, x_2(t) \, \mathrm{d}t + \frac{b_2 \, b_3}{\gamma} \, x_1^2(t) \, x_2(t) \, \mathrm{d}t + \frac{b_2 \, \sigma}{\gamma} \, x_1(t) \, \mathrm{d}W(t) \,. \tag{135}$$

This reduced model has cubic nonlinearity and multiplicative noise. Future research will explore the effects of the memory kernel since in the climate system we do not have any obvious time scale separations.

6 Outlook

In this chapter we provided a brief overview of some basic GFD issues and also discussed some current research themes like variational balance models and stochastic mode reduction. In our future research we plan to follow these three broad research directions: (i) specify the limits of validity of asymptotic regimes and the interaction of different scale regimes, (ii) examine how dynamic properties such as stability, bifurcating nonlinear waves, and nonlinear interactions out of linear waves persist across model hierarchies, (iii) include effective dissipation in scaling analysis and the role of eddy dissipation, (iv) study model hierarchies to develop parameterizations for processes not represented at a coarse level. Here we will use the separation of balanced motion from the meso-scales, where our aim is the develop new energy and momentum consistent stochastic parameterizations of the unresolved processes.

Acknowledgements We thank Sergey Danilov for helpful advice on several aspects of this paper, and Frederica Gugole for careful proof reading of the manuscript. We received funding from the Collaborative Research Centre TRR 181. CF was supported by the German Research Foundation (DFG) through the Cluster of Excellence CliSAP (EXC177) at the University of Hamburg and DFG grant FR3515/3-1. GB was partially supported by the DFG grant FOR1740.

References

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

Babin, A., Mahalov, A., and Nicolaenko, B. (1996). Global splitting, integrability and regularity of 3D Euler and Navier-Stokes equations for uniformly rotating fluids. *E. J. Mech. B Fluids*, 15(3):291–300.

Babin, A., Mahalov, A., and Nicolaenko, B. (1997). Global splitting and regularity of rotating shallow-water equations. *E. J. Mech. B Fluids*, 16(5):725–754.

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

Badin, G. (2014). On the role of non-uniform stratification and short-wave instabilities in three-layer quasi-geostrophic turbulence. *Phys. Fluids*, 26(9):096603.

Balmforth, N. J., Morrison, P. J., and Thiffeault, J.-L. (2013). Pattern formation in Hamiltonian systems with continuous spectra; a normal-form single-wave model. *arXiv preprint*, arXiv:1303.0065.

Bates, M. L. and Grimshaw, R. H. J. (2014). An extended equatorial plane: linear spectrum and resonant triads. *Geophys. Astrophys. Fluid Dyn.*, 108(1):1–19.

Beck, M., Sandstede, B., and Zumbrun, K. (2010). Nonlinear stability of time-periodic viscous shocks. *Arch. Ration. Mech. Anal.*, 196(3):1011–1076.

Becker, E. (2003). Frictional heating in global climate models. *Mon. Wea. Rev.*, 131:508–520.

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

Blender, R. and Badin, G. (2015). Hydrodynamic Nambu brackets derived by geometric constraints. *J. Phys. A: Math. Theor.*, 48(10):105501.

Blumen, W. and Wu, R. (1995). Geostrophic adjustment: Frontogenesis and energy conversion. *J. Phys. Oceanogr.*, 25(3):428–438.

Bokhove, O. and Oliver, M. (2006). Parcel Eulerian–Lagrangian fluid dynamics of rotating geophysical flows. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 462:2563–2573.

Bokhove, O., Vanneste, J., and Warn, J. (1998). A variational formulation for barotropic quasi-geostrophic flows. *Geophys. Astrophys. Fluid Dyn.*, 88(1-4):67–79.

Bouchut, F., Le Sommer, J., and Zeitlin, V. (2005). Breaking of balanced and unbalanced equatorial waves. *Chaos*, 15(1):013503.

Boyd, J. P. (1980). Equatorial solitary waves. Part I: Rossby solitons. *J. Phys. Oceano.*, 10(11):1699–1717.

Boyd, J. P. (2002). Equatorial solitary waves. Part V: Initial value experiments, coexisting branches, and tilted-pair instability. *J. Phys. Oceanogr.*, 32(9):2589–2602.

Boyd, J. P. and Zhou, C. (2008). Kelvin waves in the nonlinear shallow water equations on the sphere: nonlinear travelling waves and the corner wave bifurcation. *J. Fluid Mech.*, 617:187–205.

Burkhardt, U. and Becker, E. (2006). A consistent diffusion-dissipation parameterization in the ECHAM climate model. *Mon. Wea. Rev.*, 134:1194–1204.

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

Callaghan, T. G. and Forbes, L. K. (2006). Computing large-amplitude progressive Rossby waves on a sphere. *J. Comput. Phys.*, 217(2):845–865.

Cao, C., Li, J., and Titi, E. S. (2016). Global well-posedness of the three-dimensional primitive equations with only horizontal viscosity and diffusion. *Comm. Pure Appl. Math.*, 69(8):1492–1531.

Chan, I. H. and Shepherd, T. G. (2013). Balance model for equatorial long waves. *J. Fluid Mech.*, 725:55–90.

Chatterjee, R. (1996). Dynamical Symmetries and Nambu Mechanics. *Letters Math. Phys.*, 36:117–126.

Cheng, B. and Mahalov, A. (2013). Time-averages of fast oscillatory systems. *Discrete Contin. Dyn. Syst. Ser. S*, 6(5):1151–1162.

Chorin, A. J., Hald, O. H., and Kupferman, R. (2000). Optimal prediction and the Mori–Zwanzig representation of irreversible processes. *Proc. Nat. Acad. Sci.*, 97(7):2968–2973.

Constantin, A. (2013). Some three-dimensional nonlinear equatorial flows. *J. Phys. Oceanogr.*, 43(1):165–175.

Crommelin, D. and Majda, A. (2004). Strategies for model reduction: comparing different optimal bases. *J. Atmos. Sci.*, 61(17):2206–2217.

Cullen, M. J. P. (2008). A comparison of numerical solutions to the Eady frontogenesis problem. *Quart. J. R. Meteorol. Soc.*, 134(637):2143–2155.

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

Dalibard, A.-L. and Saint-Raymond, L. (2010). Mathematical study of the β-plane model for rotating fluids in a thin layer. *J. Math. Pures Appl. (9)*, 94(2):131–169.

Dijkstra, H. A., Wubs, F. W., Cliffe, A. K., Doedel, E., Dragomirescu, I. F., Eckhardt, B., Gelfgat, A. Y., Hazel, A. L., Lucarini, V., Salinger, A. G., et al. (2014). Numerical bifurcation methods and their application to fluid dynamics: analysis beyond simulation. *Commun. Comput. Phys*, 15(1):1–45.

Doelman, A., Sandstede, B., Scheel, A., and Schneider, G. (2009). The dynamics of modulated wave trains. *Mem. Amer. Math. Soc.*, 199(934):viii+105.

Dolaptchiev, S. I. and Klein, R. (2013). A multiscale model for the planetary and synoptic motions in the atmosphere. *J. Atmos. Sci.*, 70:2963–2981.

Dolaptchiev, S. I., Timofeyev, I., and Achatz, U. (2013). Subgrid-scale closure for the inviscid Burgers–Hopf equation. *Commun. Math. Sci.*, 11:757–777.

Dritschel, D. G., Gottwald, G. A., and Oliver, M. (2016). Comparison of variational balance models for the rotating shallow water equations. J. Fluid Mech., to appear.

Dritschel, D. G. and Viúdez, A. (2003). A balanced approach to modelling rotating stably stratified geophysical flows. *J. Fluid Mech.*, 488:123–150.

Dutrifoy, A., Majda, A. J., and Schochet, S. (2009). A simple justification of the singular limit for equatorial shallow-water dynamics. *Comm. Pure Appl. Math.*, 62(3):322–333.

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

Emanuel, K. A. (1979). Inertial instability and mesoscale convective systems. Part I: Linear theory of inertial instability in rotating viscous fluids. *J. Atmos. Sci.*, 36(12):2425–2449.

Emanuel, K. A. (1984). Comments on "inertial instability and mesoscale convective systems. Part I: Linear theory of inertial instability in rotating viscous fluids". *J. Atmos. Sci.*, 42(7):747–752.

Embid, P. F. and Majda, A. J. (1996). Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity. *Comm. Part. Diff. Eq.*, 21(3–4):619–658.

Ford, R., McIntyre, M. E., and Norton, W. A. (2000). Balance and the slow quasi-manifold: Some explicit results. *J. Atmos. Sci.*, 57(9):1236–1254.

Franzke, C. and Majda, A. J. (2006). Low-order stochastic mode reduction for a prototype atmospheric GCM. *J. Atmos. Sci.*, 63:457–479.

Franzke, C., Majda, A. J., and Vanden-Eijnden, E. (2005). Low-order stochastic mode reduction for a realistic barotropic model climate. *J. Atmos. Sci.*, 62:1722–1745.

Franzke, C., O'Kane, T., Berner, J., Williams, P., and Lucarini, V. (2015). Stochastic climate theory and modelling. *WIREs Climate Change*, 6:63–78.

Franzke, C. L. E. (2013). Predictions of critical transitions with non-stationary reduced order models. *Phys. D*, 262:35–47.

Fringer, O. B. (2009). Towards nonhydrostatic ocean modeling with large-eddy simulation. In Glickson, D., editor, *Oceanography in 2025*, pages 81–83. Natl. Academies Press.

Gardiner, C. W. (2009). *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer.

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

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

Gill, A. E. (1982). *Atmosphere-Ocean Dynamics*. Academic Press.

Givon, D., Kupferman, R., and Stuart, A. (2004). Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55.

Gottwald, G., Crommelin, D., and Franzke, C. (2017). Stochastic climate theory. In Franzke, C. and O'Kane, T., editors, *Nonlinear and Stochastic Climate Dynamics*, pages 209–240. Cambridge University Press.

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

Grooms, I., Julien, K., and Fox-Kemper, B. (2011). On the interactions between planetary geostrophy and mesoscale eddies. *Dynam. Atmos. Oceans*, 51(3):109–136.

Haragus, M. and Iooss, G. (2011). *Local bifurcations, center manifolds, and normal forms in infinite-dimensional dynamical systems*. Springer.

Hasselmann, K. (1976). Stochastic climate models. Part I. Theory. *Tellus*, 28(6):473–485.

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 471(2176):20140963, 19.

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

Holm, D. D., Schmah, T., and Stoica, C. (2009). *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford University Press.

Holm, D. D. and Zeitlin, V. (1998). Hamilton's principle for quasigeostrophic motion. *Phys. Fluids*, 10(4):800–806.

Holton, J. R. and Hakim, G. J. (2012). *An Introduction To Dynamic Meteorology*. Academic Press.

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

Hsu, H.-C. (2014). An exact solution for nonlinear internal equatorial waves in the f-plane approximation. *J. Math. Fluid Mech.*, 16(3):463–471.

Ibragimov, R. N. (2011). Nonlinear viscous fluid patterns in a thin rotating spherical domain and applications. *Phys. Fluids*, 23(12):123102.

Julien, K. and Knobloch, E. (2007). Reduced models for fluid flows with strong constraints. *J. Math. Phys.*, 48(6).

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

Kamenkovich, V. M., Koshlyakov, M. N., and Monin, A. S. (1986). *Synoptic Eddies in the Ocean*. D. Reidel Publishing Company.

Kapitula, T. and Promislow, K. (2013). *Spectral and Dynamical Stability of Nonlinear Waves*. Springer.

Kasahara, A. and Gary, J. M. (2010). Studies of inertio-gravity waves on midlatitude beta-plane without the traditional approximation. *Quart. J. R. Meteorol. Soc.*, 136(647):517–536.

Kawahara, G., Uhlmann, M., and van Veen, L. (2012). The significance of simple invariant solutions in turbulent flows. *Ann. Rev. Fluid Mech.*, 44(1):203–225.

Khouider, B., Majda, A. J., and Stechmann, S. N. (2013). Climate science in the tropics: waves, vortices and PDEs. *Nonlinearity*, 26(1):R1.

Klein, R. (2010). Scale-dependent models for atmospheric flows. *Ann. Rev. Fluid Mech.*, 42:249–274.

Klingbeil, K. and Burchard, H. (2013). Implementation of a direct nonhydrostatic pressure gradient discretisation into a layered ocean model. *Ocean Model.*, 65:64–77.

Koide, T. and Kodama, T. (2012). Navier–Stokes, Gross–Pitaevskii and generalized diffusion equations using the stochastic variational method. *J. Phys. A*, 45(25):255204, 18.

Kurtz, T. G. (1973). A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Func. Anal.*, 12(1):55–67.

Kwasniok, F. (2004). Empirical low-order models of barotropic flow. *J. Atmos. Sci.*, 61(2):235–245.

Majda, A. (2003). *Introduction to PDEs and Waves for the Atmosphere and Ocean*. American Mathematical Society.

Majda, A., Franzke, C., and Crommelin, D. (2009). Normal forms for reduced stochastic climate models. *Proc. Natl. Acad. Sci. USA*, 106:3649–3653.

Majda, A., Timofeyev, I., and Vanden-Eijnden, E. (2002). A priori tests of a stochastic mode reduction strategy. *Phys. D*, 170:206–252.

Majda, A. J., Franzke, C., and Khouider, B. (2008). An applied mathematics perspective on stochastic modelling for climate. *Phil. Trans. R. Soc. A*, 366:2429–2455.

Majda, A. J. and Klein, R. (2003). Systematic multiscale models for the tropics. *J. Atmos. Sci.*, 60(2):393–408.

Majda, A. J., Timofeyev, I., and Vanden-Eijnden, E. (1999). Models for stochastic climate prediction. *Proc. Nat. Acad. Sci. USA*, 96(26):14687–14691.

Majda, A. J., Timofeyev, I., and Vanden-Eijnden, E. (2001). A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math.*, 54(8):891–974.

Majda, A. J., Timofeyev, I., and Vanden-Eijnden, E. (2003). Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.*, 60(14):1705–1722.

McIntyre, M. (2015). Dynamical meteorology – balanced flow. In Pyle, J. and Zhang, F., editors, *Encyclopedia of Atmospheric Sciences*, pages 298–303. Academic Press, Oxford, second edition.

McIntyre, M. E. and Norton, W. A. (2000). Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, 57(9):1214–1235.

McWilliams, J. C. (1977). A note on a consistent quasigeostrophic model in a multiply connected domain. *Dynam. Atmos. Oceans*, 1(5):427–441.

Meyries, M., Rademacher, J., and Siero, E. (2014). Quasilinear parabolic reaction-diffusion systems: User's guide to well-posedness, spectra and stability of travelling waves. *SIAM J. Appl. Dyn. Sys.*, 13:249–275.

Mohebalhojeh, A. R. and Dritschel, D. G. (2001). Hierarchies of balance conditions for the f-plane shallow-water equations. *J. Atmos. Sci.*, 58(16):2411–2426.

Monahan, A. H. and Culina, J. (2011). Stochastic averaging of idealized climate models. *J. Climate*, 24(12):3068–3088.

Moon, W. and Wettlaufer, J. S. (2014). On the interpretation of Stratonovich calculus. *New J. Phys.*, 16(5):055017.

Mori, H. (1965). Transport, collective motion, and Brownian motion. *Progress Theor. Phys.*, 33(3):423–455.

Nambu, Y. (1973). Generalized Hamiltonian dynamics. *Phys. Rev. D*, 7(8):2405.

Olbers, D. and Eden, C. (2013). A global model for the diapycnal diffusivity induced by internal gravity waves. *J. Phys. Oceanogr.*, 43(8):1759–1779.

Olbers, D., Willebrand, J., and Eden, C. (2012). *Ocean Dynamics*. Springer.

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

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

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

Palmer, T., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G., Steinheimer, M., and Weisheimer, A. (2009). Stochastic parametrization and model uncertainty. Technical report, ECMWF.

Papanicolaou, G. C. (1976). Some probabilistic problems and methods in singular perturbations. *Rocky Mountain J. Math.*, 6(4):653–674. Summer Research Conference on Singular Perturbations: Theory and Applications (Northern Arizona Univ., Flagstaff, Ariz., 1975).

Pavliotis, G. A. and Stuart, A. (2008). *Multiscale Methods: Averaging and Homogenization*. Springer.

Pedlosky, J. (1987). *Geophysical Fluid Dynamics*. Springer, second edition.

Plougonven, R. and Zeitlin, V. (2005). Lagrangian approach to geostrophic adjustment of frontal anomalies in a stratified fluid. *Geophys. Astrophys. Fluid Dyn.*, 99(2):101–135.

Ragone, F. and Badin, G. (2016). A study of surface semi-geostrophic turbulence: freely decaying dynamics. *J. Fluid Mech.*, 792:740–774.

Reznik, G. M. (2015). Wave adjustment: general concept and examples. *J. Fluid Mech.*, 779:514–543.

Risken, H. (1996). *The Fokker–Planck Equation*. Springer.

Sagaut, P. (2006). *Large Eddy Simulation for Incompressible Flows: An Introduction*. Springer.

Saint-Raymond, L. (2010). Lecture notes: Mathematical study of singular perturbation problems. Applications to large-scale oceanography. *Journées Eq. Deriv. Part.*, pages 1–49.

Saito, K., Ishida, J., Aranami, K., Hara, T., Segawa, T., Narita, M., and Honda, Y. (2007). Nonhydrostatic atmospheric models and operational development at JMA. *J. Meteor. Soc. Japan*, 85B:271–304.

Salmon, R. (1982). The shape of the main thermocline. *J. Phys. Oceanogr.*, 12:1458–1479.

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

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

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

Salmon, R. (1998). *Lectures on Geophysical Fluid Dynamics*. Oxford University Press.

Salmon, R. (2005). A general method for conserving quantities related to potential vorticity in numerical models. *Nonlinearity*, 18(5):R1.

Salmon, R. (2007). A general method for conserving energy and potential enstrophy in shallow water models. *J. Atmos. Sci.*, 64:515–531.

Sandstede, B. (2002). Stability of travelling waves. In *Handbook of Dynamical Systems*, volume 2, pages 983–1055. North-Holland, Amsterdam.

Schaefer-Rolffs, U. and Becker, E. (2013). Horizontal momentum diffusion in GCMs using the dynamic Smagorinsky model. *Mon. Weather Rev.*, 141(3):887–899.

Schaefer-Rolffs, U., Knöpfel, R., and Becker, E. (2015). A scale invariance criterion for LES parametrizations. *Meteorol. Z.*, 24(1):3–13.

Schneider, G. (1994). Error estimates for the Ginzburg–Landau approximation. *Z. Angew. Math. Phys.*, 45(3):433–457.

Schubert, W. H., Taft, R. K., and Silvers, L. G. (2009). Shallow water quasi-geostrophic theory on the sphere. *J. Adv. Model. Earth Syst.*, 1(2). 2.

Shaw, T. A. and Shepherd, T. G. (2009). A theoretical framework for energy and momentum consistency in subgrid-scale parameterization for climate models. *J. Atmos. Sci.*, 66:3095–3114.

Shepherd, T. G. (1990). Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys.*, 32:287–338.

Simonnet, E., Dijkstra, H. A., and Ghil, M. (2009). Bifurcation analysis of ocean, atmosphere, and climate models. In Ciarlet, P., editor, *Handbook of Numerical Analysis*, volume 14, pages 187–229. Elsevier.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Mon. Weather Rev.*, 91(3):99–164.

Smith, R. K. and Dritschel, D. G. (2006). Revisiting the Rossby–Haurwitz wave test case with contour advection. *J. Comput. Phys.*, 217(2):473–484.

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

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

Stinis, P. (2006). A comparative study of two stochastic mode reduction methods. *Phys. D*, 213(2):197–213.

Takhtajan, L. (1994). On foundation of the generalized Nambu mechanics. *Comm. Math. Phys.*, 160(2):295–315.

Temam, R. and Wang, S. H. (1993). Inertial forms of Navier–Stokes equations on the sphere. *J. Funct. Anal.*, 117(1):215–242.

Temam, R. and Wirosoetisno, D. (2010). Slow manifolds and invariant sets of the primitive equations. *J. Atmos. Sci.*, 68(3):675–682.

Theiss, J. and Mohebalhojeh, A. R. (2009). The equatorial counterpart of the quasi-geostrophic model. *J. Fluid Mech.*, 637:327–356.

Thuburn, J. and Li, Y. (2000). Numerical simulations of Rossby–Haurwitz waves. *Tellus A*, 52(2):181–189.

Tort, M. and Dubos, T. (2014). Usual approximations to the equations of atmospheric motion: A variational perspective. *J. Atmos. Sci.*, 71(7):2452–2466.

Tort, M., Dubos, T., Bouchut, F., and Zeitlin, V. (2014). Consistent shallow-water equations on the rotating sphere with complete Coriolis force and topography. *J. Fluid Mech.*, 748:789–821.

Tort, M., Ribstein, B., and Zeitlin, V. (2016). Symmetric and asymmetric inertial instability of zonal jets on the f-plane with complete Coriolis force. *J. Fluid Mech.*, 788:274–302.

Trias, F. X., Folch, D., Gorobets, A., and Oliva, A. (2015). Building proper invariants for eddy-viscosity subgrid-scale models. *Phys. Fluids*, 27(6).

Vallis, G. K. (2006). *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. Cambridge University Press.

Vanderbauwhede, A. (2012). Lyapunov–Schmidt method for dynamical systems. In *Mathematics of complexity and dynamical systems. Vols. 1–3*, pages 937–952. Springer, New York.

Vanneste, J. (2013). Balance and spontaneous wave generation in geophysical flows. *Ann. Rev. Fluid Mech.*, 45(1):147–172.

Verkley, W. and van der Velde, I. (2010). Balanced dynamics in the tropics. *Quart. J. R. Meteorol. Soc.*, 136(646):41–49.

Warn, T., Bokhove, O., Shepherd, T., and Vallis, G. (1995). Rossby number expansions, slaving principles, and balance dynamics. *Quart. J. R. Meteorol. Soc.*, 121(523):723–739.

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

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

Wouters, J. and Lucarini, V. (2013). Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori–Zwanzig approach. *J. Stat. Phys.*, 151(5):850–860.

Zeitlin, V., Medvedev, S. B., and Plougonven, R. (2003). Frontal geostrophic adjustment, slow manifold and nonlinear wave phenomena in one-dimensional rotating shallow water. Part 1. Theory. *J. Fluid Mech.*, 481:269–290.

Zwanzig, R. (1973). Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9(3):215–220.

Zwanzig, R. (2001). *Nonequilibrium Statistical Mechanics*. Oxford University Press.