J. Fluid Mech. (2006), *vol.* 551, *pp.* 197–234. © 2006 Cambridge University Press
doi:10.1017/S0022112005008256 Printed in the United Kingdom

Variational asymptotics for rotating shallow water near geostrophy: a transformational approach

By MARCEL OLIVER

School of Engineering and Science, International University Bremen, 28759 Bremen, Germany
oliver@member.ams.org

(Received 17 March 2005 and in revised form 5 September 2005)

We introduce a unified variational framework in which the classical balance models for nearly geostrophic shallow water as well as several new models can be derived. Our approach is based on consistently truncating an asymptotic expansion of a near-identity transformation of the rotating shallow-water Lagrangian. Model reduction is achieved by imposing either degeneracy (for models in a semi-geostrophic scaling) or incompressibility (for models in a quasi-geostrophic scaling) with respect to the new coordinates.

At first order, we recover the classical semi-geostrophic and quasi-geostrophic equations, Salmon's L_1 and large-scale semi-geostrophic equations, as well as a one-parameter family of models that interpolate between the two. We identify one member of this family, different from previously known models, that promises better regularity – hence consistency with large-scale vortical motion – than all other first-order models. Moreover, we explicitly derive second-order models for all cases considered. While these second-order models involve nonlinear potential vorticity inversion and do not obviously share the good properties or their first-order counterparts, we offer an explicit survey of second-order models and point out several avenues for exploration.

1. Introduction

In a series of articles, Salmon proposed new approximate models for nearly geostrophic flow in a layer of shallow water (1983, 1985), and in a layer of stratified fluid of finite depth (1996). The derivation is an example of *variational asymptotics*: all approximations are performed on the Lagrangian of the parent fluid model before Hamilton's principle is applied to yield new equations of motion. One of the chief advantages of this approach is that preservation of time and particle relabelling symmetries guarantees exact conservation of a new energy and potential vorticity in the approximate system.

Salmon's approximation consists of two steps. First, noting that the stationary leading-order geostrophic balance defines a submanifold in phase space, he constrains the full Lagrangian to this 'slow manifold'. The symplectic structure of the constrained variational principle is typically non-canonical. Salmon therefore suggests applying, in a second step, a near-identity transformation to simpler, possibly canonical, coordinates. Although a transformation to canonical coordinates must exist, an explicit expression can only be given to some order in the Rossby number ε, the formal small parameter; higher-order terms are consistently dropped.

While built-in structure preservation is clearly an attractive feature, it does not guarantee even well-posedness of the resulting balance models. In fact, the large-scale semi-geostrophic (LSG) equations of Salmon (1985), as well as their generalizations to stratified flows, turn out to be ill posed (Shepherd & Ford 2001; Ford, personal communication 2000). In the hope of turning the LSG equations into a well-behaved model without losing their simple structure, we noted that the second-order generalization of LSG possesses a positive definite Hamiltonian – clearly a desirable feature, but insufficient to guarantee existence of a flow. Regularity of potential vorticity inversion is equally crucial, but explicitly violated in LSG.

The new idea presented here is that the two steps in Salmon's procedure – constraining and transforming – can be reversed in order. We will start out with an arbitrary change of coordinates which reduces to the identity when the perturbation parameter ε vanishes. Both the transformation and the Lagrangian of the parent shallow-water equations can thus be expanded in powers of ε and consistently truncated at the desired order of accuracy. At this point, the transformation is completely arbitrary, so that we can impose, order by order, conditions on the transformation that assure that the system is constrained to a submanifold in phase space, or that the correct leading-order balance is maintained. The advantages are threefold. (i) We can systematically identify degrees of freedom that leave structure and formal order of the reduced model invariant, but can be tuned to optimize desirable features such as the regularity of the potential vorticity inversion. (ii) We have a procedure that allows us, at least in principle, to develop higher-order models in a systematic fashion. (iii) We can study balance models in a unified framework that includes all the classical balance models for rotating shallow water: the semi-geostrophic and quasi-geostrophic equations, Salmon's L_1 and LSG models, and many new ones.

In the last two decades, a large number of authors have explored the variational route to deriving or analysing balance models for rotating fluids. Allen & Holm (1996) derive a class of balance models by imposing second-order constraints on the variational principle. The authors also note the distinct role of affine Lagrangians very explicitly. Their work differs from ours in that they treat the approximation of the symplectic structure and of the Hamiltonian as independent. Our point of view is that the concept of consistently truncating a change of coordinates provides a rigid dependence between the respective approximations; in other words, we supply a systematic way of deriving dependences between some of Allen & Holm's free parameters. Holm & Zeitlin (1998) introduce the variational formulation for the quasi-geostrophic equations; independently, Bokhove, Vanneste & Warn (1998) give a derivation of the quasi-geostrophic equations via a constrained expansion of the shallow-water variational principle. McIntyre & Roulstone (2002) review and systematically explain the structure of models based on workless momentum–configuration constraints, and suggest several generalizations of classical semi-geostrophic theory. Using the language of 'velocity splits' coined by McIntyre & Roulstone, Wunderer (2001) and G. Roullet (personal communication 2004) generalized Salmon's L_1 equations to second order. Roullet's L_2 equations, being non-local in time, clearly differ from ours which do not have non-local terms. The relative merits of the two approaches are currently not well understood and remain to be explored. Finally, Vanneste & Bokhove (2002) show how to translate Salmon's variational asymptotics into asymptotics on the corresponding Poisson structure, and also suggest a generalization to higher order.

The present paper is laid out as follows. Section 2 reviews the two most important models for rotating shallow water, the semi-geostrophic and the quasi-geostrophic equations. In § 3, we explain our new approach for a finite-dimensional linear toy problem. In this simple situation, we have the opportunity to compare the reduced model with explicitly computed solutions of the parent dynamics. Section 4 introduces the Lagrangian formalism for fluids with particular emphasis on affine and incompressible fluid Lagrangians, which will play a major role as target Lagrangians leading to model reduction in the semi-geostrophic and the quasi-geostrophic scaling, respectively. We discuss asymptotics in the variational principle as a means of deriving reduced models, and give a brief derivation of Salmon's L_1 and LSG models within this general framework.

The main part of the paper is the derivation of the following three distinct model hierarchies.

The *LSG hierarchy* includes Salmon's L_1 and LSG equations at first order, as well as a one-parameter family of models interpolating between the two. It is characterized by the condition that the reduced Lagrangian is affine, i.e. linear in the velocities. This implies that the resulting equation of 'motion' does not include time derivatives of the velocity field $\boldsymbol{u}$ – it defines a kinematic relationship between $\boldsymbol{u}$ and the mass configuration h. Dynamics enters via the continuity equation or, equivalently, via the advection of potential vorticity. Since the reduced Lagrangian is always degenerate, a Dirac constraint is implied by construction. Finally, time derivatives of $\boldsymbol{u}$ generally enter when transforming back to physical coordinates although they are absent from the equations of motion to any order. Section 5 details the derivation of the LSG hierarchy for the rotating shallow-water equations.

The *quasi-geostrophic hierarchy*, introduced in § 6, yields the classical quasi-geostrophic equations at first order. It is characterized by the condition that the transformed dynamics be incompressible up to the required order. The resulting reduced Lagrangian is always a regular incompressible fluid Lagrangian. Hence, the dynamics resides in the momentum equation, while the continuity equation reduces to the zero divergence condition. In physical coordinates, of course, weak compressibility is recovered.

Finally, § 7 recovers the classical semi-geostrophic equations as the first order of the *semi-geostrophic hierarchy*. While the scaling is the same as for the LSG hierarchy, the conditions we impose are subtly different. We require that the new coordinates are canonical, and the velocity in new coordinates is equal to the geostrophic velocity in old coordinates. Our transform generalizes the Hoskins (1975) transform at order two and higher.

In each case, we explicitly compute to second order. At first order, only the LSG hierarchy yields something new: a model, in a certain sense half-way between L_1 and LSG dynamics, that promises superior regularity properties relative to all other first-order models. The first-order computations in the remaining two cases yield well-known models. However, our approach still provides a constructive derivation for the variational formulation of quasigeostrophy, and we obtain an interpretation of the geostrophic momentum approximation as a truncated near-identity change of coordinates.

At second order, we derive the corresponding models of each hierarchy; in the case of the LSG approach there is a five-parameter family of models, while the other two hierarchies are unique at second order as well. Except for trivial examples in the LSG hierarchy, all second-order models require nonlinear and apparently non-elliptic potential vorticity inversion. Therefore, well-posedness and numerical implementation are not obvious, and we mainly point out the questions that need to be asked. Thus, with regard to second-order models, this paper raises more questions than it answers.

In the final discussion, § 8, we point out possible approaches to second-order models and other extensions of our ideas.

2. The classical nearly geostrophic limits

2.1. *Distinguished scaling limits*

We first sketch the two main distinguished scaling limits of the rotating shallow-water equations, the semi-geostrophic and the quasi-geostrophic equations.

We take the simplest possible non-trivial case – the rotating shallow-water equations with constant Coriolis parameter on the plane. In this model, which we regard as the standard against which the accuracy of all other models must be judged, the evolution of the horizontal velocity $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t)$ and fluid depth $h = h(\boldsymbol{x}, t)$ is governed by

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + f \boldsymbol{u}^\perp + g \nabla h = 0, \tag{2.1a}$$

$$\partial_t h + \nabla \cdot (h \boldsymbol{u}) = 0, \tag{2.1b}$$

where $\boldsymbol{u}^\perp = (-u_2, u_1)$, f is the Coriolis parameter, and g the constant of gravity. We assume that h approaches a constant, and $\boldsymbol{u}$ vanishes at infinity. In all of the following, we take f to be constant, though the fundamental ideas extend to the general case.

We first non-dimensionalize the shallow-water equations. Let U be the horizontal velocity scale, L the horizontal geometric length scale, and H the mean layer depth. Throughout, we take the advective time scale $T = L/U$ and assume that the *Rossby number* ε is small, i.e.

$$\varepsilon = \frac{U}{fL} \ll 1. \tag{2.2}$$

We also define the *Burger number*

$$B = \frac{gH}{f^2 L^2}. \tag{2.3}$$

The shallow-water equations in non-dimensionalized variables then read

$$\varepsilon(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u}) + \boldsymbol{u}^\perp + \frac{B}{\varepsilon} \nabla h = 0, \tag{2.4a}$$

$$\partial_t h + \nabla \cdot (h \boldsymbol{u}) = 0. \tag{2.4b}$$

We are interested in the physical regime where the pressure gradient force balances the Coriolis force to leading order. In other words, we seek a leading-order *geostrophic balance* relation of the form

$$\boldsymbol{u}_{\mathrm{G}} = \nabla^\perp h. \tag{2.5}$$

The resulting geostrophic motion is stationary, as can be checked by substituting (2.5) back into the continuity equation (2.4*b*).

There are two distinguished scaling limits that result in leading-order geostrophic balance. If we admit order one variations in the total depth, balance requires that $B = \varepsilon$. This is called the *semi-geostrophic scaling*. On the other hand, we can allow for a Burger number of order one if the total depth is an $O(\varepsilon)$ variation of a constant mean depth. Thus, in this so-called *quasi-geostrophic scaling* we keep $B = 1$ and

$$h = 1 + \varepsilon h_1, \tag{2.6}$$

so that $\nabla h = O(\varepsilon)$.

Before proceeding further, we set up notation that is crucial later, but already useful now. We then present traditional derivations of the next order corrections to geostrophic balance in each of the two scalings.

2.2. *Notation*

Throughout this paper, we adapt conventions that are less used in the geophysical literature, but have proved – conceptionally as well as regarding the ease of symbolic manipulation – extremely useful. We generally view velocities as vector fields and transformations as diffeomorphisms of the plane, avoiding explicitly working in coordinates whenever possible. Most of the following could easily be written in geometrically intrinsic notation; this, however, is not the point here.

First, we employ fixed-slot notation, always stating changes of variables explicitly. Thus, if $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t)$ denotes the Eulerian velocity of a fluid, and $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{a}, t)$ the corresponding flow map – the fluid particle initially at location $\boldsymbol{a}$ is at location $\boldsymbol{x} = \boldsymbol{\eta}(\boldsymbol{a}, t)$ at time t – then the Lagrangian velocity of this fluid particle must be

$$\partial_t \boldsymbol{\eta}(\boldsymbol{a}, t) = \boldsymbol{u}(\boldsymbol{\eta}(\boldsymbol{a}, t), t), \tag{2.7}$$

which we abbreviate, throughout, by

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}. \tag{2.8}$$

In this notation, the continuity equation (2.4*b*) is equivalent to

$$h \circ \boldsymbol{\eta} = \frac{1}{\det \nabla \boldsymbol{\eta}}, \tag{2.9}$$

where derivatives are always taken with respect to the natural arguments.

Secondly, throughout this paper, we will encounter near-identity changes of coordinates which are eventually expanded in the perturbation parameter ε. When such a transformation $\boldsymbol{\xi}_\varepsilon$ is introduced, we endow quantities in old (physical) coordinates with an ε subscript, and leave the corresponding quantities in the new (computational) coordinates unsubscripted. In particular, flow maps then transform as

$$\boldsymbol{\eta}_\varepsilon = \boldsymbol{\xi}_\varepsilon \circ \boldsymbol{\eta}. \tag{2.10}$$

Thirdly, being sloppy about the distinction between vector fields and forms, we write an explicit '$\cdot$' to denote the dot product between two vectors, and no multiplication sign for vector–matrix multiplication, which takes precedence. Thus, for example,

$$\boldsymbol{u} \cdot \nabla \boldsymbol{v} \boldsymbol{w} = (\nabla \boldsymbol{v})^T \boldsymbol{u} \cdot \boldsymbol{w} = u_i (\partial_j v_i) w_j. \tag{2.11}$$

2.3. *The semi-geostrophic equations*

The semi-geostrophic equations arise from a single approximation, the *geostrophic momentum approximation*, where the advected velocity, but not the advecting velocity, is replaced by the geostrophic velocity (Eliassen 1948, 1962):

$$(\partial_t + \boldsymbol{u} \cdot \nabla)\boldsymbol{u} \to (\partial_t + \boldsymbol{u} \cdot \nabla)\boldsymbol{u}_{\mathrm{G}}. \tag{2.12}$$

Keeping with the conventions introduced in the previous section, we endow all quantities in old coordinates with an ε subscript, so that the semi-geostrophic momentum equation reads

$$\varepsilon(\partial_t + \boldsymbol{u}_\varepsilon \cdot \nabla)\nabla^\perp h_\varepsilon + \boldsymbol{u}_\varepsilon^\perp + \nabla h_\varepsilon = 0. \tag{2.13}$$

This equations combines with the continuity equation into a single prognostic equation for the layer depth h_ε, whose remarkable structure is exposed through the so-called

Hoskins transformation. Hoskins (1975) introduced new *semi-geostrophic coordinates* via

$$\boldsymbol{\eta} = \boldsymbol{\eta}_\varepsilon + \varepsilon \nabla h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon, \tag{2.14}$$

where the transformation is written in terms of the Lagrangian flows, and $\nabla h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon = (\nabla h_\varepsilon)(\boldsymbol{\eta}_\varepsilon(\boldsymbol{a}, t), t)$. Going to Eulerian positions, the transformation $\boldsymbol{\xi}_\varepsilon$ is implicitly defined through

$$\mathrm{id} = \boldsymbol{\xi}_\varepsilon + \varepsilon \nabla h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon. \tag{2.15}$$

By differentiating (2.14) in time, we obtain

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta} = \boldsymbol{u}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon + \varepsilon(\nabla \dot{h}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon + (\nabla\nabla h_\varepsilon) \circ \boldsymbol{\eta}_\varepsilon \dot{\boldsymbol{\eta}}_\varepsilon) = \nabla^\perp h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon, \tag{2.16}$$

where the last equality is due to the semi-geostrophic momentum equation (2.13). In other words,

$$\boldsymbol{u} = \nabla^\perp h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon; \tag{2.17}$$

the new velocity $\boldsymbol{u}$ is equal to the geostrophic velocity in the old coordinates. Further, (2.9) and (2.10) imply that

$$h = h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon \det \nabla \boldsymbol{\xi}_\varepsilon. \tag{2.18}$$

The right-hand expression can be closed in geostrophic coordinates as follows. First, taking the gradient of (2.15), using (2.17), yields

$$\boldsymbol{I} = \nabla \boldsymbol{\xi}_\varepsilon - \varepsilon \nabla \boldsymbol{u}^\perp, \tag{2.19}$$

where $\boldsymbol{I}$ denotes the 2×2 identity matrix, so that

$$\det \nabla \boldsymbol{\xi}_\varepsilon = \det(\boldsymbol{I} + \varepsilon \nabla \boldsymbol{u}^\perp). \tag{2.20}$$

Secondly,

$$\nabla(h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon) = (\nabla \boldsymbol{\xi}_\varepsilon)^T (\nabla h_\varepsilon) \circ \boldsymbol{\xi}_\varepsilon = -(\boldsymbol{I} + \varepsilon \nabla \boldsymbol{u}^\perp)^T \boldsymbol{u}^\perp = -\boldsymbol{u}^\perp - \tfrac{1}{2} \varepsilon \nabla |\boldsymbol{u}|^2. \tag{2.21}$$

Thus, if we define a streamfunction ψ by

$$h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon = \psi - \tfrac{1}{2} \varepsilon |\nabla \psi|^2, \tag{2.22}$$

then $\boldsymbol{u} \equiv \nabla^\perp \psi$ satisfies (2.17). Inserting (2.20) and (2.21) back into (2.18), we obtain

$$h = \left(\psi - \tfrac{1}{2} \varepsilon |\nabla \psi|^2\right) \det(\boldsymbol{I} - \varepsilon \nabla\nabla \psi). \tag{2.23}$$

Direct computation shows that the potential vorticity $q = 1/h$ is materially conserved, so that

$$(\partial_t + \nabla^\perp \psi \cdot \nabla) h = 0. \tag{2.24}$$

Potential vorticity advection together with the nonlinear elliptic Monge–Ampère equation (2.23) are a closed system for the semi-geostrophic dynamics in geostrophic coordinates.

The Hoskins transform can also be interpreted as a Legendre transformation; see Cullen & Purser (1984, 1989), and Benamou & Brenier (1998) for a proof of well-posedness based on this structure.

For later reference, we remark that the conservation of potential vorticity is easily translated back into physical coordinates. From (2.19), we infer that

$$\boldsymbol{I} = (\boldsymbol{I} + \varepsilon \nabla\nabla h_\varepsilon) \circ \boldsymbol{\xi}_\varepsilon \nabla \boldsymbol{\xi}_\varepsilon, \tag{2.25}$$

so that

$$q = \frac{1}{h} = \frac{1}{h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon \det \nabla \boldsymbol{\xi}_\varepsilon} = \frac{\det(\boldsymbol{I} + \varepsilon \nabla\nabla h_\varepsilon) \circ \boldsymbol{\xi}_\varepsilon}{h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon} \equiv q_\varepsilon \circ \boldsymbol{\xi}_\varepsilon \tag{2.26}$$

and conservation of potential vorticity takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(q_\varepsilon \circ \boldsymbol{\eta}_\varepsilon) = \frac{\mathrm{d}}{\mathrm{d}t}(q \circ \boldsymbol{\eta}) = 0. \tag{2.27}$$

Moreover, the semi-geostrophic equations conserve the energy

$$H_\varepsilon = \tfrac{1}{2} \int [\varepsilon |\nabla h_\varepsilon|^2 + h_\varepsilon] h_\varepsilon \mathrm{d}\boldsymbol{x}. \tag{2.28}$$

Both conservation laws arise naturally when we derive the semi-geostrophic equations variationally in § 7.

2.4. *The quasi-geostrophic equations*

In the second important distinguished scaling limit, the *quasi-geostrophic scaling*, the Burger number is of order one, but variations of the surface amplitude are small. When, as in (2.6), the deviation of the surface amplitude from equilibrium is denoted by εh_1, the quasigeostrophically rescaled shallow-water equations read

$$\varepsilon(\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u}) + \boldsymbol{u}^\perp + \varepsilon^{-1} \nabla (1 + \varepsilon h_1) = 0, \tag{2.29a}$$

$$\varepsilon \partial_t h_1 + \nabla \cdot ((1 + \varepsilon h_1) \boldsymbol{u}) = 0. \tag{2.29b}$$

At the lowest order $\varepsilon = 0$, (2.29*a*) again yields a geostrophic balance relation,

$$\boldsymbol{u}_{\mathrm{G}} = \varepsilon^{-1} \nabla^\perp h = \nabla^\perp h_1. \tag{2.30}$$

Substituting (2.30) back into the continuity equation (2.29*b*) simply confirms that $\boldsymbol{u}_{\mathrm{G}}$ is divergence free. The *quasi-geostrophic equations* are the next order correction to geostrophic balance. We make the ansatz

$$\boldsymbol{u} = \boldsymbol{u}_{\mathrm{G}} + \varepsilon \boldsymbol{u}_{\mathrm{A}}, \tag{2.31}$$

where $\boldsymbol{u}_{\mathrm{A}}$ denotes the ageostrophic component of the velocity field, substitute into (2.29), and collect first-order terms. The contributions from momentum and continuity equation, respectively, are

$$\boldsymbol{u}_{\mathrm{A}} = -(\partial_t + \boldsymbol{u}_{\mathrm{G}} \cdot \nabla) \nabla h_1, \tag{2.32a}$$

$$\partial_t h_1 + \nabla \cdot \boldsymbol{u}_{\mathrm{A}} = 0. \tag{2.32b}$$

Substituting the former equation into the latter, we obtain the quasi-geostrophic potential vorticity equation

$$(\partial_t + \boldsymbol{u}_{\mathrm{G}} \cdot \nabla)(h_1 - \Delta h_1) = 0. \tag{2.33}$$

Finally, the quasi-geostrophic equations possess the conserved 'energy'

$$H = \tfrac{1}{2} \int \left(h_1^2 + |\nabla h_1|^2 \right) \mathrm{d}\boldsymbol{x}. \tag{2.34}$$

3. A finite-dimensional example

As a caricature of the rotating shallow-water equations in semi-geostrophic scaling, we consider the system of coupled harmonic oscillators

$$\varepsilon \ddot{\boldsymbol{q}}_\varepsilon = -\Omega \boldsymbol{q}_\varepsilon + \boldsymbol{J} \dot{\boldsymbol{q}}_\varepsilon, \tag{3.1}$$

where $\boldsymbol{q}_\varepsilon : \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{J}\boldsymbol{q} \equiv \boldsymbol{q}^\perp = (-q_2, q_1)$, and $\Omega = \operatorname{diag}\{\nu^2, \omega^2\}$ is a constant diagonal 2×2 matrix. The corresponding Lagrangian is

$$L_\varepsilon = \tfrac{1}{2}\varepsilon|\dot{\boldsymbol{q}}_\varepsilon|^2 - V(\boldsymbol{q}_\varepsilon) - \boldsymbol{R}(\boldsymbol{q}_\varepsilon)\cdot\dot{\boldsymbol{q}}_\varepsilon, \tag{3.2}$$

where

$$\boldsymbol{R}(\boldsymbol{q}) = \tfrac{1}{2}\boldsymbol{J}\boldsymbol{q}, \qquad V(\boldsymbol{q}) = \tfrac{1}{2}\boldsymbol{q}^T\Omega\boldsymbol{q}. \tag{3.3}$$

Physically, this system describes the planar motion of a charged particle with harmonic restoring forces in a magnetic field perpendicular to the plane. The mass of the particle is ϵ, while all other parameters are scaled to unity.

When ε is small, both the components of $\boldsymbol{q}_\varepsilon$ form almost decoupled fast harmonic oscillators. In addition, the matrix $\boldsymbol{J}$ on the right-hand side of (3.1) introduces an additional symplectic structure independent of ε whose canonical coordinates are the two position coordinates q_1 and q_2. We note that when $\varepsilon = 0$, this structure is the only to survive; the corresponding Lagrangian is affine, i.e. it is linear in the velocities.

Our goal is to derive an effective equation for the slow evolution of q_1 and q_2 when ε is small, but different from zero. This toy model, being linear, can of course be solved explicitly by diagonalization, and the desired answer can be obtained by brute-force asymptotic expansion of the solution. However, the algebra involved is already sufficiently involved that a symbolic manipulation package is very helpful. The approach that we propose is computationally much simpler, does not depend on the linearity of the system, and will directly carry over to the rotating shallow-water equations.

The key idea is to introduce a near-identity change of (position) variables that can be expanded in powers of ε, insert this expansion into the Lagrangian, truncate to a consistent power in ε, and fix the coefficients of the transformation such that the truncated system is affine. This last step is the crucial closure assumption: the higher-order terms in the expansion are determined from the lower-order terms such that the leading-order (affine) structure is maintained. A simple application of the Hamilton principle then yields the effective equations of motion from the truncated transformed Lagrangian. The transformation can be undone *a posteriori* to the order of the approximation.

In finite dimensions, the procedure is simpler than for motion of the diffeomorphism group. The required near-identity transformation of the positions can be written straightforwardly as the asymptotic expansion

$$\boldsymbol{q}_\varepsilon = \boldsymbol{q} + \varepsilon\boldsymbol{q}' + \tfrac{1}{2}\varepsilon^2\boldsymbol{q}'' + O(\varepsilon^3). \tag{3.4}$$

We compute

$$\begin{aligned}\boldsymbol{R}(\boldsymbol{q}_\varepsilon)\cdot\dot{\boldsymbol{q}}_\varepsilon &= \tfrac{1}{2}\big(\boldsymbol{q}^\perp + \varepsilon\boldsymbol{q}'^\perp + \tfrac{1}{2}\varepsilon^2\boldsymbol{q}''^\perp\big)\cdot\big(\dot{\boldsymbol{q}} + \varepsilon\dot{\boldsymbol{q}}' + \tfrac{1}{2}\varepsilon^2\dot{\boldsymbol{q}}''\big) + O(\varepsilon^3)\\ &= \tfrac{1}{2}\boldsymbol{q}^\perp\cdot\dot{\boldsymbol{q}} + \varepsilon\boldsymbol{q}^\perp\cdot\dot{\boldsymbol{q}}' + \tfrac{1}{2}\varepsilon^2(\boldsymbol{q}^\perp\cdot\dot{\boldsymbol{q}}'' + \boldsymbol{q}'^\perp\cdot\dot{\boldsymbol{q}}') + O(\varepsilon^3)\end{aligned} \tag{3.5}$$

up to perfect time derivatives which are null-Lagrangians,

$$\begin{aligned}V(\boldsymbol{q}_\varepsilon) &= \tfrac{1}{2}\big(\boldsymbol{q} + \varepsilon\boldsymbol{q}' + \tfrac{1}{2}\varepsilon^2\boldsymbol{q}''\big)^T\Omega\big(\boldsymbol{q} + \varepsilon\boldsymbol{q}' + \tfrac{1}{2}\varepsilon^2\boldsymbol{q}''\big) + O(\varepsilon^3)\\ &= \tfrac{1}{2}\boldsymbol{q}^T\Omega\boldsymbol{q} + \varepsilon\boldsymbol{q}^T\Omega\boldsymbol{q}' + \tfrac{1}{2}\varepsilon^2(\boldsymbol{q}^T\Omega\boldsymbol{q}'' + \boldsymbol{q}'^T\Omega\boldsymbol{q}') + O(\varepsilon^3),\end{aligned} \tag{3.6}$$

and

$$\tfrac{1}{2}\varepsilon|\dot{\boldsymbol{q}}_\varepsilon|^2 = \tfrac{1}{2}\varepsilon|\dot{\boldsymbol{q}}|^2 + \varepsilon^2\dot{\boldsymbol{q}}\cdot\dot{\boldsymbol{q}}' + O(\varepsilon^3). \tag{3.7}$$

Altogether, we find the expansion

$$L_\varepsilon = L_0 + \varepsilon L_1 + \tfrac{1}{2}\varepsilon^2 L_2 + O(\varepsilon^3), \tag{3.8}$$

where, again dropping perfect time derivatives whenever convenient,

$$L_0 = -\tfrac{1}{2}\boldsymbol{q}^\perp \cdot \dot{\boldsymbol{q}} - \tfrac{1}{2}\boldsymbol{q}^T \Omega \boldsymbol{q}, \tag{3.9}$$

$$L_1 = \tfrac{1}{2}|\dot{\boldsymbol{q}}|^2 + \dot{\boldsymbol{q}}^\perp \cdot \boldsymbol{q}' - \boldsymbol{q}^T \Omega \boldsymbol{q}', \tag{3.10}$$

$$L_2 = 2\dot{\boldsymbol{q}} \cdot \dot{\boldsymbol{q}}' + \dot{\boldsymbol{q}}^\perp \cdot \boldsymbol{q}'' - \boldsymbol{q}'^\perp \cdot \dot{\boldsymbol{q}}' - \boldsymbol{q}^T \Omega \boldsymbol{q}'' - \boldsymbol{q}'^T \Omega \boldsymbol{q}'. \tag{3.11}$$

The crucial step now is to impose *degeneracy conditions*, i.e. to choose $\boldsymbol{q}'$ and $\boldsymbol{q}''$ that render the truncated Lagrangian affine to first and second order. At $O(\varepsilon)$, we must set

$$\boldsymbol{q}' = -\tfrac{1}{2}\dot{\boldsymbol{q}}^\perp + \text{any function of } \boldsymbol{q}. \tag{3.12}$$

For simplicity, we restrict ourselves to

$$\boldsymbol{q}' = -\tfrac{1}{2}\dot{\boldsymbol{q}}^\perp + \lambda \Omega \boldsymbol{q}. \tag{3.13}$$

This choice is motivated by the observation that $\boldsymbol{q}'$ vanishes – at least for a particular value of λ – when the toy model is in ‘geostrophic balance’. With this choice of $\boldsymbol{q}'$, we obtain

$$L_1 = -\left(\tfrac{1}{2} + \lambda\right)(\Omega \boldsymbol{q})^\perp \cdot \dot{\boldsymbol{q}} - \lambda \boldsymbol{q}^T \Omega^2 \boldsymbol{q}. \tag{3.14}$$

It is easily verified that the Euler–Lagrange equations for an affine Lagrangian of the form

$$L = \dot{\boldsymbol{q}} \cdot \boldsymbol{F}(\boldsymbol{q}) + g(\boldsymbol{q}) \tag{3.15}$$

are

$$\nabla^\perp \cdot \boldsymbol{F} \dot{\boldsymbol{q}}^\perp = \nabla g, \tag{3.16}$$

so that the reduced dynamics for our toy model including terms up to $O(\varepsilon)$ reads

$$\left[1 + \varepsilon\left(\tfrac{1}{2} + \lambda\right)(\omega^2 + \nu^2)\right] \dot{\boldsymbol{q}}^\perp = (\Omega + 2\varepsilon\lambda\Omega^2)\boldsymbol{q}. \tag{3.17}$$

This is a harmonic oscillator with frequency

$$\mu = \omega\nu \frac{\sqrt{1 + 2\varepsilon\lambda\nu^2}\sqrt{1 + 2\varepsilon\lambda\omega^2}}{1 + \varepsilon\left(\tfrac{1}{2} + \lambda\right)(\omega^2 + \nu^2)}. \tag{3.18}$$

Note that the first-order contribution is independent of λ, and coincides with the expansion of the slow eigenvalues of the full system to this order. In other words, λ is indeed a free parameter. In the special case when $\lambda = 1/2$, the frequency given by (3.18) is accurate even to $O(\varepsilon^2)$. This case corresponds to $\boldsymbol{q}' = 0$ in (3.13) if the dynamics were exactly following the leading-order ‘geostrophic balance’ dynamics.

Note further that the reduced dynamics does not represent the full system up to and including $O(\varepsilon)$ terms – the fast contributions to q_1 and q_2 are $O(\varepsilon)$, but are absent in the reduced system. Finally, the reconstruction of the solution in the original coordinates via (3.4) adds only amplitude corrections and is therefore only of interest with regard to the initialization of the reduced dynamics.

The second-order computation is only slightly more involved. By inserting the first-order degeneracy condition into L_2, we obtain

$$\begin{aligned} L_2 = \dot{\boldsymbol{q}}^\perp \cdot \left[\boldsymbol{q}'' + \tfrac{3}{4}\ddot{\boldsymbol{q}} - \tfrac{1}{4}\Omega\dot{\boldsymbol{q}}^\perp + \lambda(\Omega\dot{\boldsymbol{q}})^\perp\right] \\ - \boldsymbol{q}^T \Omega \boldsymbol{q}'' + \lambda \dot{\boldsymbol{q}}^\perp \cdot \Omega^2 \boldsymbol{q} - \lambda^2 (\Omega \boldsymbol{q})^\perp \cdot \Omega \dot{\boldsymbol{q}} - \lambda^2 \boldsymbol{q}^T \Omega^3 \boldsymbol{q}. \end{aligned} \tag{3.19}$$

Choosing

$$\boldsymbol{q}'' = -\tfrac{3}{4}\ddot{\boldsymbol{q}} + \tfrac{1}{4}\Omega\dot{\boldsymbol{q}}^{\perp} + \left(\tfrac{3}{4} - \lambda\right)(\Omega\dot{\boldsymbol{q}})^{\perp} \tag{3.20}$$

will render L_2 affine. Of course, as in the first-order degeneracy condition, we could add arbitrary functions of $\boldsymbol{q}$ only – we will do so when we apply the method to the rotating shallow-water equations. Staying with (3.20) for the time being, the resulting degenerate L_2 Lagrangian is

$$L_2 = \dot{\boldsymbol{q}} \cdot \left[\left(\tfrac{1}{4} - \lambda\right)(\Omega^2\boldsymbol{q})^{\perp} + \left(\tfrac{3}{4} - \lambda - \lambda^2\right)\Omega(\Omega\boldsymbol{q})^{\perp}\right] - \lambda^2\boldsymbol{q}^T\Omega^3\boldsymbol{q}. \tag{3.21}$$

Completing the Euler–Lagrange equations to second order yields a harmonic oscillator equation with frequency

$$\mu = \frac{\omega\nu\sqrt{1 + 2\varepsilon\lambda\nu^2 + \varepsilon^2\lambda^2\nu^4}\sqrt{1 + 2\varepsilon\lambda\omega^2 + \varepsilon^2\lambda^2\omega^4}}{1 + \varepsilon\left(\tfrac{1}{2} + \lambda\right)(\omega^2 + \nu^2) - \tfrac{1}{2}\varepsilon^2\left(\tfrac{1}{4} - \lambda\right)(\omega^4 + \nu^4) - \varepsilon^2\left(\tfrac{3}{4} - \lambda - \lambda^2\right)\nu^2\omega^2}. \tag{3.22}$$

By explicitly expanding in powers of ε, we can show that this expression is independent of λ up to second order, i.e. λ remains a free parameter.

A complete analysis of this toy system for linear and nonlinear potentials, including proofs of convergence which generalize the above observations, is provided in forthcoming work (Gottwald & Oliver 2005; Gottwald, Oliver & Tecu 2005). We finally remark that this system is more appropriate for illustrating the working of our method than the elastic pendulum, which has been explored as a simple model for atmospheric balance by Lynch (2002). Moreover, our model does not intend to address the issue of spontaneous generation of inertia–gravity waves, which has been studied in low-dimensional models starting with Lorenz (1980). For recent results and a more complete history of this line of research see, for example, Vanneste (2004).

4. Variational principles in fluid dynamics

This section introduces the variational framework for equations of rotating fluid flow. None of this material is original; the goal of this section is to collect pertinent results in consistent notation.

4.1. *Rotating shallow-water Lagrangians*

Our parent system, the rotating shallow-water equations (2.4), are the equations of a barotropic fluid with pressure function $\pi = \ln h$. The configuration space is formally the semidirect product of the group of diffeomorphisms of the plane, the space of flow maps $\boldsymbol{\eta}$, with the vector space of smooth functions, the space of densities h.

Fluid Lagrangians are invariant with respect to the tangent lift of the natural group action on this semidirect product – simply speaking, they depend on Eulerian velocities and advected quantities only – and can therefore be treated in the framework of Euler–Poincaré reduction (see, e.g. Arnold & Khesin 1998; Holm, Marsden & Ratiu 1998). In practical terms, this means that the equations of motion, the *Euler–Poincaré equations*, can be obtained by taking variations, fundamentally with respect to the flow map $\boldsymbol{\eta}$ and vanishing at the temporal endpoints, of the action

$$S = \int_{t_1}^{t_2} L[\boldsymbol{u}, h]\,\mathrm{d}t. \tag{4.1}$$

The Lagrangian variations $\delta\boldsymbol{\eta}$ induce variations of the Eulerian quantities $\boldsymbol{u}$ and h as follows. First, taking the variational derivative $\delta\boldsymbol{\eta}$ means differentiating along a curve on the diffeomorphism group. Hence, we can associate an Eulerian vector field

$\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{x})$ via

$$\delta\boldsymbol{\eta} = \boldsymbol{w} \circ \boldsymbol{\eta}. \tag{4.2}$$

For compressible flow, $\boldsymbol{w}$ is arbitrary, while for incompressible flow, variations of the flow map must remain area preserving – the corresponding vector field $\boldsymbol{w}$ must be divergence free. By differentiating (4.2) in time and taking the variational derivative of (2.7), we obtain the so-called *Lin constraint* (Bretherton 1970),

$$\delta\boldsymbol{u} = \dot{\boldsymbol{w}} + \nabla\boldsymbol{w}\boldsymbol{u} - \nabla\boldsymbol{u}\boldsymbol{w}. \tag{4.3}$$

Since $h^{-1} \circ \boldsymbol{\eta} = \det \nabla\boldsymbol{\eta}$, the Liouville theorem applied to the flow generated by $\boldsymbol{w}$, where the variational parameter is playing the role of time, directly implies the 'continuity equation'

$$\delta h + \nabla \cdot (\boldsymbol{w}h) = 0. \tag{4.4}$$

Thus, we have a way of translating between the Lagrangian variation $\delta\boldsymbol{\eta}$ and the associated Eulerian variations $\delta\boldsymbol{u}$ and δh. In practice, we will choose whichever formulation is more convenient for the task at hand, and move freely between the two.

As a first example, take the semigeostrophically scaled rotating shallow-water Lagrangian (Salmon 1983),

$$\begin{aligned} L &= \int \left[\left(\boldsymbol{R} + \tfrac{1}{2}\varepsilon\boldsymbol{u}\right) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - \tfrac{1}{2}h \circ \boldsymbol{\eta}\right] \mathrm{d}\boldsymbol{a} \\ &= \int h\left[\boldsymbol{R} \cdot \boldsymbol{u} + \tfrac{1}{2}\varepsilon|\boldsymbol{u}|^2 - \tfrac{1}{2}h\right] \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{4.5}$$

where $\boldsymbol{R}$ denotes the vector potential of the Coriolis parameter, so that $\nabla^\perp \cdot \boldsymbol{R} = f \equiv 1$. Plugging L into the action integral and taking variations, we find

$$\delta S = \int_{t_1}^{t_2} \int \left[\delta h\left(\boldsymbol{R} \cdot \boldsymbol{u} + \tfrac{1}{2}\varepsilon|\boldsymbol{u}|^2 - h\right) + h(\boldsymbol{R} + \varepsilon\boldsymbol{u}) \cdot \delta\boldsymbol{u}\right] \mathrm{d}\boldsymbol{x}\, \mathrm{d}t. \tag{4.6}$$

Inserting the constrained variations (4.3) and (4.4), integrating by parts in space and time, using the continuity equation $\dot{h} + \nabla \cdot (\boldsymbol{u}h) = 0$, the time independence of $\boldsymbol{R}$, and collecting terms, we obtain straightforwardly that

$$\delta S = \int_{t_1}^{t_2} \int h\boldsymbol{w} \cdot \left[(\nabla\boldsymbol{R}^T - \nabla\boldsymbol{R})\boldsymbol{u} - \varepsilon\dot{\boldsymbol{u}} - \varepsilon\nabla\boldsymbol{u}\boldsymbol{u} - \nabla h\right] \mathrm{d}\boldsymbol{x}\, \mathrm{d}t. \tag{4.7}$$

Owing to identity (A 6), the terms in the square bracket yield precisely the shallow-water momentum equation (2.4*a*) with semi-geostrophic scaling $B = \varepsilon$.

One of the main advantages of the variational route is that the conservation of energy and potential vorticity is built into the formalism. This can be made explicit by writing out an extended Lagrangian that separates symplectic structure from the Hamiltonian,

$$L = \int \boldsymbol{F}(\boldsymbol{u}) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - H[\boldsymbol{u}, h], \tag{4.8}$$

where *a priori* $\boldsymbol{u}$ and $\dot{\boldsymbol{\eta}}$ are treated as independent quantities; the relationship $\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}$ is recovered by taking variations in $\boldsymbol{u}$ (see, e.g. Salmon 1983). We can then show that H is the Noetherian conserved quantity arising from time translation invariance, and that the *potential vorticity*

$$q = \frac{\nabla^\perp \cdot \boldsymbol{F}}{h} \tag{4.9}$$

is the materially conserved quantity arising from the invariance under particle relabelling (see, e.g. Ripa 1981). A fully variational derivation of the conservation of potential vorticity involves taking variations that do not vanish at the temporal endpoints along a trajectory satisfying the Euler–Poincaré equations of motion, so that only boundary terms arising from integration by parts remain.

For the semigeostrophically scaled rotating shallow-water equations, the extended Lagrangian reads

$$L = \int (\boldsymbol{R} + \varepsilon \boldsymbol{u}) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} \, \mathrm{d}\boldsymbol{a} - H, \tag{4.10a}$$

$$H = \tfrac{1}{2} \int [\varepsilon |\boldsymbol{u}|^2 + h] \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}, \tag{4.10b}$$

with the well-known potential vorticity

$$q = \frac{1 + \varepsilon \nabla^\perp \cdot \boldsymbol{u}}{h}. \tag{4.11}$$

We now discuss two important special cases: affine Lagrangians and Lagrangians for incompressible fluids, which will arise as Lagrangians of nearly geostrophic models in the semi-geostrophic and the quasi-geostrophic limit, respectively. We derive general equations of motion and the corresponding conservation laws for each.

4.2. *Affine Lagrangians*

Consider an affine (degenerate) Lagrangian of the form

$$L = \int (\boldsymbol{F}(h) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - g(h) \circ \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{a} = \int h(\boldsymbol{F}(h) \cdot \boldsymbol{u} - g(h)) \, \mathrm{d}\boldsymbol{x}, \tag{4.12}$$

where $\boldsymbol{F}$ and g are arbitrary functionals of the layer depth h. We insert this Lagrangian into the action integral and take variations with respect to $\boldsymbol{\eta}$, using $\mathrm{D}\boldsymbol{F}$ to denote the Fréchet-derivative of $\boldsymbol{F}$ and $\mathrm{D}\boldsymbol{F}^*$ to denote the formal L^2 adjoint thereof:

$$\begin{aligned}
\delta S &= \delta \int_{t_1}^{t_2} \int (\boldsymbol{F}(h) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - g(h) \circ \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int ((\delta \boldsymbol{F}) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + (\nabla \boldsymbol{F}) \circ \boldsymbol{\eta} \delta \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \boldsymbol{F} \circ \boldsymbol{\eta} \cdot \delta \dot{\boldsymbol{\eta}} \\
&\quad - (\delta g) \circ \boldsymbol{\eta} - (\nabla g) \circ \boldsymbol{\eta} \delta \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int ((\mathrm{D}\boldsymbol{F}(h) \delta h) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - \dot{\boldsymbol{F}} \circ \boldsymbol{\eta} \cdot \delta \boldsymbol{\eta} - (\nabla \boldsymbol{F} - \nabla \boldsymbol{F}^T) \circ \boldsymbol{\eta} \dot{\boldsymbol{\eta}} \cdot \delta \boldsymbol{\eta} \\
&\quad - (\mathrm{D}g(h) \delta h) \circ \boldsymbol{\eta} - (\nabla g) \circ \boldsymbol{\eta} \delta \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int h(\mathrm{D}\boldsymbol{F}(h) \delta h \cdot \boldsymbol{u} - \mathrm{D}\boldsymbol{F}(h) \dot{h} \cdot \boldsymbol{w} - \boldsymbol{w} \cdot \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} \\
&\quad - \mathrm{D}g(h) \delta h - \boldsymbol{w} \cdot \nabla g) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \int (\delta h \mathrm{D}\boldsymbol{F}^*(h) \cdot (h \boldsymbol{u}) - h \boldsymbol{w} \cdot \dot{\boldsymbol{F}} - h \boldsymbol{w} \cdot \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} \\
&\quad - \delta h \mathrm{D}g^*(h) h - h \boldsymbol{w} \cdot \nabla g) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t
\end{aligned}$$

$$
= \int_{t_1}^{t_2} \int h\boldsymbol{w} \cdot (\nabla(\mathrm{D}\boldsymbol{F}^*(h) \cdot (h\boldsymbol{u})) - \dot{\boldsymbol{F}} - \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} - \nabla(\mathrm{D}g^*(h)h) - \nabla g) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{4.13}
$$

Since $\boldsymbol{w}$ is arbitrary, the vanishing of δS yields the degenerate Euler–Poincaré equation

$$
\nabla(\mathrm{D}\boldsymbol{F}^*(h) \cdot (h\boldsymbol{u})) - \dot{\boldsymbol{F}} - \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} = \nabla(\mathrm{D}g^*(h)h) + \nabla g. \tag{4.14}
$$

Note that $\dot{\boldsymbol{F}} = \mathrm{D}\boldsymbol{F}(h)\dot{h}$, so that all time derivatives can be eliminated via the continuity equation – we obtain a diagnostic relationship between h and $\boldsymbol{u}$.

To derive the expression for the potential vorticity, we could follow the Noetherian approach and work explicitly with the particle relabelling symmetry; see, for example, Bridges, Hydon & Reich (2005). However, it turns out to be much easier in this case to directly take the curl of (4.14),

$$
\nabla^\perp \cdot \dot{\boldsymbol{F}} + \nabla \cdot (\boldsymbol{u} \nabla^\perp \cdot \boldsymbol{F}) = 0, \tag{4.15}
$$

whence, dividing through by h and using that $\dot{h} + \nabla \cdot (h\boldsymbol{u}) = 0$, we find that the potential vorticity

$$
q = \frac{\nabla^\perp \cdot \boldsymbol{F}(h)}{h} \tag{4.16}
$$

is advected by the velocity field $\boldsymbol{u}$,

$$
\partial_t q + \boldsymbol{u} \cdot \nabla q = 0. \tag{4.17}
$$

Similarly, the conservation of the Hamiltonian

$$
H = \int h g(h) \,\mathrm{d}\boldsymbol{x} \tag{4.18}
$$

follows from Noether's theorem, or can easily be verified by direct computation.

4.3. *Incompressible fluid Lagrangians*

We consider the general case of an incompressible, rotating fluid with a Lagrangian of the form

$$
L = \int (\boldsymbol{R} + \boldsymbol{N}(\boldsymbol{u})) \cdot \boldsymbol{u} \,\mathrm{d}\boldsymbol{x}, \tag{4.19}
$$

where $\boldsymbol{N}$ is a potentially nonlinear operator acting on $\boldsymbol{u}$. We take variations that are, as before, subject to the Lin constraint (4.3). Since the flow is incompressible, the vector fields $\boldsymbol{u}$ and $\boldsymbol{w}$ are divergence free, so that

$$
\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \int (\boldsymbol{R} \cdot \delta\boldsymbol{u} + \boldsymbol{N}(\boldsymbol{u}) \cdot \delta\boldsymbol{u} + \mathrm{D}\boldsymbol{N}(\boldsymbol{u})\delta\boldsymbol{u} \cdot \boldsymbol{u}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\
&= \int_{t_1}^{t_2} \int (\boldsymbol{R} + \boldsymbol{N}(\boldsymbol{u}) + \mathrm{D}\boldsymbol{N}^*(\boldsymbol{u})\boldsymbol{u}) \cdot \delta\boldsymbol{u} \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\
&= \int_{t_1}^{t_2} \int \boldsymbol{F} \cdot (\dot{\boldsymbol{w}} + \nabla\boldsymbol{w}\boldsymbol{u} - \nabla\boldsymbol{u}\boldsymbol{w}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\
&= -\int_{t_1}^{t_2} \int \boldsymbol{w} \cdot (\dot{\boldsymbol{F}} + (\nabla\boldsymbol{F} - \nabla\boldsymbol{F}^T)\boldsymbol{u}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\
&= -\int_{t_1}^{t_2} \int \boldsymbol{w} \cdot (\dot{\boldsymbol{F}} + \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t,
\end{aligned} \tag{4.20}
$$

where we have used identity (A 6) in the last step, and

$$\boldsymbol{F} \equiv \boldsymbol{R} + \boldsymbol{N}(\boldsymbol{u}) + \mathrm{D}\boldsymbol{N}^*(\boldsymbol{u})\boldsymbol{u}. \tag{4.21}$$

Since $\boldsymbol{w}$ is an arbitrary divergence-free vector field, the expression in parentheses on the right-hand side of (4.20) must be zero modulo gradients. Therefore, the Euler–Poincaré equations of motion are

$$\partial_t \boldsymbol{F} + \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} + \nabla p = 0, \tag{4.22}$$

p being the pressure which is determined via the incompressibility constraint $\nabla \cdot \boldsymbol{u} = 0$. The corresponding potential vorticity equation is most easily obtained by taking the curl of this expression,

$$(\partial_t + \boldsymbol{u} \cdot \nabla)\nabla^\perp \cdot \boldsymbol{F} = 0, \tag{4.23}$$

and the conserved energy takes the form

$$H = \tfrac{1}{2} \int \boldsymbol{u} \cdot \boldsymbol{N}(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{x}. \tag{4.24}$$

A general computation of this type in the context of the quasi-geostrophic equations has previously appeared in Holm & Zeitlin (1998).

4.4. *Variational asymptotics: L_1 and LSG dynamics*

The fundamental idea, pioneered by Salmon (1993, 1985, 1996) is to derive approximate equations for nearly geostrophic flow by approximating the Lagrangian rather than the equations of motions directly. If the approximation preserves time translation and particle relabelling symmetries, the resulting approximate system will possess proper analogues of the original conserved energy and potential vorticity.

In this section, we first recall the approach of Salmon, who proceeds in two steps. He initially constrains the Hamiltonian phase space to the submanifold defined by geostrophic motion. The resulting system is called the L_1 equations. In a second step, he introduces a near-identity change of variables that, when only keeping terms to the same consistent asymptotic order, yields a simpler system in canonical coordinates, the large-scale semi-geostrophic (LSG) equations.

Any imposed functional dependence $\boldsymbol{u} = \boldsymbol{F}(h)$ of the Hamiltonian momentum on the mass configuration in the extended variational principle (4.8) defines a constraint manifold in the Hamiltonian phase space. In particular, choosing geostrophic balance

$$\boldsymbol{u} = \nabla^\perp h \tag{4.25}$$

as the constraint, we obtain the affine Lagrangian

$$L_{\mathrm{c}} = \int \left[(\boldsymbol{R} + \varepsilon \nabla^\perp h) \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - \tfrac{1}{2}(\varepsilon |\nabla h|^2 + h) \circ \boldsymbol{\eta} \right] \mathrm{d}\boldsymbol{a}. \tag{4.26}$$

The resulting Euler–Poincaré equations of motion are

$$[1 - \varepsilon(h\Delta + 2\nabla h \cdot \nabla)]\boldsymbol{u} = \nabla^\perp \left[h - \varepsilon \left(h\Delta h + \tfrac{1}{2}|\nabla h|^2 \right) \right], \tag{4.27}$$

which, for given h, is a second-order elliptic problem for the velocity $\boldsymbol{u}$. The computation, using the formalism set up in § 4.2, is a special case of § 5.2; the reader is referred to this later section for details. The corresponding potential vorticity, computed directly from (4.16), reads

$$q = \frac{1 + \varepsilon \Delta h}{h}. \tag{4.28}$$

Setting

$$\boldsymbol{u} = \nabla^{\perp} h + \boldsymbol{u}_{\mathrm{A}}, \tag{4.29}$$

where $\boldsymbol{u}_{\mathrm{A}}$ is the ageostrophic part of the velocity field, and using identity (A 10), we can rewrite (4.27) as an elliptic equation for $\boldsymbol{u}_{\mathrm{A}}$,

$$\begin{aligned}[1 - \varepsilon(h\Delta + 2\nabla h \cdot \nabla)]\boldsymbol{u}_{\mathrm{A}} &= \varepsilon\left[h\Delta\nabla^{\perp} h + 2\nabla h \cdot \nabla\nabla^{\perp} h - \nabla^{\perp}\left(h\Delta h + \tfrac{1}{2}|\nabla h|^2\right)\right] \\ &= \varepsilon(\nabla h \cdot \nabla\nabla^{\perp} h - \nabla^{\perp} h \Delta h) \\ &= \varepsilon \nabla^{\perp} h \cdot \nabla\nabla h. \end{aligned} \tag{4.30}$$

This coincides with Salmon's (1985, equation 2.27) expression for the ageostrophic velocity. We note that, at best, the ageostrophic velocity is of the same regularity class as h; since the geostrophic velocity is a skew-gradient of h, the full inversion from h to $\boldsymbol{u}$ therefore loses one derivative. Further, since (4.28) implies

$$(q - \varepsilon\Delta)h = 1, \tag{4.31}$$

and this equation is elliptic and positive as long as the initial potential vorticity is positive, the full potential vorticity inversion gains one derivative – the functional setting is similar to that of the two-dimensional incompressible Euler equations.

On the other hand, the L_1 equations, involving variable coefficient elliptic equations, are harder to implement numerically than two-dimensional ideal fluid equations. Salmon (1985) therefore suggested further approximating the system by applying a truncated near-identity transformation to canonical coordinates. The Euler–Poincaré equations of the resulting so-called large-scale semi-geostrophic equations are

$$\boldsymbol{u} = \nabla^{\perp}\left(h + \varepsilon h\Delta h + \tfrac{1}{2}|\nabla h|^2\right), \tag{4.32}$$

and the potential vorticity

$$q = \frac{1}{h}. \tag{4.33}$$

Since both L_1 and LSG arise as special cases in our setting, we will not work through the details of the construction. Equation (4.32) shows that the advecting velocity field is less smooth than the advected quantity. As a consequence, standard arguments for proving well-posedness of such equations fail, and numerical simulations indicate that the LSG equations, though much simpler than the L_1 equations, are indeed ill posed even for short times. Unfortunately, ill-posedness extends to Salmon's (1996) large-scale semi-geostrophic dynamics for rotating stratified flow (R. Ford, personal communication 2000), and to Ford, Malham & Oliver's (2002) attempt to fix the indefiniteness of the LSG energy by adding higher-order terms.

5. The LSG hierarchy

In this section, we apply the procedure outlined in § 3 to the rotating shallow-water equations in semi-geostrophic scaling. At first order in ε, we obtain a one-parameter family of models that includes Salmon's L_1 and LSG equations, motivating the name LSG hierarchy. We also carry the computation to second order, where we discuss models with altogether five free parameters. The LSG hierarchy does not include the Hoskins semi-geostrophic equations, even though these equations are based on the same scaling. We take up this issue in § 7, where we present a different ansatz for the variational asymptotics that yields the classical semi-geostrophic equations.

We are motivated by the question of whether Salmon's idea of using truncated transformations into convenient coordinates can be generalized in a way that does not necessarily lead to ill-posed models. The crucial observation is that we need not constrain the dynamics explicitly – we can let consistent truncation to a certain asymptotic order do all the work. If, by means of a clever choice of transformation, the truncated system degenerates, constraints will appear naturally by the Dirac (1966) theory of constraints. However, since all we require is a reduced set of equations, we need not compute any constraints explicitly.

5.1. *Set-up*

We follow the conventions introduced in § 2.2, where $\boldsymbol{u}_\varepsilon$ denotes the velocity in physical coordinates, and $\boldsymbol{u}$ the velocity in a new, yet-to-be-determined, coordinate system. Correspondingly, h_ε denotes the layer depth in physical, and h the layer depth in the new coordinate system. Then, the full semigeostrophically scaled shallow-water Lagrangian reads

$$L_\varepsilon = \int \left[\boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon + \tfrac{1}{2}\varepsilon |\dot{\boldsymbol{\eta}}_\varepsilon|^2 - \tfrac{1}{2} h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \right] \mathrm{d}\boldsymbol{a}. \tag{5.1}$$

Recall that the flow in each coordinate system has an associated vector field via

$$\dot{\boldsymbol{\eta}} = \boldsymbol{u} \circ \boldsymbol{\eta}, \tag{5.2}$$

$$\dot{\boldsymbol{\eta}}_\varepsilon = \boldsymbol{u}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon, \tag{5.3}$$

and that the change of coordinates is expressed by the transformation

$$\boldsymbol{\eta}_\varepsilon = \boldsymbol{\xi}_\varepsilon \circ \boldsymbol{\eta}. \tag{5.4}$$

At this stage, the fundamental objects are still the flow maps $\boldsymbol{\eta}$ and $\boldsymbol{\eta}_\varepsilon$, and there is no truncation to some order of ε yet. The crucial point is that we can regard $\boldsymbol{\xi}_\varepsilon$ as a flow in ε, and associate with it a vector field $\boldsymbol{v}_\varepsilon$ via

$$\boldsymbol{\xi}_\varepsilon' = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\xi}_\varepsilon, \tag{5.5}$$

where $\boldsymbol{\xi}_0 = \mathrm{id}$ and the prime denotes a derivative with respect to ε.

This basic set-up is similar to, and has in fact been motivated by, the Lagrangian averaging construction of Marsden & Shkoller (2001, 2003), which has recently been extended to compressible fluids by Bhat *et al.* (2005). The difference is that in our case there is no explicit averaging. Instead, we have the Rossby number as the natural physical small parameter, and model reduction is achieved purely by shifting all non-degeneracy into orders beyond those that are kept.

The task is now to expand systematically all quantities in the 'old' Lagrangian L_ε in powers of ε. The computations are most easily written in terms of the Taylor coefficients of the Eulerian vector fields $\boldsymbol{u}_\varepsilon$ and $\boldsymbol{v}_\varepsilon$, which we denote by $\boldsymbol{u}$, $\boldsymbol{u}'$, $\boldsymbol{u}''$, etc. Appendix B summarizes the relationship between Eulerian and Lagrangian expansion coefficients, and gives the details of the expansion of each term in the Lagrangian. In this procedure, $\boldsymbol{v}$, $\boldsymbol{v}'$, and their higher-order cousins can be chosen by us, and we use this freedom to impose degeneracy at each relevant order of the expansion.

A lengthy, but straightforward computation, the details of which are provided in Appendix B, yields the expansion

$$L_\varepsilon = L_0 + \varepsilon L_1 + \tfrac{1}{2}\varepsilon^2 L_2 + O(\varepsilon^3), \tag{5.6}$$

where

$$L_0 = \int \left[\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} - \tfrac{1}{2} h \circ \boldsymbol{\eta} \right] \mathrm{d}\boldsymbol{a}, \tag{5.7}$$

$$L_1 = \int \left[\boldsymbol{v}^\perp \cdot \boldsymbol{u} + \tfrac{1}{2}|\boldsymbol{u}|^2 + \tfrac{1}{2} h \nabla \cdot \boldsymbol{v} \right] \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}, \tag{5.8}$$

$$L_2 = \int \left[\boldsymbol{u} \cdot (\boldsymbol{v}' + \nabla \boldsymbol{v} \boldsymbol{v})^\perp + (\boldsymbol{v}^\perp + 2\boldsymbol{u}) \cdot (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u}) + \tfrac{1}{2} h (\nabla \cdot \boldsymbol{v}' + \boldsymbol{v} \cdot \nabla \nabla \cdot \boldsymbol{v} - (\nabla \cdot \boldsymbol{v})^2) \right] \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}. \tag{5.9}$$

5.2. *First-order LSG models*

We will now look at the first- and second-order contributions in turn, fixing $\boldsymbol{v}$ and $\boldsymbol{v}'$ such that L_1 and L_2, respectively, become affine. At first order, it is immediately clear that any choice of the form

$$\boldsymbol{v} = \tfrac{1}{2} \boldsymbol{u}^\perp + \boldsymbol{F}(h) \tag{5.10}$$

will render L_1 affine. For simplicity, we restrict ourselves to the one-parameter family of transformations

$$\boldsymbol{v} = \tfrac{1}{2} \boldsymbol{u}^\perp + \lambda \nabla h. \tag{5.11}$$

This restriction is motivated by the observation that under geostrophic balance, the second-order term is a scalar multiple of the first. Thus, when diagnosing the transformation with geostrophic balance, the factor $(\frac{1}{2} - \lambda)$ is scaling the transformation vector field linearly to leading order.

With (5.11), the first-order Lagrangian reads

$$\begin{aligned} L_1 &= \int \left[\lambda \nabla^\perp h \cdot \boldsymbol{u} + \tfrac{1}{4} h \nabla \cdot \boldsymbol{u}^\perp + \tfrac{1}{2} \lambda h \Delta h \right] h \, \mathrm{d}\boldsymbol{x} \\ &= \left(\lambda + \tfrac{1}{2}\right) \int h \nabla^\perp h \cdot \boldsymbol{u} \, \mathrm{d}\boldsymbol{x} - \lambda \int h |\nabla h|^2 \, \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{5.12}$$

We use the general Euler–Poincaré equations (4.14) to compute the equations of motion. In the notation of § 4.2,

$$\boldsymbol{F}(h) = \boldsymbol{R} + \varepsilon \left(\lambda + \tfrac{1}{2}\right) \nabla^\perp h, \tag{5.13}$$

$$g(h) = \lambda \varepsilon |\nabla h|^2 - \tfrac{1}{2} h, \tag{5.14}$$

so that, for some scalar function ϕ,

$$\mathrm{D}\boldsymbol{F}(h)\phi = \varepsilon \left(\lambda + \tfrac{1}{2}\right) \nabla^\perp \phi, \tag{5.15}$$

$$\mathrm{D}g(h)\phi = 2\lambda \varepsilon \nabla h \cdot \nabla \phi - \tfrac{1}{2} \phi, \tag{5.16}$$

and, for some vector field $\boldsymbol{w}$ and scalar function ψ,

$$\mathrm{D}\boldsymbol{F}^*(h) \cdot \boldsymbol{w} = -\varepsilon \left(\lambda + \tfrac{1}{2}\right) \nabla^\perp \cdot \boldsymbol{w}, \tag{5.17}$$

$$\mathrm{D}g^*(h)\psi = -2\lambda \varepsilon \nabla \cdot (\psi \nabla h) - \tfrac{1}{2} \psi. \tag{5.18}$$

Therefore, the terms involving $\boldsymbol{F}$ of (4.14) read

$$\begin{aligned} &\nabla(\mathrm{D}\boldsymbol{F}^*(h) \cdot (h\boldsymbol{u})) - \dot{\boldsymbol{F}} - \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{F} \\ &= \varepsilon \left(\lambda + \tfrac{1}{2}\right) (-\nabla \nabla^\perp \cdot (h\boldsymbol{u}) + \nabla^\perp \nabla \cdot (h\boldsymbol{u}) - \boldsymbol{u}^\perp \nabla^\perp \cdot \nabla^\perp h) - \boldsymbol{u}^\perp \nabla^\perp \cdot \boldsymbol{R} \\ &= \varepsilon \left(\lambda + \tfrac{1}{2}\right) (\Delta(h\boldsymbol{u}^\perp) - \boldsymbol{u}^\perp \Delta h) - \boldsymbol{u}^\perp \\ &= \left(\varepsilon \left(\lambda + \tfrac{1}{2}\right) (h\Delta + 2\nabla h \cdot \nabla) - 1\right) \boldsymbol{u}^\perp. \end{aligned} \tag{5.19}$$

Similarly, the terms involving g are

$$\begin{aligned}\nabla(\mathrm{D}g^*(h)h + \nabla g) &= -\nabla\big(2\lambda\varepsilon\nabla\cdot(h\nabla h) + \tfrac{1}{2}h - \lambda\varepsilon|\nabla h|^2 + \tfrac{1}{2}h\big)\\ &= -\nabla(h + \lambda\varepsilon(2h\Delta h + |\nabla h|^2)).\end{aligned} \tag{5.20}$$

Equating (5.19) with (5.20), we obtain

$$\big[1 - \varepsilon\big(\lambda + \tfrac{1}{2}\big)(h\Delta + 2\nabla h\cdot\nabla)\big]\boldsymbol{u} = \nabla^{\perp}[h - \varepsilon\lambda(2h\Delta h + |\nabla h|^2)]. \tag{5.21}$$

Moreover, (4.16) yields the potential vorticity

$$q = \frac{1 + \varepsilon\big(\lambda + \frac{1}{2}\big)\Delta h}{h}. \tag{5.22}$$

Here, $\lambda = -1/2$ corresponds to a complete loss of relative vorticity, while $\lambda = 1/2$ includes relative vorticity with the same weight as for the parent dynamics in physical coordinates.

When $\lambda > -1/2$ and provided h and q are positive and sufficiently smooth, equation (5.21) is not only elliptic, but its weak formulation is also coercive in the (Sobolev) space H^1 of square integrable functions with square integrable first derivatives. This key requisite for proving the existence of unique weak solutions via the Lax–Milgram theorem (see, e.g. Evans 1998) is shown as follows.

We say that $\boldsymbol{u} \in H^1$ solves the weak form of (5.21) if

$$B(\boldsymbol{u}, \boldsymbol{v}) = \int \boldsymbol{v}\cdot\nabla^{\perp}[h - \varepsilon\lambda(2h\Delta h + |\nabla h|^2)]\,\mathrm{d}\boldsymbol{x} \tag{5.23}$$

for every $\boldsymbol{v} \in H^1$, where

$$B(\boldsymbol{u}, \boldsymbol{v}) \equiv \int \boldsymbol{v}\cdot[1 - \sigma(h\Delta + 2\nabla h\cdot\nabla)]\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}. \tag{5.24}$$

Then, after integration by parts,

$$\begin{aligned}B(\boldsymbol{u}, \boldsymbol{u}) &= \int \big[\boldsymbol{u}\cdot\boldsymbol{u} + \sigma h\nabla\boldsymbol{u} : \nabla\boldsymbol{u} - \tfrac{1}{2}\sigma\nabla h\cdot\nabla|\boldsymbol{u}|^2\big]\,\mathrm{d}\boldsymbol{x}\\ &= \int \big[\tfrac{1}{2}(1 + hq)|\boldsymbol{u}|^2 + \sigma h|\nabla\boldsymbol{u}|^2\big]\,\mathrm{d}\boldsymbol{x},\end{aligned} \tag{5.25}$$

which defines a norm equivalent to the canonical H^1 norm so long as $hq > 1$ and $\sigma h > 1$, uniformly on the plane. This is true, in particular, if h, q and $\sigma \equiv \varepsilon(\lambda + 1/2)$ are positive and $h \to 1$ as $|\boldsymbol{x}| \to \infty$.

We now consider three special choices for λ. When $\lambda = 1/2$, the transformation is, *a posteriori*, the identity up to terms of order ε^2. The equations of motion in this case are elliptic, coercive, and given by (4.27) – we have recovered Salmon's L_1 dynamics. In other words, Salmon's constraint to geostrophic balance has been replaced by choosing a transform to an affine Lagrangian that is near-identity to one order higher than generically expected for our ansatz. Whether the L_1 model is also more accurate, perhaps by one order as for the frequency of the linear toy problem in § 3, remains to be investigated.

When $\lambda = -1/2$, the function $\boldsymbol{F}$ which defines the symplectic structure becomes very simple, namely $\boldsymbol{F} = \boldsymbol{R}$ – the symplectic structure is canonical. This corresponds to the case of Salmon's LSG equations. However, the relation between h and $\boldsymbol{u}$, equation (4.32), ceases to be second-order elliptic and, as mentioned previously, the resulting system of equations is ill posed. In fact, owing to the restriction on coercivity, none of the models with $\lambda \leqslant -1/2$ can be well posed.

Half way between L_1 and LSG, when $\lambda = 0$, lies another special case. Here, both the transformation (5.11) and the right-hand side of the Euler–Poincaré equation (5.21) take the simplest possible form, while the left-hand side of (5.21) remains a coercive elliptic operator with non-constant coefficients, i.e.

$$\left[1 - \tfrac{1}{2}\varepsilon(h\Delta + 2\nabla h \cdot \nabla)\right]\boldsymbol{u} = \nabla^{\perp} h, \tag{5.26}$$

and the potential vorticity reads

$$q = \frac{1 + \frac{1}{2}\varepsilon\Delta h}{h}, \tag{5.27}$$

so that

$$\left(q - \tfrac{1}{2}\varepsilon\Delta\right)h = 1. \tag{5.28}$$

The remarkable consequence is that now potential vorticity inversion 'gains' three derivatives, the maximum possible for first-order models of this type. Two derivatives are gained by inverting (5.28), and one derivative is gained through the inversion of (5.26).

We conclude that the $\lambda = 0$ case resembles the regularity type of the two-dimensional Lagrangian-averaged Euler equations; see Holm *et al.* (1998) and Holm (1999) for a derivation, and Oliver & Shkoller (2001) for their analytical properties. Although these equations are, in principle, as difficult to solve as the L_1 equations, we expect that the built-in non-dissipative smoothing will make the new model numerically much better behaved.

5.3. *Second-order LSG models*

The derivation of the second-order transformation that yields an affine L_2 Lagrangian is substantially more involved, and therefore relegated to Appendix C.

Our ansatz identifies four more naturally occurring free parameters α, β, γ and μ, and yields

$$\begin{aligned} L_2 &= L_{22}^{\text{deg}} - \int h(\boldsymbol{u}^{\perp} + \nabla h) \cdot \boldsymbol{v}'_{\text{free}} \, \mathrm{d}\boldsymbol{x} \\ &= \int h\boldsymbol{u} \cdot \left[\left(\alpha + \lambda^2 - \tfrac{1}{2}\right)\nabla^{\perp} h \Delta h + \left(\beta - \lambda + \tfrac{1}{4}\right) h \nabla^{\perp}\Delta h \right. \\ &\qquad \left. + (\gamma - 2\lambda + 1)\nabla^{\perp}\nabla^{\perp} h \nabla^{\perp} h + \mu h^{-1} \nabla^{\perp} h |\nabla h|^2\right] \mathrm{d}\boldsymbol{x} - H_2, \end{aligned} \tag{5.29}$$

where

$$\boldsymbol{v}'_{\text{free}} = \alpha \nabla h \Delta h + \beta h \nabla \Delta h + \gamma \nabla\nabla h \nabla h + \mu h^{-1} \nabla h |\nabla h|^2, \tag{5.30}$$

and

$$\begin{aligned} H_2 &= \int \left[(\lambda^2 - \beta) h^2 (\Delta h)^2 + \left(\lambda^2 + \alpha - 2\beta - \tfrac{\gamma}{2}\right) h |\nabla h|^2 \Delta h + \left(\mu - \tfrac{\gamma}{2}\right)|\nabla h|^4\right] \mathrm{d}\boldsymbol{x} \\ &= \int h\left[\left(\tfrac{2}{3}\lambda^2 - \tfrac{1}{3}\alpha - \tfrac{1}{3}\beta + \tfrac{1}{6}\gamma\right) h (\Delta h)^2 \right. \\ &\qquad + \left(\tfrac{1}{3}\lambda^2 + \tfrac{1}{3}\alpha - \tfrac{2}{3}\beta - \tfrac{1}{6}\gamma\right) h \nabla\nabla h : \nabla\nabla h \\ &\qquad \left. + \left(\mu - \tfrac{1}{3}\lambda^2 - \tfrac{1}{3}\alpha + \tfrac{2}{3}\beta - \tfrac{1}{3}\gamma\right) h^{-1} |\nabla h|^4\right] \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{5.31}$$

Our task is now to compute the equations of motion to second order, and to find 'good' choices for the parameters α, β, γ, λ and μ. However, before resuming the computation, we remark on two special choices that directly yield known models.

5.4. *Second-order L_1 dynamics*

When

$$\alpha = \tfrac{1}{4}, \quad \beta = \tfrac{1}{4}, \quad \gamma = 0, \quad \lambda = \tfrac{1}{2}, \quad \mu = 0, \tag{5.32}$$

the L_2 Lagrangian vanishes identically – the resulting dynamics in new coordinates is still Salmon's L_1 dynamics. However, the corresponding near-identity transformation back to 'physical' coordinates has a non-vanishing generating vector field at second order, namely

$$\begin{aligned} \boldsymbol{v}' = & -\tfrac{3}{4}\dot{\boldsymbol{u}} - \tfrac{3}{4}\nabla\boldsymbol{u}\boldsymbol{u} - \tfrac{1}{4}\nabla\boldsymbol{u}^\perp\boldsymbol{u}^\perp + \tfrac{1}{4}\nabla^\perp\nabla h\boldsymbol{u} + \tfrac{1}{4}\nabla\boldsymbol{u}^\perp\nabla h \\ & + \tfrac{1}{4}h\Delta\boldsymbol{u}^\perp + \tfrac{1}{4}\nabla h\Delta h + \tfrac{1}{4}h\nabla\Delta h, \end{aligned} \tag{5.33}$$

where we used identities (A 2), (A 3) and (A 4) to simplify the expression. We can now solve Salmon's L_1 equation of motion and then obtain a second-order *a posteriori* correction using (5.33).

5.5. *Second-order LSG*

We take $\lambda = -1/2$ as in Salmon's LSG model and require that there is no contribution to the potential vorticity at second order, i.e. the resulting symplectic structure is canonical. This necessitates the choice

$$\alpha = \tfrac{1}{4}, \quad \beta = -\tfrac{3}{4}, \quad \gamma = -2, \quad \mu = 0. \tag{5.34}$$

Substitution into (5.29) then gives

$$-L_2 = H_2 = \int h^2 \nabla\nabla h : \nabla\nabla h \, \mathrm{d}\boldsymbol{x}. \tag{5.35}$$

This is precisely the L_2 Lagrangian derived by Ford, Malham & Oliver (2002). In this earlier work, we had directly followed Salmon's procedure of first constraining to geostrophic balance and later transforming – in this case up to second order – to canonical coordinates. We then observed, as can also be seen from (5.35), that the second-order contribution to the Hamiltonian is positive, which can be shown to render the entire Hamiltonian positive definite. Although this appears to stabilize the dynamics, the kinematic potential vorticity inversion yields advecting velocity fields that are insufficiently smooth to generate a flow – both first- and second-order LSG are ill posed. This example nonetheless demonstrates that the formal steps of constraining and transforming up to a given asymptotic order commute.

5.6. *Second-order Euler–Poincaré equations*

We write the second-order Lagrangian (5.29) in the form

$$L_2 = \int h\boldsymbol{u} \cdot [\sigma_1 \boldsymbol{F}_1 + \sigma_2 \boldsymbol{F}_2 + \sigma_3 \boldsymbol{F}_3 + \sigma_4 \boldsymbol{F}_4] \, \mathrm{d}\boldsymbol{x} - H_2, \tag{5.36a}$$

$$H_2 = \int h[\rho_1 g_1 + \rho_2 g_2 + \rho_3 g_3] \, \mathrm{d}\boldsymbol{x}, \tag{5.36b}$$

where

$$\boldsymbol{F}_1(h) = \nabla^\perp h \Delta h, \tag{5.37a}$$

$$\boldsymbol{F}_2(h) = h \nabla^\perp \Delta h, \tag{5.37b}$$

$$\boldsymbol{F}_3(h) = \nabla\nabla^{\perp}h\nabla h, \tag{5.37c}$$
$$\boldsymbol{F}_4(h) = h^{-1}\nabla^{\perp}h|\nabla h|^2, \tag{5.37d}$$

and

$$g_1(h) = h(\Delta h)^2, \tag{5.38a}$$
$$g_2(h) = h\nabla\nabla h : \nabla\nabla h, \tag{5.38b}$$
$$g_3(h) = h^{-1}|\nabla h|^4. \tag{5.38c}$$

By direct calculation,

$$\mathrm{D}\boldsymbol{F}_1(h)\phi = \nabla^{\perp}h\Delta\phi + \nabla^{\perp}\phi\Delta h, \tag{5.39a}$$
$$\mathrm{D}\boldsymbol{F}_2(h)\phi = h\nabla^{\perp}\Delta\phi + \phi\nabla^{\perp}\Delta h, \tag{5.39b}$$
$$\mathrm{D}\boldsymbol{F}_3(h)\phi = \nabla\nabla^{\perp}h\nabla\phi + \nabla\nabla^{\perp}\phi\nabla h, \tag{5.39c}$$
$$\mathrm{D}\boldsymbol{F}_4(h)\phi = 2h^{-1}\nabla^{\perp}h\nabla h\cdot\nabla\phi + h^{-1}\nabla^{\perp}\phi|\nabla h|^2 - \phi h^{-2}\nabla^{\perp}h|\nabla h|^2, \tag{5.39d}$$

and therefore

$$\mathrm{D}\boldsymbol{F}_1^*(h)\cdot\boldsymbol{w} = \Delta(\boldsymbol{w}\cdot\nabla^{\perp}h) - \nabla^{\perp}\cdot(\boldsymbol{w}\Delta h), \tag{5.40a}$$
$$\mathrm{D}\boldsymbol{F}_2^*(h)\cdot\boldsymbol{w} = \Delta\nabla\cdot(h\boldsymbol{w}^{\perp}) - \boldsymbol{w}^{\perp}\cdot\nabla\Delta h, \tag{5.40b}$$
$$\mathrm{D}\boldsymbol{F}_3^*(h)\cdot\boldsymbol{w} = \nabla\cdot(\nabla\nabla h\boldsymbol{w}^{\perp}) - \nabla\nabla : (\nabla h\otimes\boldsymbol{w}^{\perp}), \tag{5.40c}$$
$$\mathrm{D}\boldsymbol{F}_4^*(h)\cdot\boldsymbol{w} = \nabla\cdot(2h^{-1}\nabla h\boldsymbol{w}^{\perp}\cdot\nabla h + \boldsymbol{w}^{\perp}h^{-1}|\nabla h|^2) + h^{-2}\boldsymbol{w}^{\perp}\cdot\nabla h|\nabla h|^2. \tag{5.40d}$$

We can now plug these expressions into the respective terms of the Euler–Poincaré equation (4.14). For $\boldsymbol{F}_1$, we obtain

$$\begin{aligned} S_1 &\equiv \nabla(\mathrm{D}\boldsymbol{F}_1^*(h)\cdot(h\boldsymbol{u})) - \mathrm{D}\boldsymbol{F}_1(h)\dot{h} - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot\boldsymbol{F}_1 \\ &= \nabla\nabla\cdot(h\boldsymbol{u}^{\perp}\Delta h) - \nabla\Delta(h\boldsymbol{u}^{\perp}\cdot\nabla h) + \nabla^{\perp}h\Delta\nabla\cdot(h\boldsymbol{u}) \\ &\quad + \nabla^{\perp}\nabla\cdot(h\boldsymbol{u})\Delta h - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot(\nabla^{\perp}h\Delta h). \end{aligned} \tag{5.41}$$

Similarly,

$$\begin{aligned} S_2 &\equiv \nabla(\mathrm{D}\boldsymbol{F}_2^*(h)\cdot(h\boldsymbol{u})) - \mathrm{D}\boldsymbol{F}_2(h)\dot{h} - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot\boldsymbol{F}_2 \\ &= \nabla\Delta\nabla\cdot(h^2\boldsymbol{u}^{\perp}) - \nabla(h\boldsymbol{u}^{\perp}\cdot\nabla\Delta h) + h\nabla^{\perp}\Delta\nabla\cdot(h\boldsymbol{u}) \\ &\quad + \nabla\cdot(h\boldsymbol{u})\nabla^{\perp}\Delta h - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot(h\nabla^{\perp}\Delta h), \end{aligned} \tag{5.42}$$

and

$$\begin{aligned} S_3 &\equiv \nabla(\mathrm{D}\boldsymbol{F}_3^*(h)\cdot(h\boldsymbol{u})) - \mathrm{D}\boldsymbol{F}_3(h)\dot{h} - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot\boldsymbol{F}_3 \\ &= \nabla\nabla\cdot(h\nabla\nabla h\boldsymbol{u}^{\perp}) - \nabla\nabla\nabla : (h\nabla h\otimes\boldsymbol{u}^{\perp}) + \nabla\nabla^{\perp}h\nabla(\nabla\cdot(h\boldsymbol{u})) \\ &\quad + \nabla\nabla^{\perp}(\nabla\cdot(h\boldsymbol{u}))\nabla h - \boldsymbol{u}^{\perp}\nabla^{\perp}\cdot(\nabla\nabla^{\perp}h\nabla h). \end{aligned} \tag{5.43}$$

There is a similar expression for S_4 which is not used in the following.

The corresponding computation for the 'energy' terms yields

$$\mathrm{D}g_1(h)\phi = \phi(\Delta h)^2 + 2h\Delta h\Delta\phi, \tag{5.44a}$$
$$\mathrm{D}g_2(h)\phi = \phi\nabla\nabla h : \nabla\nabla h + 2h\nabla\nabla h : \nabla\nabla\phi, \tag{5.44b}$$
$$\mathrm{D}g_3(h)\phi = -h^{-2}\phi|\nabla h|^4 + 4h^{-1}|\nabla h|^2\nabla h\cdot\nabla\phi, \tag{5.44c}$$

and

$$\mathrm{D}g_1^*(h)\psi = \psi(\Delta h)^2 + 2\Delta(h\psi\Delta h), \tag{5.45a}$$
$$\mathrm{D}g_2^*(h)\psi = \psi\nabla\nabla h : \nabla\nabla h + 2\nabla\nabla(h\psi : \nabla\nabla h), \tag{5.45b}$$

$$\mathrm{D}g_3^*(h)\psi = -h^{-2}\psi|\nabla h|^4 - 4\nabla(h^{-1}\psi\nabla h|\nabla h|^2). \tag{5.45c}$$

The corresponding terms on the right-hand side of the Euler–Poincaré equation are

$$r_1 \equiv \nabla\big(\mathrm{D}g_1^*(h)h + g_1\big) = 2\nabla(h(\Delta h)^2 + \Delta(h^2\Delta h)), \tag{5.46}$$

$$r_2 \equiv \nabla\big(\mathrm{D}g_2^*(h)h + g_2\big) = 2\nabla(h\nabla\nabla h : \nabla\nabla h + \nabla\nabla : (h^2\nabla\nabla h)), \tag{5.47}$$

$$r_3 \equiv \nabla\big(\mathrm{D}g_3^*(h)h + g_3\big) = \nabla\nabla\cdot(\nabla h|\nabla h|^2). \tag{5.48}$$

5.7. L_2 *dynamics*

Since Salmon's L_1 dynamics is characterized by the transformation reducing to the identity up to terms of $O(\varepsilon)$, it is natural to define an L_2 dynamics by imposing that the transformation reduces to the identity up to terms of $O(\varepsilon^2)$. In other words, we demand that

$$\boldsymbol{v}_\varepsilon \equiv \boldsymbol{v} + \varepsilon\boldsymbol{v}' = O(\varepsilon^2) \tag{5.49}$$

when the implicit $\boldsymbol{u}$ dependence of this expansion is expressed by a consistent diagnostic relationship, which we derive in the following. Since

$$\boldsymbol{u} = \nabla^\perp h + O(\varepsilon), \tag{5.50}$$

we must set, as for Salmon's L_1 dynamics, $\lambda = 1/2$ to remove $O(1)$ terms from (5.49). Inserting this choice and the diagnostic relationship (5.50) into the first-order Euler–Poincaré equation (4.27), we obtain the next order diagnostic relationship

$$\begin{aligned}\boldsymbol{u} &= \nabla^\perp h - \varepsilon\big[\nabla^\perp\big(h\Delta h + \tfrac{1}{2}|\nabla h|^2\big) - h\Delta\nabla^\perp h - 2\nabla h\cdot\nabla\nabla^\perp h\big] + O(\varepsilon^2)\\ &= \nabla^\perp h - \varepsilon\big[\nabla\nabla^\perp h\nabla h - \nabla^\perp h\Delta h\big] + O(\varepsilon^2).\end{aligned} \tag{5.51}$$

Similarly, we diagnose

$$\begin{aligned}\boldsymbol{v}' &= \boldsymbol{v}'_{\text{free}} - \tfrac{3}{4}\nabla^\perp\dot{h} + \tfrac{1}{4}\nabla\nabla h\nabla h + \nabla^\perp\nabla^\perp h\nabla h - \tfrac{1}{4}\nabla h\Delta h - \tfrac{1}{4}h\nabla\Delta h + O(\varepsilon)\\ &= \boldsymbol{v}'_{\text{free}} + \tfrac{3}{4}\nabla h\Delta h - \nabla\nabla h\nabla h - \tfrac{1}{4}h\nabla\Delta h,\end{aligned} \tag{5.52}$$

so that, altogether,

$$\begin{aligned}\boldsymbol{v}_\varepsilon &= \boldsymbol{v} + \varepsilon\boldsymbol{v}'\\ &= \tfrac{1}{2}(\boldsymbol{u}^\perp + \nabla h) + \varepsilon\boldsymbol{v}'\\ &= \varepsilon\big[\boldsymbol{v}'_{\text{free}} + \tfrac{5}{4}\nabla h\Delta h - \tfrac{3}{2}\nabla\nabla h\nabla h - \tfrac{1}{4}h\nabla\Delta h\big] + O(\varepsilon^2).\end{aligned} \tag{5.53}$$

Thus, for this diagnostic relation to vanish at $O(\varepsilon)$, we must require that

$$\alpha = -\tfrac{5}{4}, \quad \beta = \tfrac{1}{4}, \quad \gamma = \tfrac{3}{2}, \quad \mu = 0. \tag{5.54}$$

The corresponding coefficients for the second-order contributions to $\boldsymbol{F}$ are

$$\sigma_1 = -\tfrac{3}{2}, \quad \sigma_2 = 0, \quad \sigma_3 = \tfrac{3}{2}, \quad \sigma_4 = 0, \tag{5.55}$$

and therefore the full second-order contribution to $\boldsymbol{F}$ reads

$$\tfrac{3}{2}(-\nabla^\perp h\Delta h + \nabla^\perp\nabla^\perp h\nabla^\perp h) = -\tfrac{3}{2}\nabla\nabla h\nabla^\perp h. \tag{5.56}$$

The full L_2 potential vorticity is thus given by

$$q = \frac{\nabla^\perp\cdot\boldsymbol{F}}{h} = \frac{1 + \varepsilon\Delta h - \frac{3}{2}\varepsilon^2\nabla\nabla h : \nabla^\perp\nabla^\perp h}{h} = \frac{1 + \varepsilon\Delta h - 3\varepsilon^2\det\mathrm{Hess}h}{h}, \tag{5.57}$$

where the numerator is a second-order elliptic Monge–Ampère operator (see Lychagin, Rubtsov & Chekalov 1993).

The second-order contributions to the left-hand side of the Euler–Poincaré equation (4.14) are

$$\tfrac{3}{2}\nabla\nabla\cdot(\nabla^{\perp}\boldsymbol{u}\nabla h^2) - 3\nabla\nabla h\nabla^{\perp}(\nabla\cdot(h\boldsymbol{u})) + \tfrac{3}{2}\boldsymbol{u}^{\perp}\nabla\nabla h : \nabla^{\perp}\nabla^{\perp}h. \tag{5.58}$$

Similarly, we find that the coefficients corresponding to the components of the H_2 Hamiltonian (5.36*b*) are

$$\rho_1 = \tfrac{3}{4}, \quad \rho_2 = -\tfrac{3}{4}, \quad \rho_3 = 0, \tag{5.59}$$

so that the second-order contributions to the right-hand side of the Euler–Poincaré equation (4.14) are

$$-\tfrac{3}{2}\nabla[h\nabla\nabla h : \nabla^{\perp}\nabla^{\perp}h + \nabla\nabla : (h^2\nabla^{\perp}\nabla^{\perp}h)]. \tag{5.60}$$

Unfortunately, the resulting equation for $\boldsymbol{u}$ in terms of h is third order, not elliptic, and cannot be written as an operator solely acting on $\boldsymbol{u}^{\perp}$. The natural generalization, in our framework, of the L_1 setting therefore does not appear to yield a useful model. However, if we are prepared to make further approximations, consistent with the order of the model, we might be able to remove all of the 'bad' terms at the likely expense of losing the Hamiltonian structure.

5.8. *Order limitations*

We now ask more generally what order of potential vorticity inversion can be expected from an optimal choice of parameters. There are three competing considerations: the order of differentiation on the right-hand side of the Euler–Poincaré equation, ellipticity and regularity of the operator on the left-hand side of the Euler–Poincaré equation, and ellipticity and regularity of the q–h inversion.

Since the left-hand side of the Euler–Poincaré equation and the q–h inversion are each fourth-order at best, improving on the third-order regularity of potential vorticity inversion of the first-order model with $\lambda = 0$ requires that the right-hand side of the Euler–Poincaré equation does not contain derivatives of the maximum order five. These terms come from the symmetric second-order term in the H_2 Hamiltonian (5.31). We must hence require its coefficient to vanish, i.e. $\beta = \lambda^2$ or, in the notation of § 5.6, $\rho_1 + \rho_2 = 0$. However, this choice immediately implies that $\sigma_2 = (\lambda - 1/2)^2 \geqslant 0$. We note that σ_2 is the coefficient multiplying S_2 on the left-hand side of the Euler–Poincaré equation, which contains all possible fourth-order terms on $\boldsymbol{u}$. Dropping all lower-order contributions, these terms are

$$\sigma_2 h^2(\Delta\nabla\nabla\cdot\boldsymbol{u}^{\perp} + \Delta\nabla^{\perp}\nabla\cdot\boldsymbol{u}) = \sigma_2 h^2\Delta^2\boldsymbol{u}^{\perp}. \tag{5.61}$$

After being left-multiplied by $\boldsymbol{J}$, this expression enters the Euler–Poincaré equations with a negative sign, causing the combined operator to lose positivity unless $\sigma_2 = 0$. In the latter case, however, the operator on the left-hand side can be elliptic of order two at best, and the q–h inversion also can not reach maximal order.

We conclude that none of the second-order models will be able to exceed the degree of smoothness afforded by the most regular first-order model. However, this does not mean second-order models cannot be accurate – this question is entirely open to investigation. Moreover, when splitting off the ageostrophic velocity component, cancellations of higher-order terms similar to those in (4.30) occur provided that

$$\beta = 2\lambda^2 - \lambda + \tfrac{1}{4}, \tag{5.62}$$

with the possibility that the ageostrophic velocity may be smoother than the overall velocity field. Moreover, further approximation may also result in second-order accurate as well as regular models.

6. The quasi-geostrophic hierarchy

The shallow-water Lagrangian in quasi-geostrophic scaling is

$$L_\varepsilon = \int \left[\boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon + \tfrac{1}{2}\varepsilon |\dot{\boldsymbol{\eta}}_\varepsilon|^2 - \tfrac{1}{2}\varepsilon^{-1} h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \right] \mathrm{d}\boldsymbol{a}. \tag{6.1}$$

If we expand the quasi-geostrophic Lagrangian in powers of ε, the term at $O(\varepsilon^{-1})$ reads

$$L_{-1} = -\tfrac{1}{2} \int h \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}. \tag{6.2}$$

Taking arbitrary variations on any finite subdomain forces $h = 1$, i.e. the flow is incompressible. We now seek new coordinates in which $h = 1$ to all orders. Thus, the transformation cannot be area preserving, and we will be able to recover the weakly compressible effects of the parent dynamics by changing back into physical coordinates *a posteriori*. Thus, for a model in the quasi-geostrophic hierarchy, the continuity equation will always be trivial, while the momentum equation, once higher-order terms are included, remains prognostic. This should be contrasted with the approach taken in the LSG hierarchy, where the leading-order defining feature is that the Lagrangian is affine. This feature of the leading order was then imposed on the higher-order Lagrangians, resulting in a kinematic relationship between h and $\boldsymbol{u}$, while the continuity equation remained a prognostic equation. In each case, the shallow-water system is reduced to a single prognostic equation.

Once incompressibility is imposed, the L_{-1} contribution can be normalized out. Collecting terms at the remaining orders gives

$$L_\varepsilon = L_0 + \varepsilon L_1 + \tfrac{1}{2}\varepsilon^2 L_2 + O(\varepsilon^3), \tag{6.3}$$

with

$$L_0 = \int \left[\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \tfrac{1}{2}(h \nabla \cdot \boldsymbol{v}) \circ \boldsymbol{\eta} \right] \mathrm{d}\boldsymbol{a}, \tag{6.4}$$

$$L_1 = \int \left[\boldsymbol{v}^\perp \cdot \boldsymbol{u} + \tfrac{1}{2}|\boldsymbol{u}|^2 + \tfrac{1}{4} h (\nabla \cdot \boldsymbol{v}' + \boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} - (\nabla \cdot \boldsymbol{v})^2) \right] \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}, \tag{6.5}$$

$$\begin{aligned} L_2 = \int \Big[& \boldsymbol{u} \cdot (\boldsymbol{v}' + \nabla \boldsymbol{v} \boldsymbol{v})^\perp + (\boldsymbol{v}^\perp + 2\boldsymbol{u}) \cdot (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u}) \\ & + \tfrac{1}{6} h \big(\nabla \cdot \boldsymbol{v}'' + 2\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}' + \boldsymbol{v}' \cdot \nabla\nabla \cdot \boldsymbol{v} - 3 \nabla \cdot \boldsymbol{v} \nabla \cdot \boldsymbol{v}' \\ & + \boldsymbol{v} \cdot \nabla(\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}) - 3 \nabla \cdot \boldsymbol{v} \boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} + (\nabla \cdot \boldsymbol{v})^3 \big) \Big] \circ \boldsymbol{\eta} \, \mathrm{d}\boldsymbol{a}. \end{aligned} \tag{6.6}$$

Incompressibility also allows us to considerably simplify the expanded Lagrangians. Changing to Eulerian variables, eliminating perfect derivatives, and integrating by parts in various terms, we find

$$L_0 = \int \boldsymbol{R} \cdot \boldsymbol{u} \, \mathrm{d}\boldsymbol{x}, \tag{6.7}$$

$$L_1 = \int \left[\boldsymbol{v}^\perp \cdot \boldsymbol{u} + \tfrac{1}{2}|\boldsymbol{u}|^2 - \tfrac{1}{2}(\nabla \cdot \boldsymbol{v})^2 \right] \mathrm{d}\boldsymbol{x}, \tag{6.8}$$

$$\begin{aligned} L_2 &= \int \left[\boldsymbol{u} \cdot (\boldsymbol{v}' + \nabla \boldsymbol{v} \boldsymbol{v})^\perp + (\boldsymbol{v}^\perp + 2\boldsymbol{u}) \cdot (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u}) - \nabla \cdot \boldsymbol{v} \nabla \cdot \boldsymbol{v}' + \tfrac{1}{2}(\nabla \cdot \boldsymbol{v})^3 \right] \mathrm{d}\boldsymbol{x} \\ &= \int \left[(\nabla \nabla \cdot \boldsymbol{v} - \boldsymbol{u}^\perp) \cdot \boldsymbol{v}' + 2\boldsymbol{u} \cdot \nabla \boldsymbol{v} \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{v}^\perp \nabla \cdot \boldsymbol{v} + \tfrac{1}{2}(\nabla \cdot \boldsymbol{v})^3 \right] \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{6.9}$$

where, in the last equality, we have used identity (A 7).

Variations of the L_0 Lagrangian (6.7) simply confirm that $\boldsymbol{u}$ is divergence free. We also note that the quasi-geostrophic scaling has $\boldsymbol{v}$ appear at $O(1)$ – the change of variables is no longer small. In the variational principle, however, this contribution is lost as a result of imposing incompressibility, and consequently, leading-order geostrophic balance is lost.

In the next section, we show that geostrophic balance can be restored through conditions on $\boldsymbol{v}$, $\boldsymbol{v}'$, etc. from independent considerations.

6.1. *Balance conditions*

We first note that h_ε satisfies a continuity equation with respect to the change of variables,

$$h_\varepsilon' + \nabla \cdot (h_\varepsilon \boldsymbol{v}_\varepsilon) = 0, \tag{6.10}$$

as a direct consequence of the definitions for h_ε and $\boldsymbol{v}_\varepsilon$. Differentiating (6.10) and setting $\varepsilon = 0$, we obtain

$$h' + \nabla \cdot \boldsymbol{v} = 0, \tag{6.11}$$

$$h'' + \nabla \cdot \boldsymbol{v}' - \nabla \cdot (\boldsymbol{v} \nabla \cdot \boldsymbol{v}) = 0, \tag{6.12}$$

where once more we have used that $h = 1$ in the quasi-geostrophic scaling. Similarly, noting that $\dot{\boldsymbol{\eta}}_\varepsilon = \boldsymbol{u}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon$ and $\boldsymbol{\eta}_\varepsilon' = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon$, we find by cross-differentiation that

$$\boldsymbol{u}' + \nabla \boldsymbol{u} \boldsymbol{v} = \dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u}. \tag{6.13}$$

We now substitute the power series expansions

$$\boldsymbol{u}_\varepsilon = \boldsymbol{u} + \varepsilon \boldsymbol{u}' + O(\varepsilon^2), \tag{6.14}$$

$$h_\varepsilon = 1 + \varepsilon h' + \tfrac{1}{2}\varepsilon^2 h'' + O(\varepsilon^3), \tag{6.15}$$

into the quasigeostrophically rescaled shallow-water equations and collect terms at each power of ε. At order ε^0, we find

$$\boldsymbol{u} = \nabla^\perp h' = -\nabla^\perp \nabla \cdot \boldsymbol{v}, \tag{6.16}$$

or

$$\boldsymbol{v} = -\nabla \Delta^{-2} \nabla^\perp \cdot \boldsymbol{u}. \tag{6.17}$$

At order ε, the balance condition is

$$\dot{\boldsymbol{u}} + \nabla \boldsymbol{u} \boldsymbol{u} + \boldsymbol{u}'^\perp + \tfrac{1}{2}\nabla h'' = 0. \tag{6.18}$$

We eliminate $\boldsymbol{u}'$ and h'' via (6.13) and (6.12), respectively, whence

$$\dot{\boldsymbol{u}} + \nabla \boldsymbol{u} \boldsymbol{u} + (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u} - \nabla \boldsymbol{u} \boldsymbol{v})^\perp - \tfrac{1}{2}\nabla \nabla \cdot \boldsymbol{v}' + \tfrac{1}{2}\nabla \nabla \cdot (\boldsymbol{v} \nabla \cdot \boldsymbol{v}) = 0. \tag{6.19}$$

The divergence of this expression then yields

$$\tfrac{1}{2}\Delta \nabla \cdot \boldsymbol{v}' = \nabla \boldsymbol{u} : \nabla \boldsymbol{u}^T + \nabla \boldsymbol{v}^\perp : \nabla \boldsymbol{u}^T - \nabla \boldsymbol{u}^\perp : \nabla \boldsymbol{v}^T + \boldsymbol{v} \cdot \nabla \nabla^\perp \cdot \boldsymbol{u} + \tfrac{1}{2}\Delta \nabla \cdot (\boldsymbol{v} \nabla \cdot \boldsymbol{v}). \tag{6.20}$$

We remark that a first-order balance condition can also be derived variationally. Take, for example, arbitrary variations of the full compressible L_0 Lagrangian (6.4) with $\boldsymbol{v}$ fixed. The resulting condition reduces to (6.16) for $h = 1$.

6.2. *First-order quasigeostrophy*

Collecting terms to first order, the truncated Lagrangian reads

$$L = \int \left[\boldsymbol{R} \cdot \boldsymbol{u} + \varepsilon \left(\boldsymbol{v}^\perp \cdot \boldsymbol{u} + \tfrac{1}{2}|\boldsymbol{u}|^2 - \tfrac{1}{2}(\nabla \cdot \boldsymbol{v})^2 \right) \right] \mathrm{d}\boldsymbol{x}. \tag{6.21}$$

Since $\boldsymbol{u}$ is divergence free in the new coordinates, it is convenient to set $\boldsymbol{u} = \nabla^\perp \psi$ for some streamfunction ψ. Similarly, noting that only the curl-free part of $\boldsymbol{v}$ contributes to the Lagrangian, we set $\boldsymbol{v} = \nabla \phi$. In this notation,

$$L = \int \left[\boldsymbol{R} \cdot \nabla^\perp \psi + \varepsilon \left(\nabla \phi \cdot \nabla \psi + \tfrac{1}{2}|\nabla \psi|^2 - \tfrac{1}{2}(\Delta \phi)^2 \right) \right] \mathrm{d}\boldsymbol{x}. \tag{6.22}$$

The leading-order balance condition (6.17) implies $\phi = -\Delta^{-1}\psi$, so that

$$\begin{aligned} L &= \int \left[\boldsymbol{R} \cdot \nabla^\perp \psi + \varepsilon \left(-\nabla \Delta^{-1} \psi \cdot \nabla \psi + \tfrac{1}{2}|\nabla \psi|^2 - \tfrac{1}{2}\psi^2 \right) \right] \mathrm{d}\boldsymbol{x} \\ &= \int \left(\boldsymbol{R} + \tfrac{1}{2}\varepsilon \left(\boldsymbol{u} - \Delta^{-1}\boldsymbol{u} \right) \right) \cdot \boldsymbol{u} \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{6.23}$$

and the potential vorticity equation reads

$$(\partial_t + \boldsymbol{u} \cdot \nabla)(1 + \varepsilon \nabla^\perp \cdot (\boldsymbol{u} - \Delta^{-1}\boldsymbol{u})) = 0, \tag{6.24}$$

or

$$(\partial_t + \nabla^\perp \psi \cdot \nabla)(\Delta \psi - \psi) = 0. \tag{6.25}$$

We have thus recovered the classical quasi-geostrophic potential vorticity equation (2.33). The variational formulation (6.23) has already been noted by Holm & Zeitlin (1998), but we believe that the constructive derivation is new.

We remark that the balance condition (6.16) is crucial to derive a meaningful model for rotating shallow water. Choosing $\phi = \psi$, for example, yields the Lagrangian

$$\begin{aligned} L &= \int \left[\boldsymbol{R} \cdot \nabla^\perp \psi + \varepsilon \left(\nabla \psi \cdot \nabla \psi + \tfrac{1}{2}|\nabla \psi|^2 - \tfrac{1}{2}(\Delta \psi)^2 \right) \right] \mathrm{d}\boldsymbol{x} \\ &= \int \left(\boldsymbol{R} + \varepsilon \left(\tfrac{3}{2}\boldsymbol{u} - \tfrac{1}{2}\Delta \boldsymbol{u} \right) \cdot \boldsymbol{u} \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{6.26}$$

and the resulting potential vorticity equation reads

$$(\partial_t + \nabla^\perp \psi \cdot \nabla)\Delta \left(\psi - \tfrac{1}{3}\Delta \psi \right) = 0. \tag{6.27}$$

This corresponds to the Lagrangian-averaged Euler equations with $\alpha^2 = 1/3$, see Holm *et al.* (1998), Oliver & Shkoller (2001), and references cited therein, which even at leading order describe entirely different physics.

6.3. *Second-order quasigeostrophy*

To obtain the second order of the quasi-geostrophic hierarchy, we first substitute the leading-order balance condition into the second-order quasi-geostrophic Lagrangian (6.9). It is most convenient to work in terms of $\boldsymbol{v}$ rather than $\boldsymbol{u}$, so that we use the balance condition in the form (6.16), and obtain

$$L_2 = \int \left[2\boldsymbol{u} \cdot \nabla \boldsymbol{v} \boldsymbol{u} - \nabla \cdot \boldsymbol{v} \boldsymbol{v}^\perp \cdot \nabla^\perp \nabla \cdot \boldsymbol{v} + \tfrac{1}{2}(\nabla \cdot \boldsymbol{v})^3\right] \mathrm{d}\boldsymbol{x}$$
$$= \int \left[2\boldsymbol{u} \otimes \boldsymbol{u} : \nabla \boldsymbol{v} + (\nabla \cdot \boldsymbol{v})^3\right] \mathrm{d}\boldsymbol{x}. \tag{6.28}$$

The contribution involving $\boldsymbol{v}'$ has dropped entirely from the Lagrangian – we need the second-order balance condition only for the transformation back into the old coordinate system.

To derive the potential vorticity at order ε^2, it is easiest to take directly variations of (6.28), which are again subject to the Lin constraint (4.3). Since $\boldsymbol{v}$ is curl free, the matrix $\nabla \boldsymbol{v}$ is symmetric, so that

$$\delta L_2 = \int [4\delta\boldsymbol{u} \otimes \boldsymbol{u} : \nabla \boldsymbol{v} + 2\boldsymbol{u} \otimes \boldsymbol{u} : \nabla \delta \boldsymbol{v} + 3(\nabla \cdot \boldsymbol{v})^2 \nabla \cdot \delta \boldsymbol{v}] \, \mathrm{d}\boldsymbol{x}$$
$$= \int [4\delta \boldsymbol{u} \cdot \nabla \boldsymbol{v} \boldsymbol{u} - 2\boldsymbol{u} \otimes \boldsymbol{u} : \nabla\nabla \Delta^{-2} \nabla^\perp \cdot \delta \boldsymbol{u} - 3(\nabla \cdot \boldsymbol{v})^2 \Delta^{-1} \nabla^\perp \cdot \delta \boldsymbol{u}] \, \mathrm{d}\boldsymbol{x}$$
$$= \int \delta \boldsymbol{u} \cdot [4\nabla \boldsymbol{v} \boldsymbol{u} + 2\nabla^\perp \Delta^{-2} (\nabla \boldsymbol{u} : \nabla \boldsymbol{u}^T) + 3\nabla^\perp \Delta^{-1} (\nabla \cdot \boldsymbol{v})^2] \, \mathrm{d}\boldsymbol{x}$$
$$\equiv \int \delta \boldsymbol{u} \cdot \boldsymbol{F}_2 \, \mathrm{d}\boldsymbol{x}. \tag{6.29}$$

Note that we used the leading-order balance condition (6.17) to substitute for $\delta \boldsymbol{v}$ in the second step, and integrated by parts in the third. Hence, the order ε^2 contribution to the potential vorticity $\nabla^\perp \cdot \boldsymbol{F}_\varepsilon$, where

$$\boldsymbol{F}_\varepsilon = \boldsymbol{F}_0 + \varepsilon \boldsymbol{F}_1 + \tfrac{1}{2}\varepsilon^2 \boldsymbol{F}_2, \tag{6.30}$$

must be

$$\nabla^\perp \cdot \boldsymbol{F}_2 = 4\nabla \boldsymbol{v} : (\nabla^\perp \boldsymbol{u})^T + 2\Delta^{-1}(\nabla \boldsymbol{u} : \nabla \boldsymbol{u}^T) + 3(\nabla \cdot \boldsymbol{v})^2. \tag{6.31}$$

Using $\boldsymbol{u} = \nabla^\perp \psi$ and $\boldsymbol{v} = -\nabla \Delta^{-1} \psi$, we can also write

$$\nabla^\perp \cdot \boldsymbol{F}_2 = 3\psi^2 - 4\nabla^\perp \nabla^\perp \psi : \nabla\nabla \Delta^{-1} \psi - 2\Delta^{-1}(\nabla^\perp \nabla^\perp \psi : \nabla\nabla\psi). \tag{6.32}$$

We see that potential vorticity inversion is now nonlinear, but its regularity cannot be of higher order than that of the standard quasi-geostrophic model. Since the operator is not obviously elliptic, well-posedness of the second-order model remains open.

We note that there are no obvious free parameters in the quasi-geostrophic hierarchy, even for models beyond order two.

7. The semi-geostrophic hierarchy

We finally seek to identify the Hoskins semi-geostrophic equations and higher-order generalizations thereof within our variational framework. It may see natural to conjecture that the semi-geostrophic equations in physical coordinates – before the Hoskins transformation is applied – can be recovered as a particular case of the second-order LSG hierarchy. (In fact, this conjecture provided the initial motivation for going to second order in § 5.) It turns out, however, that this is not the case, as can be seen by the following argument.

The semi-geostrophic potential vorticity in physical coordinates, given by (2.26), can be recovered from our general L_2 Lagrangian (5.29) by the unique choice of parameters

$$\alpha = \tfrac{3}{4}, \quad \beta = \tfrac{1}{4}, \quad \gamma = -\tfrac{1}{2}, \quad \mu = 0, \tag{7.1}$$

which determines the associated Hamiltonian completely. In particular, the second-order contribution to H_ε reads

$$H_2 = \tfrac{1}{4} \int \left[3h|\nabla h|^2 \Delta h + |\nabla h|^4\right] \mathrm{d}\boldsymbol{x}. \tag{7.2}$$

This Hamiltonian is not even sign definite, and clearly differs from the semi-geostrophic Hamiltonian (2.28). Thus, classical semigeostrophy cannot arise as a second-order LSG model in the sense of § 5. (Changing procedure, however, and imposing different constraints on the symplectic structure and on the Hamiltonian, we can indeed derive the semi-geostrophic equations as has been noted by McIntyre & Roulstone 2002.)

On the other hand, there are two key features of semigeostrophy written in Hoskins coordinates that we can replicate in our transformational approach. First, the symplectic structure is canonical, so that the potential vorticity is $q = 1/h$. Secondly, the transformed velocity is equal to the geostrophic velocity in old coordinates. In the following, we show that these two conditions can be applied at any order of the asymptotics. The challenge, however, is closing the equations in transformed coordinates beyond order two.

7.1. *General setting*

The key observation – implicit, for example, in Appendix B of Salmon (1985) – is that

$$\delta \int h_\varepsilon^2 \,\mathrm{d}\boldsymbol{x} = 2 \int h_\varepsilon \delta h_\varepsilon \,\mathrm{d}\boldsymbol{x} = -2 \int h_\varepsilon \nabla \cdot (h_\varepsilon \boldsymbol{w}_\varepsilon) \,\mathrm{d}\boldsymbol{x} = 2 \int h_\varepsilon \boldsymbol{w}_\varepsilon \cdot \nabla h_\varepsilon \,\mathrm{d}\boldsymbol{x}, \tag{7.3}$$

which, in Lagrangian coordinates, reads

$$\delta \int h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \,\mathrm{d}\boldsymbol{a} = 2 \int (\nabla h_\varepsilon) \circ \boldsymbol{\eta}_\varepsilon \cdot \delta \boldsymbol{\eta}_\varepsilon \,\mathrm{d}\boldsymbol{a}. \tag{7.4}$$

We now impose that the velocity in new coordinates is equal to the old geostrophic velocity,

$$\boldsymbol{u} = \nabla^\perp h_\varepsilon \circ \boldsymbol{\xi}_\varepsilon, \tag{7.5}$$

and therefore

$$\delta \int h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \,\mathrm{d}\boldsymbol{a} \equiv 2 \int \boldsymbol{u} \circ \boldsymbol{\eta} \cdot \delta \boldsymbol{\eta}_\varepsilon^\perp \,\mathrm{d}\boldsymbol{a}. \tag{7.6}$$

Hence, we proceed as follows. We take the variation of the full non-transformed action and apply (7.6). Only then do we expand all terms in powers of ε. We finally impose canonical coordinates by choosing $\boldsymbol{\eta}'$, $\boldsymbol{\eta}''$, etc. such that the variation of the action, when truncated to consistent order, is of the form

$$\delta S = -\iint \dot{\boldsymbol{\eta}}^\perp \cdot \delta \boldsymbol{\eta} \,\mathrm{d}\boldsymbol{a}\,\mathrm{d}t - \iint \boldsymbol{F}_\varepsilon \cdot \delta \boldsymbol{\eta} \,\mathrm{d}\boldsymbol{a}\,\mathrm{d}t. \tag{7.7}$$

The first term in this expression guarantees canonicity. The resulting Euler–Lagrange equations, in connection with (7.5), then tell us that

$$\boldsymbol{F}_\varepsilon = \nabla h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon. \tag{7.8}$$

The variations of each term in the action corresponding to the semigeostrophically scaled Lagrangian (5.1) are

$$\delta \iint \boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t = \iint \dot{\boldsymbol{\eta}}_\varepsilon \cdot \delta \boldsymbol{\eta}_\varepsilon^\perp \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t$$
$$= \iint \big[\dot{\boldsymbol{\eta}} \cdot \delta \boldsymbol{\eta}^\perp + \varepsilon (\dot{\boldsymbol{\eta}}' \cdot \delta \boldsymbol{\eta}^\perp + \dot{\boldsymbol{\eta}} \cdot \delta \boldsymbol{\eta}'^\perp)$$
$$+ \varepsilon^2 \big(\tfrac{1}{2} \dot{\boldsymbol{\eta}}'' \cdot \delta \boldsymbol{\eta}^\perp + \dot{\boldsymbol{\eta}}' \cdot \delta \boldsymbol{\eta}'^\perp + \tfrac{1}{2} \dot{\boldsymbol{\eta}} \cdot \delta \boldsymbol{\eta}''^\perp \big) \big] \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t + O(\varepsilon^3), \tag{7.9}$$

where, in the first equality, we have used identities similar to those applied in (B 15). Next,

$$\tfrac{1}{2} \varepsilon \delta \iint |\dot{\boldsymbol{\eta}}_\varepsilon|^2 \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t = -\varepsilon \iint \ddot{\boldsymbol{\eta}}_\varepsilon^\perp \cdot \delta \boldsymbol{\eta}_\varepsilon^\perp \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t$$
$$= - \iint [\varepsilon \ddot{\boldsymbol{\eta}}^\perp \cdot \delta \boldsymbol{\eta}^\perp + \varepsilon^2 (\ddot{\boldsymbol{\eta}}'^\perp \cdot \delta \boldsymbol{\eta}^\perp + \ddot{\boldsymbol{\eta}}^\perp \cdot \delta \boldsymbol{\eta}'^\perp)] \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t + O(\varepsilon^3), \tag{7.10}$$

and, using (7.6),

$$-\tfrac{1}{2} \delta \iint h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t = - \iint \boldsymbol{u} \circ \boldsymbol{\eta} \cdot \big[\delta \boldsymbol{\eta}^\perp + \varepsilon \delta \boldsymbol{\eta}'^\perp + \tfrac{1}{2} \varepsilon^2 \delta \boldsymbol{\eta}''^\perp \big] \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t + O(\varepsilon^3). \tag{7.11}$$

7.2. *First-order semigeostrophy*

We now look at each order in the variation of the action in turn. At leading order, we recover our ansatz, since

$$\delta S_0 = \iint (\dot{\boldsymbol{\eta}} - \boldsymbol{u} \circ \boldsymbol{\eta}) \cdot \delta \boldsymbol{\eta}^\perp \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t = 0. \tag{7.12}$$

At the next order,

$$\delta S_1 = \iint [(\dot{\boldsymbol{\eta}}' - \ddot{\boldsymbol{\eta}}^\perp) \cdot \delta \boldsymbol{\eta}^\perp + (\dot{\boldsymbol{\eta}} - \boldsymbol{u} \circ \boldsymbol{\eta}) \cdot \delta \boldsymbol{\eta}'^\perp] \, \mathrm{d}\boldsymbol{a} \, \mathrm{d}t \equiv 0. \tag{7.13}$$

Therefore, we need to impose that

$$\dot{\boldsymbol{\eta}} + \boldsymbol{\eta}'^\perp = 0. \tag{7.14}$$

Since, up to first order,

$$\boldsymbol{\xi}_\varepsilon \circ \boldsymbol{\eta} = \boldsymbol{\eta}_\varepsilon = \boldsymbol{\eta} + \varepsilon \boldsymbol{\eta}', \tag{7.15}$$

substituting (7.14) into this expression yields the classical Hoskins transformation

$$\boldsymbol{\xi}_\varepsilon = \mathrm{id} + \varepsilon \boldsymbol{u}^\perp. \tag{7.16}$$

Thus, we have recovered the semi-geostrophic equations without imposing the geostrophic momentum approximation, but simply by systematically truncating the Hoskins transformed variation of the action at first order. In other words, while Hoskins (1975) combined an exact transformation with an independently motivated approximation, our approximation lies entirely with the truncation of the transformation. With the conservation laws already contained in our ansatz, the non-obvious and perhaps surprising aspect of the semi-geostrophic equations is that they can be closed in the new semi-geostrophic coordinates as explained in § 2.3.

7.3. *Higher-order semigeostrophy*

At second order, we have

$$\delta S_2 = \iint \left[\left(\tfrac{1}{2}\dot{\boldsymbol{\eta}}'' - \ddot{\boldsymbol{\eta}}'^{\perp} \right) \cdot \delta\boldsymbol{\eta}^{\perp} + (\dot{\boldsymbol{\eta}}' - \ddot{\boldsymbol{\eta}}^{\perp}) \cdot \delta\boldsymbol{\eta}'^{\perp} + \tfrac{1}{2}(\dot{\boldsymbol{\eta}} - \boldsymbol{u} \circ \boldsymbol{\eta}) \cdot \delta\boldsymbol{\eta}''^{\perp} \right] \mathrm{d}\boldsymbol{a}\, \mathrm{d}t \equiv 0, \quad (7.17)$$

where, as before, only the term multiplying $\delta\boldsymbol{\eta}$ yields new information, and we find that

$$\tfrac{1}{2}\boldsymbol{\eta}'' - \dot{\boldsymbol{\eta}}'^{\perp} = 0 \quad (7.18)$$

and therefore

$$\tfrac{1}{2}\boldsymbol{\eta}'' + \ddot{\boldsymbol{\eta}} = 0. \quad (7.19)$$

The corresponding second-order transformation reads

$$\boldsymbol{\xi}_{\varepsilon} = \mathrm{id} + \varepsilon \boldsymbol{u}^{\perp} + \varepsilon^2 (\dot{\boldsymbol{u}} + \boldsymbol{u} \cdot \nabla \boldsymbol{u}). \quad (7.20)$$

Continuing this way, we find that

$$\boldsymbol{\xi}_{\varepsilon} \circ \boldsymbol{\eta} = \boldsymbol{\eta}_{\varepsilon} = \boldsymbol{\eta} + \varepsilon \dot{\boldsymbol{\eta}}^{\perp} - \varepsilon^2 \ddot{\boldsymbol{\eta}} - \varepsilon^3 \dddot{\boldsymbol{\eta}}^{\perp} + \varepsilon^4 \boldsymbol{\eta}^{(4)} + \dots . \quad (7.21)$$

We notice that t-derivatives of $\boldsymbol{u}$ start to appear, so that potential vorticity inversion is non-local in time, and cannot be done in any obvious way for models of order two or higher.

However, if we are prepared to make further approximations which potentially destroy the Hamiltonian structure, the equations can at least formally be closed. In particular, at second order, we can remove the time derivative by noting that

$$\dot{\boldsymbol{u}} = \nabla^{\perp} \dot{h} + O(\varepsilon) = O(\varepsilon). \quad (7.22)$$

However, at this level of approximation, the resulting system is not elliptic. Thus, although the generalized Hoskins transformation (7.21) has a very simple structure, it is not clear whether the corresponding models are useful or even well posed.

8. Discussion and outlook

We introduced a unified framework in which the classical balance models as well as new ones – of the same formal order of accuracy – can be derived by consistently truncating a near-identity change of coordinates in the variational formulation of the rotating shallow-water equations. Model reduction is achieved by imposing either degeneracy or incompressibility on the truncated expansion of the Lagrangian.

This approach has a number of advantages.

(i) Since all approximations are interpreted as arising through a change of coordinates, we have a formalism for *a posteriori* next order correction of numerically computed solutions.

(ii) We have derived several new models, at least one of which has promising analytical properties.

(iii) The unified formulation provides a framework for computational benchmarking of the different models against the full shallow-water parent model.

Future work may take a number of different directions, in particular the following.

(i) Inclusion of bottom topography, stratification, boundary conditions, and spatial variations in the Coriolis parameter.

(ii) Can more general choices than (5.11) for the first-order transformation yield interesting models or connections with more classical Hamiltonian approximation theory?

(iii) Investigation of the well-posedness of the reduced models and analytical estimates of the modelling error.

(iv) Numerical benchmarking of the different models.

(v) Investigation of connections to Lagrangian averaging (cf. Holm 1999; Marsden & Shkoller 2003).

(vi) Investigation of connections between our quasi-geostrophic hierarchy and the theory of nearly incompressible flow.

(vii) Systematic study of more interesting finite-dimensional models than the simple toy presented here.

Some of the key ideas in this paper have developed out of discussions with Rupert Ford, Georg Gottwald, Simon Malham and Matthew West. I also thank Onno Bokhove for a thorough reading of an earlier version of the manuscript, and Mike Cullen, David Dritschel, Michael McIntyre, Sebastian Reich, Guillaume Roullet, Ian Roulstone, and all participants of the EPSRC network 'Geometric methods in geophysical fluid dynamics' for numerous stimulating discussions, and the network organizers for the opportunity to participate. Finally, I thank the anonymous referees for their numerous very constructive comments.

Appendix A. Useful identities

For arbitrary, sufficiently smooth functions f, g and h, and an arbitrary vector field $\boldsymbol{u}$, the following identities hold:

$$\nabla^\perp g \boldsymbol{u} \cdot \nabla f - \nabla f \boldsymbol{u} \cdot \nabla^\perp g = \boldsymbol{u}^\perp \nabla f \cdot \nabla g, \tag{A 1}$$

$$\nabla\nabla \cdot \boldsymbol{u}^\perp + \nabla^\perp \nabla \cdot \boldsymbol{u} = \Delta \boldsymbol{u}^\perp, \tag{A 2}$$

$$\nabla h \nabla \cdot \boldsymbol{u}^\perp + (\nabla^\perp \boldsymbol{u})^T \nabla h = \nabla h \cdot \nabla \boldsymbol{u}^\perp, \tag{A 3}$$

$$\nabla^\perp h \nabla \cdot \boldsymbol{u} + (\nabla \boldsymbol{u}^\perp)^T \nabla h = \nabla h \cdot \nabla \boldsymbol{u}^\perp, \tag{A 4}$$

Further, with

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{A 5}$$

so that $\boldsymbol{J}\boldsymbol{u} = \boldsymbol{u}^\perp$,

$$\nabla \boldsymbol{u} - (\nabla \boldsymbol{u})^T = \boldsymbol{J} \nabla^\perp \cdot \boldsymbol{u}, \tag{A 6}$$

$$\nabla^\perp \boldsymbol{u}^\perp + (\nabla \boldsymbol{u})^T = \boldsymbol{I} \nabla \cdot \boldsymbol{u}. \tag{A 7}$$

Equations (A 6) and (A 7) imply that

$$\nabla\nabla^\perp h - \nabla^\perp \nabla h = \Delta h \boldsymbol{J}, \tag{A 8}$$

$$\nabla\nabla h + \nabla^\perp \nabla^\perp h = \Delta h \boldsymbol{I}, \tag{A 9}$$

and therefore, in particular,

$$\nabla\nabla^\perp h \nabla h + \nabla\nabla h \nabla^\perp h = \nabla^\perp h \Delta h. \tag{A 10}$$

All identities can easily be verified by direct calculation.

Appendix B. Details of the expansion

The expansions of each term in the shallow-water Lagrangian is most easily written in terms of the Eulerian vector fields $\boldsymbol{u}$ and $\boldsymbol{v}$. Thus, we first establish relationships between derivatives of the diffeomorphisms $\boldsymbol{\eta}_\varepsilon$ and $\boldsymbol{\xi}_\varepsilon$, and the corresponding vector fields $\boldsymbol{u}$ and $\boldsymbol{v}$. Differentiating $\boldsymbol{\xi}'_\varepsilon = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\xi}_\varepsilon$ with respect to t and ε, respectively, gives

$$\dot{\boldsymbol{\xi}}'_\varepsilon = \dot{\boldsymbol{v}}_\varepsilon \circ \boldsymbol{\xi}_\varepsilon + (\nabla \boldsymbol{v}_\varepsilon) \circ \boldsymbol{\xi}_\varepsilon \dot{\boldsymbol{\xi}}_\varepsilon, \tag{B 1}$$

$$\boldsymbol{\xi}''_\varepsilon = \boldsymbol{v}'_\varepsilon \circ \boldsymbol{\xi}_\varepsilon + (\nabla \boldsymbol{v}_\varepsilon) \circ \boldsymbol{\xi}_\varepsilon \boldsymbol{\xi}'_\varepsilon. \tag{B 2}$$

Setting $\varepsilon = 0$ and using that, by definition, $\boldsymbol{\xi} \equiv \boldsymbol{\xi}_0 = \mathrm{id}$ and therefore $\dot{\boldsymbol{\xi}} = 0$, we obtain

$$\boldsymbol{\xi}' = \boldsymbol{v}, \tag{B 3}$$

$$\dot{\boldsymbol{\xi}}' = \dot{\boldsymbol{v}}, \tag{B 4}$$

$$\boldsymbol{\xi}'' = \boldsymbol{v}' + \nabla \boldsymbol{v} \boldsymbol{v}. \tag{B 5}$$

(Quantities without subscript are taken to be evaluated at $\varepsilon = 0$.) Similarly, successive differentiation of $\boldsymbol{\eta}_\varepsilon = \boldsymbol{\xi}_\varepsilon \circ \boldsymbol{\eta}$ gives

$$\boldsymbol{\eta}'_\varepsilon = \boldsymbol{\xi}'_\varepsilon \circ \boldsymbol{\eta}, \tag{B 6}$$

$$\boldsymbol{\eta}''_\varepsilon = \boldsymbol{\xi}''_\varepsilon \circ \boldsymbol{\eta}, \tag{B 7}$$

$$\dot{\boldsymbol{\eta}}'_\varepsilon = \dot{\boldsymbol{\xi}}'_\varepsilon \circ \boldsymbol{\eta} + (\nabla \boldsymbol{\xi}'_\varepsilon) \circ \boldsymbol{\eta} \dot{\boldsymbol{\eta}}, \tag{B 8}$$

whence, setting $\varepsilon = 0$,

$$\boldsymbol{\eta}' = \boldsymbol{v} \circ \boldsymbol{\eta}, \tag{B 9}$$

$$\boldsymbol{\eta}'' = (\boldsymbol{v}' + \nabla \boldsymbol{v} \boldsymbol{v}) \circ \boldsymbol{\eta}, \tag{B 10}$$

$$\dot{\boldsymbol{\eta}}' = (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v} \boldsymbol{u}) \circ \boldsymbol{\eta}. \tag{B 11}$$

We now look at each term of the rotating shallow-water Lagrangian separately. First, consider the Coriolis term. Since f is constant, second derivatives of $\boldsymbol{R}$ vanish, and a straightforward Taylor expansion of $\boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon$ about $\varepsilon = 0$ gives

$$\boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon = \boldsymbol{R} \circ \boldsymbol{\eta} + \varepsilon (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' + \tfrac{1}{2} \varepsilon^2 (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}'' + O(\varepsilon^3). \tag{B 12}$$

Thus,

$$\begin{aligned} \boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon = {} & \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \varepsilon (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}} + \varepsilon \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}}' + \tfrac{1}{2} \varepsilon^2 \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}}'' \\ & + \tfrac{1}{2} \varepsilon^2 (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}'' \cdot \dot{\boldsymbol{\eta}} + \varepsilon^2 (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}}' + O(\varepsilon^3). \end{aligned} \tag{B 13}$$

We can pull out of this expression some full time derivatives which do not contribute to the variational principle. For any vector $\boldsymbol{w}$,

$$\partial_t (\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \boldsymbol{w}) = (\nabla \boldsymbol{R})^T \circ \boldsymbol{\eta} \boldsymbol{w} \cdot \dot{\boldsymbol{\eta}} + \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{w}}, \tag{B 14}$$

so that

$$\begin{aligned} \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{w}} + (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{w} \cdot \dot{\boldsymbol{\eta}} &= (\nabla \boldsymbol{R} - (\nabla \boldsymbol{R})^T) \circ \boldsymbol{\eta} \boldsymbol{w} \cdot \dot{\boldsymbol{\eta}} + \partial_t (\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \boldsymbol{w}) \\ &= \boldsymbol{w}^\perp \cdot \dot{\boldsymbol{\eta}} + \partial_t (\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \boldsymbol{w}). \end{aligned} \tag{B 15}$$

Similarly, we compute, again under the assumption that f is constant (when f is arbitrary, the additional terms that arise do not combine in the same way),

$$\partial_t ((\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \boldsymbol{\eta}') = (\nabla \boldsymbol{R})^T \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}}' + (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}}', \tag{B 16}$$

so that

$$\begin{aligned} (\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}}' &= \tfrac{1}{2} (\nabla \boldsymbol{R} - (\nabla \boldsymbol{R})^T) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \dot{\boldsymbol{\eta}}' + \tfrac{1}{2} \partial_t ((\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \boldsymbol{\eta}') \\ &= \tfrac{1}{2} \boldsymbol{\eta}'^\perp \cdot \dot{\boldsymbol{\eta}}' + \tfrac{1}{2} \partial_t ((\nabla \boldsymbol{R}) \circ \boldsymbol{\eta} \boldsymbol{\eta}' \cdot \boldsymbol{\eta}'). \end{aligned} \tag{B 17}$$

We now apply (B 15) with $\boldsymbol{w} = \boldsymbol{\eta}'$ and $\boldsymbol{w} = \boldsymbol{\eta}''$, respectively, and (B 17) to rewrite (B 13) as follows:

$$\begin{aligned}\boldsymbol{R} \circ \boldsymbol{\eta}_\varepsilon \cdot \dot{\boldsymbol{\eta}}_\varepsilon &= \boldsymbol{R} \circ \boldsymbol{\eta} \cdot \dot{\boldsymbol{\eta}} + \varepsilon \boldsymbol{\eta}'^{\perp} \cdot \dot{\boldsymbol{\eta}} + \varepsilon \partial_t (\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \boldsymbol{\eta}') + \tfrac{1}{2}\varepsilon^2 \boldsymbol{\eta}''^{\perp} \cdot \dot{\boldsymbol{\eta}} \\ &\quad + \tfrac{1}{2}\varepsilon^2 \partial_t (\boldsymbol{R} \circ \boldsymbol{\eta} \cdot \boldsymbol{\eta}'') + \tfrac{1}{2}\varepsilon^2 \boldsymbol{\eta}'^{\perp} \cdot \dot{\boldsymbol{\eta}}' + \tfrac{1}{2}\varepsilon^2 \partial_t ((\nabla \boldsymbol{R}) \circ \boldsymbol{\eta}\boldsymbol{\eta}' \cdot \boldsymbol{\eta}') + O(\varepsilon^3) \\ &= \left[\boldsymbol{R} \cdot \boldsymbol{u} + \varepsilon \boldsymbol{u} \cdot \boldsymbol{v}^{\perp} + \tfrac{1}{2}\varepsilon^2 (\boldsymbol{u} \cdot (\boldsymbol{v}' + \nabla \boldsymbol{v}\boldsymbol{v})^{\perp} + \boldsymbol{v}^{\perp} \cdot (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v}\boldsymbol{u}))\right] \circ \boldsymbol{\eta} \\ &\quad + O(\varepsilon^3) + \dot{F},\end{aligned} \tag{B 18}$$

where $\dot{F}$ is a total time derivative which does not contribute to the variational principle, and will be dropped hereinafter.

Next, the kinetic energy can be expanded directly,

$$\begin{aligned}\tfrac{1}{2}\varepsilon |\dot{\boldsymbol{\eta}}_\varepsilon|^2 &= \tfrac{1}{2}\varepsilon |\dot{\boldsymbol{\eta}} + \varepsilon \dot{\boldsymbol{\eta}}' + O(\varepsilon^2)|^2 \\ &= \tfrac{1}{2}\varepsilon |\dot{\boldsymbol{\eta}}|^2 + \varepsilon^2 \dot{\boldsymbol{\eta}} \cdot \dot{\boldsymbol{\eta}}' + O(\varepsilon^3) \\ &= \left[\tfrac{1}{2}\varepsilon |\boldsymbol{u}|^2 + \varepsilon^2 \boldsymbol{u} \cdot (\dot{\boldsymbol{v}} + \nabla \boldsymbol{v}\boldsymbol{u})\right] \circ \boldsymbol{\eta} + O(\varepsilon^3),\end{aligned} \tag{B 19}$$

where we have used identity (B 11) to substitute for $\dot{\boldsymbol{\eta}}'$.

Finally, the potential energy term is expanded by noting that (5.4) and (5.5) combine to $\boldsymbol{\eta}'_\varepsilon = \boldsymbol{v}_\varepsilon \circ \boldsymbol{\eta}_\varepsilon$. Setting $J_\varepsilon \equiv h_\varepsilon^{-1} \circ \boldsymbol{\eta}_\varepsilon$, the Liouville theorem for the flow of $\boldsymbol{v}_\varepsilon$ reads

$$J'_\varepsilon = (\nabla \cdot \boldsymbol{v}_\varepsilon) \circ \boldsymbol{\eta}_\varepsilon J_\varepsilon. \tag{B 20}$$

After differentiating with respect to ε, setting $\varepsilon = 0$ yields the relations

$$\begin{aligned}J' &= (\nabla \cdot \boldsymbol{v}) \circ \boldsymbol{\eta} J \\ &\equiv \sigma_1 J,\end{aligned} \tag{B 21}$$

$$\begin{aligned}J'' &= [\nabla \cdot \boldsymbol{v}' + \boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} + (\nabla \cdot \boldsymbol{v})^2] \circ \boldsymbol{\eta} J \\ &\equiv \sigma_2 J,\end{aligned} \tag{B 22}$$

$$\begin{aligned}J''' &= [\nabla \cdot \boldsymbol{v}'' + 2\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}' + \boldsymbol{v}' \cdot \nabla\nabla \cdot \boldsymbol{v} + 3\nabla \cdot \boldsymbol{v}\nabla \cdot \boldsymbol{v}' \\ &\qquad + \boldsymbol{v} \cdot \nabla(\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}) + 3\nabla \cdot \boldsymbol{v}\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} + (\nabla \cdot \boldsymbol{v})^3] \circ \boldsymbol{\eta} J \\ &\equiv \sigma_3 J.\end{aligned} \tag{B 23}$$

The power series,

$$J_\varepsilon = J\left[1 + \sigma_1 \varepsilon + \tfrac{1}{2}\sigma_2 \varepsilon^2 + \tfrac{1}{6}\sigma_3 \varepsilon^3 + O(\varepsilon^4)\right], \tag{B 24}$$

is easily inverted. Setting $J^{-1} \equiv h \circ \boldsymbol{\eta}$, we find

$$\begin{aligned}h_\varepsilon \circ \boldsymbol{\eta}_\varepsilon &= J_\varepsilon^{-1} \\ &= J^{-1}\left[1 - \sigma_1 \varepsilon + \left(\sigma_1^2 - \tfrac{1}{2}\sigma_2\right)\varepsilon^2 - \left(\sigma_1^3 - \sigma_1\sigma_2 + \tfrac{1}{6}\sigma_3\right)\varepsilon^3 + O(\varepsilon^4)\right] \\ &= h \circ \boldsymbol{\eta}\left[1 - \varepsilon \nabla \cdot \boldsymbol{v} - \tfrac{1}{2}\varepsilon^2 (\nabla \cdot \boldsymbol{v}' + \boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} - (\nabla \cdot \boldsymbol{v})^2)\right. \\ &\quad - \tfrac{1}{6}\varepsilon^3 (\nabla \cdot \boldsymbol{v}'' + 2\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}' + \boldsymbol{v}' \cdot \nabla\nabla \cdot \boldsymbol{v} - 3\nabla \cdot \boldsymbol{v}\nabla \cdot \boldsymbol{v}' \\ &\quad \left. + \boldsymbol{v} \cdot \nabla(\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v}) - 3\nabla \cdot \boldsymbol{v}\boldsymbol{v} \cdot \nabla\nabla \cdot \boldsymbol{v} + (\nabla \cdot \boldsymbol{v})^3) + O(\varepsilon^4)\right] \circ \boldsymbol{\eta}.\end{aligned} \tag{B 25}$$

Appendix C. Derivation of the second-order LSG transformation

The identification of transformation which renders the L_2 Lagrangian (5.9) affine requires some preparatory work. There are three distinct groups of terms which we consider separately – terms involving $\boldsymbol{v}'$, terms involving $\dot{\boldsymbol{v}}$, and all others:

$$\left.\begin{aligned} L_{21} &= \int h\big(\boldsymbol{u}\cdot\boldsymbol{v}'^{\perp} + \tfrac{1}{2}h\nabla\cdot\boldsymbol{v}'\big)\,\mathrm{d}\boldsymbol{x}, \\ L_{221} &= \int h(\boldsymbol{v}^{\perp} + 2\boldsymbol{u})\cdot\dot{\boldsymbol{v}}\mathrm{d}\boldsymbol{x}, \\ L_{222} &= \int h\big[\boldsymbol{u}\cdot\nabla\boldsymbol{v}^{\perp}\boldsymbol{v} + (\boldsymbol{v}^{\perp}+2\boldsymbol{u})\cdot\nabla\boldsymbol{v}\boldsymbol{u} + \tfrac{1}{2}h(\boldsymbol{v}\cdot\nabla\nabla\cdot\boldsymbol{v} - (\nabla\cdot\boldsymbol{v})^2)\big]\,\mathrm{d}\boldsymbol{x}. \end{aligned}\right\} \tag{C 1}$$

First, using integration by parts, we can write

$$L_{21} = -\int h(\boldsymbol{u}^{\perp} + \nabla h)\cdot\boldsymbol{v}'\,\mathrm{d}\boldsymbol{x}. \tag{C 2}$$

The other terms involve $\boldsymbol{v}$, so that we must insert our first-order ansatz (5.11). We begin by computing

$$\begin{aligned} L_{221} &= \int h\big[\big(\tfrac{3}{2}\boldsymbol{u} + \lambda\nabla^{\perp}h\big)\cdot\big(\tfrac{1}{2}\dot{\boldsymbol{u}}^{\perp} + \lambda\nabla\dot{h}\big)\big]\,\mathrm{d}\boldsymbol{x} \\ &= \int h\big[\big(\tfrac{3}{4}\boldsymbol{u} + \tfrac{1}{2}\lambda\nabla^{\perp}h\big)\cdot(\dot{\boldsymbol{u}}^{\perp} + \nabla\dot{h} + (2\lambda-1)\nabla\dot{h})\big]\,\mathrm{d}\boldsymbol{x} \\ &= \int h\big[\tfrac{3}{4}\boldsymbol{u}\cdot(\dot{\boldsymbol{u}}^{\perp}+\nabla\dot{h}) + \tfrac{1}{2}\lambda\nabla h\cdot\dot{\boldsymbol{u}} + \big(\tfrac{3}{2}\lambda - \tfrac{3}{4}\big)\boldsymbol{u}\cdot\nabla\dot{h}\big]\,\mathrm{d}\boldsymbol{x} \\ &= \int \big[-\tfrac{3}{4}h\dot{\boldsymbol{u}}\cdot(\boldsymbol{u}^{\perp}+\nabla h) - \big(\tfrac{3}{4}+\tfrac{1}{2}\lambda\big)\dot{h}\boldsymbol{u}\cdot\nabla h - \big(\tfrac{3}{4}-\lambda\big)h\boldsymbol{u}\cdot\nabla\dot{h}\big]\,\mathrm{d}\boldsymbol{x} \\ &= \int \big[-\tfrac{3}{4}h\dot{\boldsymbol{u}}\cdot(\boldsymbol{u}^{\perp}+\nabla h) + \big(\tfrac{3}{4}+\tfrac{1}{2}\lambda\big)(\boldsymbol{u}\cdot\nabla h)^2 + \big(\tfrac{3}{4}+\tfrac{1}{2}\lambda\big)h\boldsymbol{u}\cdot\nabla h\nabla\cdot\boldsymbol{u} \\ &\qquad + \big(\tfrac{3}{4}-\lambda\big)h\boldsymbol{u}\cdot\nabla(\boldsymbol{u}\cdot\nabla h) + \big(\tfrac{3}{4}-\lambda\big)h\boldsymbol{u}\cdot\nabla(h\nabla\cdot\boldsymbol{u})\big]\,\mathrm{d}\boldsymbol{x} \\ &= \int h\big[-\tfrac{3}{4}\dot{\boldsymbol{u}}\cdot(\boldsymbol{u}^{\perp}+\nabla h) + \boldsymbol{u}\cdot\big(-\tfrac{3}{2}\lambda(\nabla\boldsymbol{u})^T\nabla h - \tfrac{3}{2}\lambda\nabla\nabla h\boldsymbol{u} \\ &\qquad + \big(\tfrac{3}{4}-\lambda\big)\nabla h\nabla\cdot\boldsymbol{u} + \big(\tfrac{3}{4}-\lambda\big)h\nabla\nabla\cdot\boldsymbol{u}\big)\big]\mathrm{d}\boldsymbol{x}, \end{aligned} \tag{C 3}$$

where, in the second to last step, we have used the continuity equation to eliminate time derivatives of h. In the final step, we have used integration by parts on the integral of $(\boldsymbol{u}\cdot\nabla h)^2$. The above computation already outlines our general strategy: Our goal is eventually to factor out $h(\boldsymbol{u}^{\perp}+\nabla h)$ from all expressions – this is completed now for the $\dot{\boldsymbol{u}}$-term. For the remaining terms, we must first factor out $h\boldsymbol{u}$, and then iteratively complete to the full $h(\boldsymbol{u}^{\perp}+\nabla h)$, starting from the terms that are cubic in $\boldsymbol{u}$.

To start this procedure for L_{222}, we substitute in the expression for $\boldsymbol{v}$ and expand:

$$\begin{aligned} L_{222} = \int h\big[&\big(-\tfrac{1}{4}\boldsymbol{u}\cdot\nabla\boldsymbol{u}\boldsymbol{u}^{\perp} + \tfrac{3}{4}\boldsymbol{u}\cdot\nabla\boldsymbol{u}^{\perp}\boldsymbol{u}\big) \\ &+ \big(\tfrac{1}{2}\lambda\boldsymbol{u}\cdot\nabla\nabla^{\perp}h\boldsymbol{u}^{\perp} - \tfrac{1}{2}\lambda\boldsymbol{u}\cdot\nabla\boldsymbol{u}\nabla h + \tfrac{3}{2}\lambda\boldsymbol{u}\cdot\nabla\nabla h\boldsymbol{u} + \tfrac{1}{2}\lambda\nabla h\cdot\nabla\boldsymbol{u}\boldsymbol{u} \\ &+ \tfrac{1}{8}h\boldsymbol{u}^{\perp}\cdot\nabla\nabla\cdot\boldsymbol{u}^{\perp} - \tfrac{1}{8}h(\nabla\cdot\boldsymbol{u}^{\perp})^2\big) \\ &+ \big(\lambda^2\Delta h\boldsymbol{u}\cdot\nabla^{\perp}h + \tfrac{1}{4}\lambda h\boldsymbol{u}^{\perp}\cdot\nabla\Delta h + \tfrac{1}{4}\lambda h\nabla h\cdot\nabla\nabla\cdot\boldsymbol{u}^{\perp} - \tfrac{1}{2}\lambda h\Delta h\nabla\cdot\boldsymbol{u}^{\perp}\big) \\ &+ \big(\tfrac{1}{2}\lambda^2 h\nabla h\cdot\nabla\Delta h - \tfrac{1}{2}\lambda^2 h(\Delta h)^2\big)\big]\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{C 4}$$

The simplification in the third group of terms is based on identity (A 10). We must further integrate by parts on the term

$$\int h^2(\nabla\cdot\boldsymbol{u}^\perp)^2\,\mathrm{d}\boldsymbol{x} = -\int h[h\boldsymbol{u}^\perp\cdot\nabla\nabla\cdot\boldsymbol{u}^\perp + 2\boldsymbol{u}^\perp\cdot\nabla h\nabla\cdot\boldsymbol{u}^\perp]\,\mathrm{d}\boldsymbol{x}. \tag{C 5}$$

We then set $L_{22} = L_{221} + L_{222}$ and combine terms:

$$\begin{aligned} L_{22} = \int h\big[&(\boldsymbol{u}^\perp + \nabla h)\cdot\big(-\tfrac{3}{4}\dot{\boldsymbol{u}}\big) + \boldsymbol{u}\cdot\big(\tfrac{3}{4}\nabla\boldsymbol{u}^\perp\boldsymbol{u} - \tfrac{1}{4}\nabla\boldsymbol{u}\boldsymbol{u}^\perp\big)\\ &+\boldsymbol{u}\cdot\big(\tfrac{1}{2}\lambda\nabla\nabla^\perp h\boldsymbol{u}^\perp - \tfrac{1}{2}\lambda\nabla\boldsymbol{u}\nabla h - \lambda(\nabla\boldsymbol{u})^T\nabla h + \big(\tfrac{3}{4}-\lambda\big)\nabla h\nabla\cdot\boldsymbol{u}\\ &+\big(\tfrac{3}{4}-\lambda\big)h\nabla\nabla\cdot\boldsymbol{u} - \tfrac{1}{4}h\nabla^\perp\nabla\cdot\boldsymbol{u}^\perp - \tfrac{1}{4}\nabla^\perp h\nabla\cdot\boldsymbol{u}^\perp\big)\\ &+\big(\lambda^2\Delta h\boldsymbol{u}\cdot\nabla^\perp h + \tfrac{1}{4}\lambda h\boldsymbol{u}^\perp\cdot\nabla\Delta h + \tfrac{1}{4}\lambda h\nabla h\cdot\nabla\nabla\cdot\boldsymbol{u}^\perp - \tfrac{1}{2}\lambda h\Delta h\nabla\cdot\boldsymbol{u}^\perp\big)\\ &+\big(\tfrac{1}{2}\lambda^2 h\nabla h\cdot\nabla\Delta h - \tfrac{1}{2}\lambda^2 h(\Delta h)^2\big)\big]\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{C 6}$$

Next in line are the terms that are cubic in $\boldsymbol{u}$. We write

$$\begin{aligned} L_{22} = \int h\big[&(\boldsymbol{u}^\perp + \nabla h)\cdot\big(-\tfrac{3}{4}\dot{\boldsymbol{u}} - \tfrac{3}{4}\nabla\boldsymbol{u}\boldsymbol{u} - \tfrac{1}{4}\nabla\boldsymbol{u}^\perp\boldsymbol{u}^\perp\big)\\ &+\boldsymbol{u}^\perp\cdot\big(\tfrac{1}{2}\lambda\nabla^\perp\nabla h\boldsymbol{u} - \tfrac{1}{2}\lambda\nabla\boldsymbol{u}^\perp\nabla h + \big(\tfrac{3}{4}-\lambda\big)\big(\nabla^\perp\boldsymbol{u})^T\nabla h + \tfrac{1}{4}(\nabla\boldsymbol{u}^\perp)^T\nabla h\\ &+\big(\tfrac{3}{4}-\lambda\big)\nabla^\perp h\nabla\cdot\boldsymbol{u} + \big(\tfrac{3}{4}-\lambda\big)h\nabla^\perp\nabla\cdot\boldsymbol{u} + \tfrac{1}{4}h\nabla\nabla\cdot\boldsymbol{u}^\perp + \tfrac{1}{4}\nabla h\nabla\cdot\boldsymbol{u}^\perp\big)\\ &+\big(\lambda^2\Delta h\boldsymbol{u}\cdot\nabla^\perp h + \tfrac{1}{4}\lambda h\boldsymbol{u}^\perp\cdot\nabla\Delta h + \tfrac{1}{4}\lambda h\nabla h\cdot\nabla\nabla\cdot\boldsymbol{u}^\perp - \tfrac{1}{2}\lambda h\Delta h\nabla\cdot\boldsymbol{u}^\perp\big)\\ &+\big(\tfrac{1}{2}\lambda^2 h\nabla h\cdot\nabla\Delta h - \tfrac{1}{2}\lambda^2 h(\Delta h)^2\big)\big]\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{C 7}$$

We repeat our strategy for the quadratic-in-$\boldsymbol{u}$ terms, i.e.

$$\begin{aligned} L_{22} = \int h(\boldsymbol{u}^\perp + \nabla h)\cdot\big[&-\tfrac{3}{4}\dot{\boldsymbol{u}} - \tfrac{3}{4}\nabla\boldsymbol{u}\boldsymbol{u} - \tfrac{1}{4}\nabla\boldsymbol{u}^\perp\boldsymbol{u}^\perp\\ &+\tfrac{1}{2}\lambda\nabla^\perp\nabla h\boldsymbol{u} - \tfrac{1}{2}\lambda\nabla\boldsymbol{u}^\perp\nabla h + \big(\tfrac{3}{4}-\lambda\big)(\nabla^\perp\boldsymbol{u})^T\nabla h + \tfrac{1}{4}(\nabla\boldsymbol{u}^\perp)^T\nabla h\\ &+\big(\tfrac{3}{4}-\lambda\big)\nabla^\perp h\nabla\cdot\boldsymbol{u} + \big(\tfrac{3}{4}-\lambda\big)h\nabla^\perp\nabla\cdot\boldsymbol{u} + \tfrac{1}{4}h\nabla\nabla\cdot\boldsymbol{u}^\perp + \tfrac{1}{4}\nabla h\nabla\cdot\boldsymbol{u}^\perp\big]\,\mathrm{d}\boldsymbol{x} + L_{22}^{\mathrm{deg}}, \end{aligned} \tag{C 8}$$

where the two terms involving $\nabla^\perp h\nabla\cdot\boldsymbol{u}$ and $h\nabla^\perp\nabla\cdot\boldsymbol{u}$ do not contribute to L_{22}^{deg}, and the others expand to

$$\begin{aligned} L_{22}^{\mathrm{deg}} = \int\big[&-\tfrac{1}{2}\lambda h\nabla h\cdot\nabla^\perp\nabla h\boldsymbol{u} + \big(\tfrac{1}{2}\lambda - \tfrac{1}{4}\big)h\nabla h\cdot\nabla\boldsymbol{u}^\perp\nabla h\\ &-\big(\tfrac{3}{4}-\lambda\big)h\nabla h\cdot\nabla^\perp\boldsymbol{u}\nabla h + \big(\tfrac{1}{4}\lambda - \tfrac{1}{4}\big)h^2\nabla h\cdot\nabla\nabla\cdot\boldsymbol{u}^\perp - \tfrac{1}{4}h|\nabla h|^2\nabla\cdot\boldsymbol{u}^\perp\\ &+\lambda^2 h\Delta h\boldsymbol{u}\cdot\nabla^\perp h + \tfrac{1}{4}\lambda h^2\boldsymbol{u}^\perp\cdot\nabla\Delta h - \tfrac{1}{2}\lambda h^2\Delta h\nabla\cdot\boldsymbol{u}^\perp\\ &+\tfrac{1}{2}\lambda^2 h^2\nabla h\cdot\nabla\Delta h - \tfrac{1}{2}\lambda^2 h^2(\Delta h)^2\big]\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{C 9}$$

To bring these terms into standard form, we use the following identities:

$$\int h\nabla h\cdot\nabla\boldsymbol{u}^{\perp}\nabla h\,\mathrm{d}\boldsymbol{x} = -\int h\boldsymbol{u}^{\perp}\cdot(\nabla h\Delta h + \nabla\nabla h\nabla h + h^{-1}\nabla h|\nabla h|^2)\,\mathrm{d}\boldsymbol{x}, \tag{C 10}$$

$$\begin{aligned}\int h\nabla h\cdot\nabla^{\perp}\boldsymbol{u}\nabla h\,\mathrm{d}\boldsymbol{x} &= -\int h\boldsymbol{u}^{\perp}\cdot\nabla^{\perp}\nabla^{\perp}h\nabla h\,\mathrm{d}\boldsymbol{x}\\ &= \int h\boldsymbol{u}^{\perp}\cdot(\nabla\nabla h\nabla h - \nabla h\Delta h)\,\mathrm{d}\boldsymbol{x},\end{aligned} \tag{C 11}$$

$$\int h|\nabla h|^2\nabla\cdot\boldsymbol{u}^{\perp}\,\mathrm{d}\boldsymbol{x} = -\int h\boldsymbol{u}^{\perp}\cdot(h^{-1}\nabla h|\nabla h|^2 + 2\nabla\nabla h\nabla h)\,\mathrm{d}\boldsymbol{x}, \tag{C 12}$$

$$\int h^2\Delta h\nabla\cdot\boldsymbol{u}^{\perp}\,\mathrm{d}\boldsymbol{x} = -\int h\boldsymbol{u}^{\perp}\cdot(h\nabla\Delta h + 2\nabla h\Delta h)\,\mathrm{d}\boldsymbol{x}, \tag{C 13}$$

$$\begin{aligned}\int h^2\nabla h\cdot\nabla\nabla\cdot\boldsymbol{u}^{\perp}\,\mathrm{d}\boldsymbol{x} &= -\int(h^2\Delta h\nabla\cdot\boldsymbol{u}^{\perp} + 2h|\nabla h|^2\nabla\cdot\boldsymbol{u}^{\perp})\,\mathrm{d}\boldsymbol{x}\\ &= \int h\boldsymbol{u}^{\perp}\cdot(h\nabla\Delta h + 2\nabla h\Delta h + 2h^{-1}\nabla h|\nabla h|^2 + 4\nabla\nabla h\nabla h)\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{C 14}$$

The second step in (C 11) is based on identity (A 9). Collecting terms, we find

$$\begin{aligned}L_{22}^{\mathrm{deg}} &= \int h\boldsymbol{u}^{\perp}\cdot\left[\left(\tfrac{1}{2}-\lambda^2\right)\nabla h\Delta h + \left(\lambda-\tfrac{1}{4}\right)h\nabla\Delta h + (2\lambda-1)\nabla\nabla h\nabla h\right]\mathrm{d}\boldsymbol{x}\\ &\quad + \int\left[\tfrac{1}{2}\lambda^2h^2\nabla h\cdot\nabla\Delta h - \tfrac{1}{2}\lambda^2h^2(\Delta h)^2\right]\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{C 15}$$

Since our goal is to eliminate all terms that are quadratic or cubic in $\boldsymbol{u}$, we must choose $\boldsymbol{v}'$ to be equal to the terms in the square bracket in (C 8) plus arbitrary terms that only depend on h, i.e.

$$\begin{aligned}\boldsymbol{v}' &= \boldsymbol{v}'_{\mathrm{free}} - \tfrac{3}{4}\dot{\boldsymbol{u}} - \tfrac{3}{4}\nabla\boldsymbol{u}\boldsymbol{u} - \tfrac{1}{4}\nabla\boldsymbol{u}^{\perp}\boldsymbol{u}^{\perp}\\ &\quad + \tfrac{1}{2}\lambda\nabla^{\perp}\nabla h\boldsymbol{u} - \tfrac{1}{2}\lambda\nabla\boldsymbol{u}^{\perp}\nabla h + \left(\tfrac{3}{4}-\lambda\right)(\nabla^{\perp}\boldsymbol{u})^T\nabla h + \tfrac{1}{4}(\nabla\boldsymbol{u}^{\perp})^T\nabla h\\ &\quad + \left(\tfrac{3}{4}-\lambda\right)\nabla^{\perp}h\nabla\cdot\boldsymbol{u} + \tfrac{1}{4}\nabla h\nabla\cdot\boldsymbol{u}^{\perp} + \left(\tfrac{3}{4}-\lambda\right)h\nabla^{\perp}\nabla\cdot\boldsymbol{u} + \tfrac{1}{4}h\nabla\nabla\cdot\boldsymbol{u}^{\perp},\end{aligned} \tag{C 16}$$

where we choose

$$\boldsymbol{v}'_{\mathrm{free}} = \alpha\nabla h\Delta h + \beta h\nabla\Delta h + \gamma\nabla\nabla h\nabla h + \mu h^{-1}\nabla h|\nabla h|^2. \tag{C 17}$$

As in the first-order computation, we only introduce terms that have the same homogeneity as those already present.

If we substitute in this expression for $\boldsymbol{v}'$ directly, we see that there are five different terms that are quartic in h. However, integration by parts shows that there are actually only three independent terms at this level:

$$\int h^2\nabla h\cdot\nabla\Delta h\,\mathrm{d}\boldsymbol{x} = -\int(h^2\,(\Delta h)^2 + 2h|\nabla h|^2\Delta h)\,\mathrm{d}\boldsymbol{x}, \tag{C 18}$$

$$\int h\nabla h\cdot\nabla\nabla h\nabla h = -\frac{1}{2}\int(h|\nabla h|^2\Delta h + |\nabla h|^4)\,\mathrm{d}\boldsymbol{x}. \tag{C 19}$$

We can eliminate the remaining mixed term via

$$\int h^2\nabla h\cdot\nabla\Delta h\,\mathrm{d}\boldsymbol{x} = -\int(h^2\nabla\nabla h:\nabla\nabla h + 2h\nabla h\cdot\nabla\nabla h\nabla h)\,\mathrm{d}\boldsymbol{x}, \tag{C 20}$$

so that, using identities (C 18) and (C 19), we find

$$\int h|\nabla h|^2 \Delta h \, \mathrm{d}\boldsymbol{x} = \tfrac{1}{3} \int (h^2 \nabla\nabla h : \nabla\nabla h - h^2(\Delta h)^2 - |\nabla h|^4) \, \mathrm{d}\boldsymbol{x}. \tag{C 21}$$

Equations (C 18) and (C 19) then read

$$\int h^2 \nabla h \cdot \nabla \Delta h \, \mathrm{d}\boldsymbol{x} = \int \left(\tfrac{2}{3}|\nabla h|^4 - \tfrac{1}{3}h^2(\Delta h)^2 - \tfrac{2}{3}h^2 \nabla\nabla h : \nabla\nabla h \right) \mathrm{d}\boldsymbol{x}, \tag{C 22}$$

$$\int h \nabla h \cdot \nabla\nabla h \nabla h = \int \left(\tfrac{1}{6}h^2(\Delta h)^2 - \tfrac{1}{3}|\nabla h|^4 - \tfrac{1}{6}h^2 \nabla\nabla h : \nabla\nabla h \right) \mathrm{d}\boldsymbol{x}. \tag{C 23}$$

Substituting all intermediate expressions back into (5.9), we obtain the final form (5.29) of transformed L_2 Lagrangian.

REFERENCES

Allen, J. S. & Holm, D. D. 1996 Extended-geostrophic Hamiltonian models for rotating shallow water motion. *Physica* D **98**, 229–248.

Arnold, V. I. & Khesin, B. A. 1998 *Topological Methods in Hydrodynamics*. Springer.

Benamou, J.-D. & Brenier, Y. 1998 Weak existence for the semi-geostrophic equations formulated as a coupled Monge–Ampère/transport problem. *SIAM J. Appl. Maths* **58**, 1450–1461.

Bhat, H. C., Fetecau, R. C., Marsden, J. E., Mohseni, K. & West, M. 2005 Lagrangian averaging for compressible fluids. *SIAM J. Multiscale Model. Simulation* **3**, 818–837.

Bokhove, O., Vanneste, J. & Warn, T. 1998 A variational formulation for barotropic quasi-geostrophic flows. *Geophys. Astrophys. Fluid Dyn.* **88**, 67–79.

Bretherton, F. P. 1970 A note on Hamilton's principle for perfect fluids. *J. Fluid Mech.* **44**, 19–31.

Bridges, T., Hydon, P. & Reich, S. 2005 Vorticity and symplecticity in Lagrangian fluid dynamics. *J. Phys.* A: *Math. Gen.* **38**, 1403–1418.

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

Cullen, M. J. P. & Purser, R. J. 1989 Properties of the Lagrangian semi-geostrophic equations. *J. Atmos. Sci.* **46**, 2684–2697.

Dirac, P. A. M. 1966 *Lectures on Quantum Mechanics*. Yeshiva University, New York.

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

Eliassen, A. 1962 On the vertical circulation in frontal zones. *Geofys. Publ.* **24**, 147–160.

Evans, L. C. 1998 *Partial Differential Equations*. American Mathematical Society, Providence, RI.

Ford, R., Malham, S. J. A. & Oliver, M. 2002 A new model for shallow water in the low Rossby-number limit. *J. Fluid Mech.* **450**, 287–296.

Gottwald, G. & Oliver, M. 2005 Averaged dynamics via degenerate variational asymptotics. In preparation.

Gottwald, G., Oliver, M. & Tecu, N. 2005 Long-time accuracy for approximate slow manifolds in a finite dimensional model of balance. In preparation.

Holm, D. D. 1999 Fluctuation effects on 3D Lagrangian mean and Eulerian mean fluid motion. *Physica* D **133**, 215–269.

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

Holm, D. D. & Zeitlin, V. 1998 Hamilton's principle for quasi-geostrophic motion. *Phys. Fluids* **10**, 800–806.

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

Lorenz, E. N. 1980 Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.* **37**, 1685–1699.

Lychagin, V. V., Rubtsov, V. N. & Chekalov, I. V. 1993 A classification of Monge–Ampère equations. *Ann. Sci. École Norm. Sup.* (4) **26**, 281–308.

Lynch, P. 2002 The swinging spring: a simple model of atmospheric balance. *Large-scale Atmosphere–Ocean Dynamics: II: Geometric Methods and Models* (ed. I. Roulstone & J. Norbury). Cambridge University Press.

McIntyre, M. E. & Roulstone, I. 1996 Hamiltonian balanced models: constraints, slow manifolds and velocity splitting. *Forecasting Res. Sci. Paper* 41, Met Office, UK.

McIntyre, M. E. & Roulstone, I. 2002 Are there higher-accuracy analogues of semi-geostrophic theory? *Large-scale Atmosphere–Ocean Dynamics: II: Geometric Methods and Models* (ed. I. Roulstone & J. Norbury). Cambridge University Press.

Marsden, J. E. & Shkoller, S. 2001 Global well-posedness for the Lagrangian averaged Navier–Stokes (LANS-α) equations on bounded domains. *Phil. Trans. R. Soc. Lond.* A **359**, 1449–1468.

Marsden, J. E. & Shkoller, S. 2003 The anisotropic Lagrangian averaged Euler and Navier–Stokes equations. *Arch. Ration. Mech. Anal.* **166**, 27–46.

Oliver, M. & Shkoller, S. 2001 The vortex blob method as a second-grade non-Newtonian fluid. *Commun. Partial Differential Equations* **26**, 295–314.

Ripa, P. 1981 Symmetries and conservation laws for internal gravity waves. *Nonlinear Properties of Internal Waves* (ed. B.J. West). *AIP Conf. Proc.* **76**, 281–306.

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

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

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

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

Shepherd, J. R. & Ford, R. 2001 Dynamics of an oceanic model with a single thermally-active layer. Unpublished manuscript.

Vanneste, J. 2004 Inertia-gravity-wave generation by balanced motion: revisiting the Lorenz–Krishnamurthy model. *J. Atmos. Sci.* **61**, 224–234.

Vanneste, J. & Bokhove, O. 2002 Dirac-bracket approach to nearly geostrophic Hamiltonian balanced models. *Physica* D **164**, 152–167.

Wunderer, C. F. 2001 Hamiltonian and non-Hamiltonian theories of balanced vortical flow. PhD thesis, University of Cambridge.