Extended shallow-water theories with thermodynamics and geometry

F.J. Beron-Vera

Department of Atmospheric Sciences, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida 33149, USA

(Dated: 23 September 2021)

Driven by growing momentum in two-dimensional geophysical flow modeling, this paper introduces a general family of “thermal” rotating shallow-water models. The models are capable of accommodating thermodynamic processes, such as those acting in the ocean mixed layer, by allowing buoyancy to vary in horizontal position and time as well as with depth, in a polynomial fashion up to an arbitrary degree. Moreover, the models admit Euler–Poincare variational formulation and possess Lie–Poisson Hamiltonian structure. Such a geometric property provides solid fundamental support to the theories described with consequences for numerical implementation and the construction of unresolved motion parametrizations. In particular, it is found that stratification halts the development of small-scale filament rollups recently observed in a popular model, which, having vertically homogeneous density, represents a special case of the models presented here.

PACS numbers: 02.50.Ga; 47.27.De; 92.10.Fj

I. INTRODUCTION

This paper is motivated by renewed interest in “thermal shallow-water modelling” of geophysical flows. Due to Pedro Ripa’s investigation of its geometric structure and stability properties, the core model has been known to many as Ripa’s model. Its origins can be traced back to the mid 1960s and early 1970s, when various authors independently introduced the idea of incorporating thermodynamics to the rotating shallow-water (or, more fairly, Laplace’s tidal) equations, i.e., the two-dimensional Euler equations for homogeneous fluid with Coriolis force in the horizontal and the hydrostatic approximation, by vertically averaging the pressure gradient force of the full primitive equations (or PE), i.e., the three-dimensional Euler equations for arbitrary stratified incompressible fluid with Coriolis force in the horizontal and the hydrostatic and Boussinesq approximations. The result is a model with four prognostic variables: the two components of the velocity, layer thickness, and buoyancy, all varying in the horizontal and time. The main advantage of the resulting model over the rotating shallow-water model is its ability to incorporate heat and freshwater fluxes across the air–sea interface. This feature was extensively used to simulate mixed-layer (i.e., the topmost part of the ocean, above the thermocline) dynamics, particularly through the 1980s and 1990s (e.g., Beier, McCready, Kundu, and Molinari, respectively) realistically at a relatively low computational cost. Moreover, recent high-resolution numerical simulations of a quasigeostrophic (QG) approximation of the system (i.e., with the pressure gradient force nearly balancing the Coriolis force) revealed the formation of Kelvin–Helmholtz-like rollup filaments (Fig. 1: for additional snapshots, cf. Fig. 3, left panel, multimedia view) that resemble quite well submesoscale (1–10 km) features identified in satellite-derived chlorophyll distributions on the surface of the ocean. Such submesoscale circulations are believed to play an important role in the exchange of gases through the ocean surface and thus global climate, and are the subject of intense research.

In this paper I present a family of rotating shallow-water models with thermodynamics that extend Ripa’s model by allowing buoyancy to vary in the vertical.
The dependence on the vertical coordinate is polynomial, up to an arbitrary degree. This extension enables representation of important mixed-layer dynamics processes such as restratification by baroclinic instability.\textsuperscript{15} I will spend a substantial part of the paper discussing its geometric properties (Euler–Poincare variational formulation\textsuperscript{21} and generalized (noncanonical) Hamiltonian structure\textsuperscript{40}). This will enable making an explicit connection, so far overlooked, with earlier work by Morrison and Greene\textsuperscript{40}. Ripa\textsuperscript{56} presents a rotating shallow-water model with thermodynamics with velocity shear in the vertical. Such a model, however, does not possess the geometric structure of the model presented here. I emphasize this property of the present model, which is shared with the PE from which it derives. Such a structure can be exploited to formulate integration schemes to avoid spurious numerical effects. Structure-preserving algorithm development is a topic of active investigation nowadays.\textsuperscript{12,13,23,39} Additional motivation for preserving structure in rotating shallow-water theories with thermodynamics is that this enables the application of a flow-topology-preserving framework\textsuperscript{18} for building parametrizations\textsuperscript{8} of unresolvable motions and this way investigating their contribution to transport at resolvable scales. Furthermore, the Hamiltonian formulation of fluid dynamics provides a systematic manner to find conservation laws, linked with symmetries via Noether’s theorem. These can be used to derived a-priori flow stability statements\textsuperscript{16} as well as to gain insight into the nature of growing perturbations to unstable states\textsuperscript{52} or infer saturation bounds on their nonlinear growth.\textsuperscript{65}

Following notation introduced in Ripa\textsuperscript{56}, I will use the acronym IL, for inhomogeneous layer, to refer to the rotating shallow-water models with thermodynamics discussed in this paper. Superscript(s) will be appended to IL to indicate the amount of vertical variation permitted in the horizontal velocity and buoyancy fields. With this notation, Ripa’s model will be referred to as the IL\textsuperscript{0}PE model, to mean that neither velocity nor buoyancy change with depth. In turn, the full PE will be referred to as the IL\textsuperscript{∞}PE model, since velocity and buoyancy vary with depth unrestrictedly.

The rest of the paper is organized as follows. The general IL model with thermodynamics in PE is introduced in Sec. II. Its geometric structure is discussed in extent in Sec. III. The interpretation of the model as a rotating shallow-water with thermodynamics is given in Sec. IV, where several submodels are considered along with their QG form, insightful in the study of processes such as baroclinic instability leading to the rollup filaments in Fig. 1, and corresponding geometric structure. The effects of stratification on the development of such rollup filaments are discussed in Section V. Section VI summarizes the paper in addition to offering a discussion and an outlook on avenues for future research. Finally, three appendices are included with the proof of a result that extends earlier results by other authors\textsuperscript{40} and some abstract mathematics details that provide deeper geometrical mechanics interpretations of the equations discussed.

II. THE IL\textsuperscript{(0,α)-PE} MODEL

Let $\mathbf{x} = (x, y) \in D \subseteq \mathbb{R}^2$ denote position in some fixed domain of a $\beta$-plane. Let $\tilde{\mathbf{u}}(\mathbf{x}, t)$ denote instantaneous velocity at time $t$ inside a fluid layer limited from above by a rigid lid. The bottom of the layer is soft as the layer floats atop an infinitely deep layer of quiescent fluid, i.e., a reduced gravity is assumed. However, the model equations hold for a system with a rigid bottom and a free surface. The overbar notation is used to emphasize that independence on the vertical coordinate ($z$) is achieved via (vertical) average.

If $D \subset \mathbb{R}^2$ has a solid boundary, then no normal flow through it is implied, i.e., $\tilde{\mathbf{u}} \cdot \mathbf{n}|_{\partial D} = 0$; conversely, if $D = \mathbb{R}^2$ then we will assume that $\tilde{\mathbf{u}}$ decays to zero as $|\mathbf{x}| \to \infty$. A physically relevant setting is one with $D$ representing a zonal channel, infinitely long ($D = \mathbb{R} \times [0, L_y]$) or reentering ($D = \mathbb{R}/L_z \mathbb{Z} \times [0, L_y]$). The boundary condition in the first case is a combination of the aforementioned ones; in the second case, one requires periodicity in $x$ i.e., $\tilde{\mathbf{u}}(x, y, t) = \tilde{\mathbf{u}}(x + L_x, y, t)$. Doubly reentering domains, i.e., $D = \mathbb{R}/L_z \mathbb{Z} \times \mathbb{R}/L_y \mathbb{Z}$, and domains including islands, i.e., with $\partial D$ being the union of multiply connected components, are also plausible, with boundary conditions being periodicity of $\tilde{\mathbf{u}}$ in $x$ and $y$ and no flow through each connected component of the boundary, respectively.\textsuperscript{28}

Let

$$\rho_{n_\alpha}(\mathbf{x}, t), \quad n_\alpha = 1, 2, \ldots, \alpha + 2,$$

be a finite set functions, each one satisfying

$$\int_{D_\alpha} d^2\mathbf{x} \rho_{n_\alpha} = \text{const},$$

where $D_\alpha \subset D$ is a fluid region (i.e., carried along with the flow of $\tilde{\mathbf{u}}$), and consider the function

$$\varphi_{\alpha}(\rho_1, \tilde{\rho}_2, \ldots, \tilde{\rho}_{\alpha+2}),$$

where

$$\tilde{\rho}_{n_\alpha} := \rho_1^{-1} \rho_{n_\alpha}, \quad \tilde{n}_\alpha = 2, \ldots, \alpha + 2,$$

which is assumed to be differentiable in all of its arguments. When $D \subset \mathbb{R}^2$ has a solid boundary, an additional boundary condition is

$$\nabla \tilde{\rho}_{n_\alpha} \times \tilde{\mathbf{u}}|_{\partial D} = 0$$

(the above applies on each component of the boundary of a multiply connected $D$). The nature of (5) will be clarified below.

Let

$$\mathbf{m} := \rho_1(\tilde{\mathbf{u}} + \mathbf{f}), \quad \tilde{z} \cdot \nabla \times \mathbf{f} = f_0 + \beta y,$$
where $\mathbf{f}(\mathbf{x})$ is the vector potential of the local angular velocity of the Earth relative to a fixed frame, $\frac{1}{2} f \mathbf{\hat{z}}$. Finally, define

$$\bar{q} := \frac{\mathbf{\hat{z}} \cdot \nabla \times \mathbf{\hat{u}} + f}{\rho_1} \equiv \rho_1^{-1} \mathbf{\hat{z}} \cdot \nabla \times \rho_1^{-1} \mathbf{\hat{m}}. \quad (7)$$

If $\rho_1$ is taken as the thickness of the fluid layer, as we will throughout this paper, then (7) is the vertical average of Ertel’s potential vorticity in the IL$^{(0,\alpha)}$ + PE model.\textsuperscript{39} We will refer $\bar{q}$ simply as potential vorticity.

The IL$^{(0,\alpha)}$ + PE is given by:

$$\partial_t \mathbf{\hat{m}} + \mathcal{L}_\mathbf{\hat{u}} \mathbf{\hat{m}} - \rho_1 \nabla K + \nabla \rho_1^2 \tilde{\partial}_{\rho_1} \varphi_{\alpha+} = 0,$$

$$\partial_t \rho_{n\alpha} + \nabla \cdot \rho_{n\alpha} \mathbf{\hat{u}} = 0, \quad (8)$$

where

$$K := \frac{1}{2} |\mathbf{\hat{u}}|^2 + \mathbf{\hat{u}} \cdot \mathbf{f} \quad (9)$$

and

$$\tilde{\partial}_{\rho_1} := \partial_{\rho_1} |_{\rho_{n\alpha}}, \quad (10)$$

and we have adopted the convenient notation\textsuperscript{46}

$$\mathcal{L}_\mathbf{a} \mathbf{b} := (\mathbf{a} \cdot \nabla) \mathbf{b} + b_j \nabla a^j + (\nabla \cdot \mathbf{a}) \mathbf{b} = b_j \nabla a^j + \partial_j \mathbf{b} a^j \quad (11)$$

for any vectors $\mathbf{a}, \mathbf{b}(\mathbf{x}, t)$ (summation over repeated (lowered and raised) indices is implied unless indicated in between parentheses). Note that $\mathcal{L}_\mathbf{a} \mathbf{\hat{m}}$ includes three terms. The first represents the advection of $\mathbf{\hat{m}}$ by $\mathbf{\hat{u}}$, and the second and third terms respectively account for the stretching and expansion of fluid elements as $\mathbf{\hat{m}}$ is being transported by $\mathbf{\hat{u}}$.

Dividing the first equation in (8) by $\rho_1$, and using the second equation for $n_\alpha = 1$, it takes more familiar form

$$\partial_t \mathbf{\hat{u}} + (\mathbf{\hat{u}} \cdot \nabla) \mathbf{\hat{u}} + f \mathbf{\hat{z}} \times \mathbf{\hat{u}} = -\rho_1^{-1} \nabla \rho_1^2 \tilde{\partial}_{\rho_1} \varphi_{\alpha+}. \quad (12)$$

This facilitates the interpretation of the term $-\rho_1^{-1} \nabla \rho_1^2 \tilde{\partial}_{\rho_1} \varphi_{\alpha+}$ as a generalized pressure gradient force. The second equation in (8) is a consequence of Reynolds’ transport theorem applied on (2), expressing continuity of the functions $\rho_{n\alpha}$. These functions, therefore, represent conserved densities, and thus are different than $\rho_{n\alpha}$, which represent quantities transported by the flow of $\mathbf{\hat{u}}$, i.e., by continuity of $\rho_1$,

$$\frac{\text{D} \rho_{n\alpha}}{\text{D} t} = 0, \quad (13)$$

where

$$\frac{\text{D}}{\text{D} t} := \partial_t + (\mathbf{\hat{u}} \cdot \nabla) \quad (14)$$

is the material derivative.

I will make explicit the interpretation of the IL$^{(0,\alpha)}$ + PE model (8) as a family of rotating shallow-water systems with thermodynamics after discussing its geometric structure first.

III. GEOMETRY

The IL$^{(0,\alpha)}$ + PE model both admits an Euler–Poincare variational formulation and possesses a generalized Hamiltonian structure, as it is shown here. By Euler–Poincare variational principle I mean a Hamilton’s principle for fluids that leads to motion equations in Eulerian variables. This is accomplished by writing the Lagrangian in terms of Eulerian variables and then extremizing the corresponding action under constrained variations representing fluid particle path variations at fixed Lagrangian labels and time. This principle, as just described, was apparently first discussed in Newcomb.\textsuperscript{44} Yet it belongs to a much general variational formulation of mechanics, written in the abstract language of differential geometry by Holm, Marsden, and Ratiu,\textsuperscript{21, 22} who made connections with seminal work by Henri Poincare, reviewed in Marle.\textsuperscript{27} By generalized Hamiltonian structure I mean a noncanonical Hamiltonian representation of the equations of motion in terms of Eulerian variables as discussed by Morrison and Greene.\textsuperscript{45} The abstract formulation, following earlier work by Arnold,\textsuperscript{1} is due to Marsden and Weinstein.\textsuperscript{28, 31} In particular, Marsden, Ratiu, and Weinstein showed how to obtain the Euler equation for compressible fluid motion as a generalized Hamiltonian system via reduction by symmetry of the corresponding canonical Hamiltonian formulation, i.e., the Euler equation written in Lagrangian variables.

Fixing notation, calligraphic letters, e.g., $\mathcal{W}$, will be used to denote real-valued functionals of fields, e.g., $\mu(x, t) = (\mu^1(x, t), \mu^2(x, t), \ldots)$. We will restrict attention to functionals of the specific form:

$$\mathcal{W}[\mu] = \int U(x; \mu, \nabla \mu, \ldots) \quad (15)$$

where $\int := \int_D d^2 x$ operates on everything on its right. Note that $U$ in (15) does not depend explicitly on time. Otherwise it has a finite number of arguments, and is a sufficiently smooth function in each of them. The latter makes $\mathcal{W}$ smooth enough for the functional derivative wrt $\mu$, i.e., the unique element $\frac{\delta \mathcal{W}}{\delta \mu}$ satisfying

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon = 0} \mathcal{W}[\mu + \varepsilon \delta \mu] = \int \frac{\delta \mathcal{W}}{\delta \mu} \cdot \delta \mu := \frac{\delta \mathcal{W}}{\delta \mu} \delta \mu^a, \quad (16)$$

to be well defined. In particular, $\mu^a(x', t) = \int \delta(x - x') \mu^a(x, t)$; then

$$\frac{\delta \mu^a(x', t)}{\delta \mu^b(x, t)} = \delta^a_\alpha \delta(x - x'). \quad (17)$$

The shorthand $\mathcal{W}_\mu$ for $\frac{\delta \mathcal{W}}{\delta \mu}$ will be used in inline equation display or offline when notationally more convenient.

A. Euler–Poincare variational formulation

Consider the functional

$$\mathcal{L}[\mathbf{\hat{u}}, \rho] := \int \rho_1 (K - \varphi_{\alpha+}) \quad (18)$$
where $\rho = (\rho_{na})$. Its functional derivatives are:
\[
\begin{align*}
\frac{\delta L}{\delta u} &= \hat{m}, \\
\frac{\delta L}{\delta \rho_1} &= K - \varphi_{a+} - \rho_1 \tilde{\varphi}_{a+} + \tilde{\rho}_n \delta \tilde{\varphi}_{a+}, \\
\frac{\delta L}{\delta \rho_n} &= -\partial_{\tilde{\rho}_n} \varphi_{a+}.
\end{align*}
\]

Noting that
\[
\nabla \varphi_{a+} = (\tilde{\varphi}_{a+}) \nabla \rho_1 + (\partial_{\tilde{\rho}_n} \varphi_{a+}) \nabla \tilde{\rho}_n,
\]
and
\[
(\tilde{\varphi}_{a+}) \nabla \rho_1 + \nabla (\rho_1 \tilde{\varphi}_{a+}) = \rho_1^{-1} \nabla \rho_1^2 \tilde{\varphi}_{a+},
\]
the first equation of (8) follows as
\[
\partial_t \frac{\delta L}{\delta u} + L_u \frac{\delta L}{\delta u} = \rho_n \nabla \frac{\delta L}{\delta \rho_n},
\]
upon some cancellation.

The above suggests that the IL(0,a)+PE model equations (8) admit an Euler–Poincare variational formulation. In fact, interpreting $L$ as a Lagrangian, with $K$ and $\varphi_{a+}$, thus corresponding to kinetic and potential energy density, respectively, (24) extremizes the action
\[
\mathcal{S} := \int_{t_0}^{t_1} dt \left[ \frac{\delta L}{\delta \dot{u}} \dot{u} - \frac{\delta L}{\delta \varphi_{a+}} \dot{\varphi}_{a+} + \frac{\delta L}{\delta \rho_n} \dot{\rho}_n \right],
\]
under constrained variations
\[
\frac{\delta \dot{u}}{\delta \varphi_{a+}} = \partial_t \eta + \dot{u} \eta + \nabla \cdot \rho_n \eta,
\]
\[
\frac{\delta \dot{\rho}_n}{\delta \rho_n} = -\nabla \cdot \varphi_{a+} \eta,
\]
where $\eta(x,t)$ is an arbitrary vector field satisfying $\eta(x,t_0) = 0 = \eta(x,1)$ and $\nabla \cdot \hat{u} = 0$. This is replaced by the solid boundary
\[
\frac{\delta \dot{u}}{\delta \varphi_{a+}} = \partial_t \eta + \dot{u} \eta + \nabla \cdot \rho_n \eta,
\]
\[
\frac{\delta \dot{\rho}_n}{\delta \rho_n} = -\nabla \cdot \varphi_{a+} \eta,
\]
which is the commutator of vector fields. Indeed,
\[
\begin{align*}
\delta \mathcal{S} &= \int_{t_0}^{t_1} dt \int \delta \frac{\delta L}{\delta \dot{u}} \delta \dot{u} + \delta \frac{\delta L}{\delta \varphi_{a+}} \delta \dot{\varphi}_{a+} + \delta \frac{\delta L}{\delta \rho_n} \delta \dot{\rho}_n \\
&= -\int_{t_0}^{t_1} dt \int \left( \partial_t \frac{\delta L}{\delta \dot{u}} + L_u \frac{\delta L}{\delta \dot{u}} - \rho_n \nabla \frac{\delta L}{\delta \dot{\rho}_n} \right) \cdot \eta.
\end{align*}
\]

\section{Kelvin circulation theorem}

Dividing (24) by $\rho_1$ and using continuity of $\rho_1$ (the second equation in (8) for $n_a = 1$), one finds
\[
\frac{D}{Dt} \frac{\delta L}{\delta \dot{u}} + \frac{1}{\rho_1} \frac{\delta L}{\delta \dot{u}} \nabla \tilde{u} - \nabla \frac{\delta L}{\delta \dot{u}} - \tilde{\rho}_n \nabla \frac{\delta L}{\delta \dot{\rho}_n} = 0.
\]
Introducing the circulation
\[
\gamma := \oint_{\partial D_u} \frac{1}{\rho_1} \frac{\delta L}{\delta \dot{u}} \cdot \hat{n} = \oint_{\partial D_u} (\dot{u} + \hat{f}) \cdot \hat{n}
\]
$\partial \mathcal{S}$ is transported by the flow of $\hat{u}$ as its closed material boundary $\partial D_u$, (31) it follows that
\[
\frac{d\gamma}{dt} = \int_{D_u} d^2 x \left\{ \tilde{\rho}_n, \frac{\delta L}{\delta \rho_n} \right\}_{xy},
\]
upon using $a \cdot D u = a \cdot d\hat{u} = a \cdot \partial_j \hat{u} dx_j = a_j \nabla \hat{u}_j \cdot dx$ along with Stokes theorem, where
\[
\{ A, B \}_{xy} := \partial_z A \times \partial B
\]
for all $A, B(x,t)$ is the canonical Poisson bracket, which is used to denote Jacobian of functions. Equation (33) is the statement of the Kelvin circulation theorem.

In general, $\gamma$ is not conserved; it is created (or destroyed) by the misalignment of the gradients of $\tilde{\rho}_n$ and $\partial_{\tilde{\rho}_n} \varphi_{a+}$. One exception is the case in which only one density, i.e., layer thickness or $\rho_1$, is included in (8). Another exception is the situation in which (8) is initialized from $\tilde{\rho}_n = \text{const}$, since they are transported by the flow of $\hat{u}$ (13) and thus remain constant at all times. Indeed, the set $\{ \tilde{\rho}_n(x,t) = \text{const} \}$ is an invariant manifold of (8). We will get back to this below when discussing submodels of the IL(0,a)+PE model.

On the other hand, if $\partial D_u$ is replaced by the solid boundary $\partial D$ of the flow domain, then $\gamma$, as in (32) or simply in this case $\gamma = \oint_{\partial D} \hat{u} \cdot dx$, is conserved because $\oint_{\partial D} (\partial_{\tilde{\rho}_n} \varphi_{a+}) \nabla \tilde{\rho}_n \cdot dx \equiv 0$ by the boundary condition (5), clarifying the nature of this boundary condition. This holds along each connected component of $\partial D$ when $D$ has multiple islands.

\section{Potential vorticity evolution}

Finally, using Stokes theorem we have
\[
\gamma = \int_{D_u} d^2 x \rho_1 \tilde{q};
\]
then by continuity of $\rho_1$, one finds
\[
\frac{D \tilde{q}}{Dt} = \rho_1^{-1} \left\{ \partial_{\tilde{\rho}_n} \varphi_{a+}, \tilde{\rho}_n \right\}_{xy}.
\]

In general, the potential vorticity is not transported by the flow of $\hat{u}$. Exceptions are the aforementioned ones, namely, the single-density case and initialization on the constant density invariant subspace.
B. Lie–Poisson Hamiltonian structure

Consider
\[
\mathcal{E}[\mathbf{m}, \rho] := \int \mathbf{m} \cdot \dot{\mathbf{u}} - \rho_1(K - \varphi_{\alpha+}),
\]
which is nothing but the total energy of system (8), viz.,
\[
\mathcal{E} \equiv \int \rho_1 \left( \frac{1}{2} |\dot{\mathbf{u}}|^2 + \varphi_{\alpha+} \right).
\]
The functional derivatives of (37) are:
\[
\frac{\delta \mathcal{E}}{\delta \mathbf{m}} = \dot{\mathbf{u}}, \quad \frac{\delta \mathcal{E}}{\delta \rho_{\alpha+}} = -\frac{\delta L}{\delta \rho_{\alpha+}}.
\]
Then taking (22)–(23) into account, set (8) follows from
\[
\begin{align*}
\partial_t \mathbf{m} + L_{\varphi_{\alpha+}} \mathbf{m} + \rho_{\alpha+} \nabla \cdot \frac{\delta \mathcal{E}}{\delta \rho_{\alpha+}} &= 0, \\
\partial_t \rho_{\alpha+} + \nabla \cdot \rho_{\alpha+} 
\end{align*}
\] (41)

after some cancellation.

The above suggests that the IL\(^{(0, \alpha)}\) + PE model (8) possesses a generalized Hamiltonian structure, i.e., it can be cast in the form
\[
\partial_t \mu = \{ \mu, \mathcal{H} \} = \mathcal{J} \frac{\delta \mathcal{H}}{\delta \mu},
\]
where
\[
\{ \mathcal{U}, \mathcal{V} \}[\mu] := \int \frac{\delta \mathcal{U}}{\delta \mu} \cdot \mathbf{J} \frac{\delta \mathcal{V}}{\delta \mu}
\] (43)

for any functionals \( \mathcal{U}, \mathcal{V} \). Here, \( \mu \) represents the state of the system as a “point” in an infinite-dimensional phase space; \( \mathcal{H} \) is the Hamiltonian; and \( \{, \} \) and \( \mathcal{J} \) are called the Poisson bracket and operator, respectively. The bracket is assumed to satisfy two properties that do not follow from its definition, namely, \( \{ \mathcal{U} + \mathcal{V}, \mathcal{W} \} = \{ \mathcal{U}, \mathcal{W} \} + \{ \mathcal{V}, \mathcal{W} \} \) (bilinearity) and \( \{ \mathcal{U} \mathcal{V}, \mathcal{W} \} = \mathcal{U} \{ \mathcal{V}, \mathcal{W} \} + \{ \mathcal{U}, \mathcal{W} \} \mathcal{V} \) (Leibniz rule) for all functionals of state \( \mathcal{U}, \mathcal{V}, \mathcal{W} \). The properties in question are: \( \{ \mathcal{U}, \mathcal{V} \} = -\{ \mathcal{V}, \mathcal{U} \} \) (antisymmetry) and \( \{ \{ \mathcal{U}, \mathcal{V} \}, \mathcal{W} \} + \{ \mathcal{U}, \{ \mathcal{V}, \mathcal{W} \} \} = 0 \) where \( \circ \) denotes the two other terms obtained by cyclic permutation of the functionals (Jacobi identity).

Setting \( \mu = (\mathbf{m}, \rho) \) and identifying \( \mathcal{H} \) with \( \mathcal{E} \) as given by (37), from (41) it readily follows that the Poisson operator for (8) is
\[
\mathcal{J} = -
\begin{pmatrix}
L_{\varphi_{\alpha+}} \quad \rho_\alpha \nabla \cdot \rho_1 \quad \ldots \quad \rho_{\alpha+} \nabla \cdot \rho_1 \\
\nabla \cdot \rho_1 \quad 0 \quad \ldots \quad 0 \\
\vdots \quad \vdots \quad \ddots \quad \vdots \\
\nabla \cdot \rho_{\alpha+} \quad 0 \quad \ldots \quad 0
\end{pmatrix}.
\]

This leads, upon integration by parts, to the Poisson bracket
\[
\{ \mathcal{U}, \mathcal{V} \}[\mathbf{m}, \rho] = \{ \mathcal{U}, \mathcal{V} \}^\mathbf{m} + \sum \{ \mathcal{U}, \mathcal{V} \}^\rho_{\alpha+} \quad (45a)
\]

where
\[
\{ \mathcal{U}, \mathcal{V} \}^\mathbf{m} := -\int \mathbf{m} \cdot \left[ \frac{\delta \mathcal{U}}{\delta \mathbf{m}} \frac{\delta \mathcal{V}}{\delta \mathbf{m}} \right], \quad (45b)
\]
\[
\{ \mathcal{U}, \mathcal{V} \}^\rho_{\alpha+} := -\int \rho_{\alpha+} \left( \frac{\delta \mathcal{U}}{\delta \mathbf{m}} \nabla \cdot \frac{\delta \mathcal{V}}{\delta \rho_{\alpha+}} - \frac{\delta \mathcal{U}}{\delta \mathbf{m}} \nabla \cdot \frac{\delta \mathcal{V}}{\delta \rho_{\alpha+}} \right). \quad (45c)
\]

Remark 1 To guarantee cancellation of the boundary term \( \int_{\partial D} \{ \mathcal{U}, \mathcal{V} \}^\mathbf{m} \mathbf{n} \times \mathbf{V} \m \cdot \mathbf{d}x \) when the flow domain has a solid boundary, we require, as is customary, \( \mathcal{U}_{\mathbf{m} \cdot \mathbf{n}} |_{\partial D} = 0 \) for any \( \mathcal{U} \). The space of admissible functionals of course is assumed to be closed: the bracket of two admissible functionals produces an admissible functional. Alternatively, boundary terms may be added to the definition of the bracket, yet not without imposing a restriction on the class of admissible functionals.

The bracket (45) is manifestly antisymmetric, and indeed satisfies the Jacobi identity. A proof is given in App. A. This leads to an extension of an earlier geometric mechanics result by Marsden, Ratiu, and Weinstein, outlined in App. C, where a much deeper geometric interpretation of the (41) is provided.

Remark 2 The Euler–Poincare variational formulation (24) and the equivalent Lie–Poisson Hamiltonian formulation (41) of the IL\(^{(0, \alpha)}\) + PE are equivalent. The connection between them is provided by the transformation \( (\mathbf{u}, \rho) \mapsto (\mathbf{m}, \rho) \) defined by (37), which represents a partial Legendre transformation analogous to that connecting the Lagrangian and (canonical) Hamiltonian formulations of mechanical systems (cf. further details in App. C).

1. Conservation laws

More generally than (42), one has \( \tilde{\mathcal{U}} = \{ \mathcal{U}, \mathcal{H} \} \) for any (admissible) functional of state \( \mathcal{U} \). By the antisymmetry of the bracket, \( \mathcal{H} \) commutes with itself in the bracket. Thus \( \mathcal{H}(= \mathcal{E}) \) is conserved under the dynamics, which can be verified directly from (8) (after multiplying the first equation by \( \mathbf{u} \) and integrating by parts using continuity of \( \rho_1 \)). This conservation law is a consequence of Noether’s theorem, which relates it to an explicit symmetry as follows.

a. Noether’s theorem. Consider the one-parameter family of infinitesimal variations generated by a functional \( \mathcal{G} \) defined by
\[
\delta \mathcal{G} := -\varepsilon \{ \mathcal{G}, \cdot \},
\]

(46)
where \( \varepsilon > 0 \) is small. The change generated by \( \mathcal{G} \) on any functional \( \mathcal{W} \) is
\[
\Delta_\mathcal{G} \mathcal{W} := \mathcal{W} [\mu + \delta_\mathcal{G} \mu] - \mathcal{W} [\mu] \sim \varepsilon \{ \mathcal{W}, \mathcal{G} \}.
\] (47)
Consequently,
\[
\Delta_\mathcal{G} \mathcal{W} - \Delta_\mathcal{G} \bar{\mathcal{W}} \sim \varepsilon \{ \mathcal{W}, \bar{\mathcal{G}} \}.
\] (48)
The condition \( \{ \mathcal{W}, \mathcal{G} \} = 0 \) is an expression of Noether’s theorem. It says that \( \mathcal{G} \) induces a symmetry in the most general sense that applying a transformation and “letting the time run” is the same as performing these operations in reverse order.\(^{53}\) Now if \( \mathcal{G} = 0 \), i.e., \( \mathcal{G} \) is an integral of motion, then \( \mathcal{G} \) induces a symmetry. Furthermore, in such a case \( \mathcal{G} \) transforms solutions into solutions: since \( \partial_t \Delta_\mathcal{G} \mu = \Delta_\mathcal{G} \partial_t \mu \), on one hand, and \( \Delta_\mathcal{G} \{ \mu, \mathcal{H} \} = \{ \Delta_\mathcal{G} \mu, \mathcal{H} \} \), on the other. The reciprocal of this theorem is not strictly true. To see this, one first must note that nontrivial solutions \( \mathcal{G} [\mu] \), called Casimirs, of
\[
\{ \mathcal{W}, \mathcal{C} \} = 0 \forall \mathcal{W} \iff \{ \mu, \mathcal{C} \} = \int \delta \mathcal{C} / \delta \mu = 0 \tag{49}
\]
represent integrals of motion that are not related to explicit symmetries.\(^{47}\) Indeed, \( \mathcal{C} \) does not generate any variation, i.e., \( \delta \mathcal{C} / \delta \mu = -\varepsilon \{ \mathcal{C}, \mu \} = \varepsilon \mathcal{W} \mu \equiv 0 \). As for the reciprocal of the Noether’s theorem, if \( \Delta_\mathcal{G} \mathcal{W} - \Delta_\mathcal{G} \bar{\mathcal{W}} = 0 \), then \( \mathcal{G} \) is equal to a distinguished function \( F(\mathcal{C}) \), for any \( F \), since \( \{ \mathcal{W}, F(\mathcal{C}) \} \equiv 0 \). However, if \( \mathcal{W} \) is replaced by \( \bar{\mathcal{W}} \equiv \mathcal{W} - \int F \mathcal{dt} \), which does not alter (48), then \( \mathcal{G} \) is conserved.\(^{53}\) Clearly, finding the generator \( \mathcal{G} \) of a given symmetry is not in general a simple task. It is easier to start with \( \mathcal{G} \) and check if this leaves \( \mathcal{H} \) invariant. In such a case it will be conserved, by (47), and as a result it will represent a symmetry.

b. Energy and momenta. Assume that \( \mathcal{G} \) is the generator of an infinitesimal time shift \( t \to t - \varepsilon \), i.e., \( \delta_\mathcal{G} \mu = -\varepsilon \partial_t \mu \). Clearly, \( \mathcal{G} = -\mathcal{H} \) (modulo a Casimir). Now, \( \mathcal{D}_\mathcal{G} \mathcal{H} \sim \varepsilon \mathcal{H} \) by (47). So invariance of the \( \mathcal{H} \) under time shifts, i.e., generated by \( \mathcal{H} \) itself, is equivalent to conservation of \( \mathcal{H} \) (modulo a Casimir). This implies \( \mathcal{D}_\mathcal{G} \mathcal{W} \equiv \Delta_\mathcal{G} \mathcal{W} \) for any \( \mathcal{W} \), by (48).

Let \( \mathcal{C} = \mathcal{C}(\mathcal{W}) \) indicate a direction in \( \mathcal{W} \subset \mathbb{R}^2 \), and suppose that \( \mathcal{G} \) is the generator of an infinitesimal translation along \( \mathcal{C} = \mathcal{C}(\mathcal{W}) \), \( \mathcal{C} \to \mathcal{C} - \varepsilon \). Then \( \delta_\mathcal{G} \mu = -\varepsilon \partial_\mathcal{C} \mu \). Furthermore, define \( \mathcal{M} \) such that \( \partial_\mathcal{C} \mu = -\{ \mu, \mathcal{M} \} \) and identify it with the \( \mathcal{C} \)-momentum. Then \( \mathcal{G} = \mathcal{M} \). Consequently, \( \Delta_\mathcal{G} \mathcal{H} \sim \varepsilon \mathcal{M} \) by (47). In this case, invariance of the \( \mathcal{H} \) under \( \mathcal{C} \)-translations, i.e., generated by \( \mathcal{M} \), is equivalent to conservation of \( \mathcal{M} \) (modulo a Casimir). Consequently, by (48), \( \Delta_\mathcal{G} \mathcal{H} \equiv \Delta_\mathcal{G} \mathcal{M} \) for any \( \mathcal{H} \).

Examples of conserved momenta are
\[
\mathcal{M}^z := \int \mathbf{m} \cdot \mathbf{x} = \int \rho_{1}(\mathbf{u} + f_{0y}y + \frac{1}{2}f_{0}r^{2})
\] (50)
when the flow domain \( D \) is zonally symmetric, and
\[
\mathcal{M}^\phi := \int (r \mathbf{\hat{r}} \times \mathbf{m}) \cdot \mathbf{z} = \int \rho_{1}(\mathbf{u}_{\phi}r + \frac{1}{2}f_{0}r^{2})
\] (51)
where \( r \) (resp. \( \phi \)) is the radial (resp. azimuthal) coordinate, and \( \mathbf{u}_{\phi} \) is the azimuthal velocity component, which requires \( D \) to be an axisymmetric domain of the \( f \)-plane.

Now, when \( D \) is axisymmetric, \( \mathcal{H} \) is invariant under \( x \)-translations. Now, in order for \( \{ \rho_{n} \mathcal{M}^{\phi} \} = -\nabla \cdot \rho_{n} \mathcal{M}^{\phi} = -\partial_{\rho} \rho_{n} \) \( \mathcal{M}^{\phi} \equiv \mathcal{M}^{\phi} \equiv 0 \) for all \( n_{\alpha} \). Hence, (50) follows.

In turn, when \( D \) is axisymmetric, \( \mathcal{H} \) is invariant under rotations. In this case, \( \{ \rho_{n} \mathcal{M}^{\phi} \} = -\nabla \cdot \rho_{n} \mathcal{M}^{\phi} = -\partial_{\rho} \rho_{n} \) \( \mathcal{M}^{\phi} \equiv \mathcal{M}^{\phi} \equiv 0 \) \( \mathcal{M}^{\phi} \equiv \mathcal{M}^{\phi} \equiv 0 \) for all \( n_{\alpha} \), and finally (51) follows.

Remark 3 The quantity \( \rho_{1}^{-1} \mathbf{m} \cdot \mathbf{x} = \mathbf{u} + f_{0y}y + \frac{1}{2}f_{0}r^{2} \) actually is the \( 3 \)-plane representation of the angular momentum per unit mass, with respect to the center of the Earth and in the direction of its axis of rotation.\(^{4,59}\) This should not be confused with \( \rho_{1}^{-1}(r \mathbf{\hat{r}} \times \mathbf{m}) \cdot \mathbf{z} = \mathbf{u}_{\phi}r + \frac{1}{2}f_{0}r^{2} \), which has the interpretation of angular momentum per unit mass, but as a symmetry-inducing generator of infinitesimal rotations on the \( f \)-plane.

c. Casimirs. These are listed in Table I. The particular form taken by each of them depends on the number of densities considered. Conservation of these quantities can be verified directly; to obtain them from (49), as we do it below, instead of the Poisson operator (44), it is more convenient to use \( \mathcal{J} = -\int \begin{pmatrix} q_{z} \cdot \nabla \left( \cdot \right) - \left( \cdot \right) \rho_{1}^{-1} \nabla \rho_{2} & \cdots & \left( \cdot \right) \rho_{1}^{-1} \nabla \rho_{\alpha+2} \\ \nabla \cdot \left( \cdot \right) & 0 & 0 & \cdots & 0 \\ \rho_{1}^{-1} \cdot \nabla \rho_{2} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1}^{-1} \cdot \nabla \rho_{\alpha+2} & 0 & 0 & \cdots & 0 \end{pmatrix} \right)
\] (52)
this gives a Poisson bracket, \( \{ \mathcal{W}, \mathcal{V} \} [\mathbf{u}, \rho_{1}, \rho] \equiv -\int \begin{pmatrix} q_{z} \cdot \frac{\delta \mathcal{V}}{\delta \mathbf{u}} \times \frac{\delta \mathcal{W}}{\delta \mathbf{u}} + \frac{\delta \mathcal{V}}{\delta \rho_{1}} \nabla \cdot \frac{\delta \mathcal{W}}{\delta \mathbf{u}} - \left( \frac{\delta \mathcal{V}}{\delta \rho_{n_{\alpha}}} \frac{\delta \mathcal{W}}{\delta \mathbf{u}} - \frac{\delta \mathcal{V}}{\delta \rho_{n_{\alpha}}} \frac{\delta \mathcal{W}}{\delta \mathbf{u}} \right) \end{pmatrix} \right)
\] (53)
which reduces to (45) by applying the chain rule,
\[
\frac{\delta}{\delta \hat{u}_{\alpha}} = \rho_1 \frac{\delta}{\delta \hat{m}},
\]  
(54)

\[
\frac{\delta}{\delta \rho_1} \mid_{\hat{u}, \hat{\rho}} = \frac{\delta}{\delta \rho_1} \mid_{m, \rho, \hat{\alpha}} + \frac{\hat{m}}{\rho_1} \cdot \frac{\delta}{\delta \hat{m}} + \frac{\rho_{n_\alpha}}{\rho_1} \frac{\delta}{\delta \rho_{n_\alpha}},
\]
(55)

\[
\frac{\delta}{\delta \hat{\rho}_{n_\alpha}} = \rho_1 \frac{\delta}{\delta \rho_{n_\alpha}},
\]
(56)

and substantial cancellation. The \( J \) in (52) is suggested when the momentum equation of the \( \mathbb{II}^{(0, \alpha)} + \mathbb{P} \) model is written in the form \( \hat{\delta} \hat{u} + \rho_1 \hat{q} \hat{z} \times \hat{u} + \nabla_\bot^2 \hat{u} + \hat{p}_1 \nabla \hat{p}_1 \hat{\varphi} = 0 \) and using (37) as a Hamiltonian, but viewed as a functional of \( (\hat{u}, \rho_1, \hat{\rho}) \). An explicit proof for the Jacobii identity of the (Poisson) bracket (53) for \( (\hat{u}, \rho_1, \hat{\rho}) \) is given in the appendix of Ripa.\(^{55}\)

Consider the case \( \alpha \geq 1 \). Using the Poisson operator in (52) and the variables \( (\hat{u}, \rho_1, \hat{\rho}) \), we have \( \{ \hat{\rho}_{n_\alpha}, \mathcal{C} \} = -\rho_1 \mathcal{C}_{\hat{u}} \cdot \nabla \rho_{n_\alpha}, \) from which it follows that \( \mathcal{C}_{\hat{u}} = 0 \) (since it must hold for all \( \alpha \geq 1 \)). This gives \( \{ \rho_1, \mathcal{C} \} = \nabla \cdot \mathcal{C}_{\hat{u}} = 0 \), trivially. Finally, \( \{ \mathcal{C}, \mathcal{C} \} = \nabla \mathcal{C}_{\hat{u}} - \rho_1 \mathcal{C}_{\hat{n}} \nabla \rho_{n_\alpha} = 0 \) is fulfilled by the Casimir in the bottom row of Table I (where \( F \) is arbitrary) since
\[
\delta \hat{q} = \rho_1 \hat{z} \times \nabla \times \hat{u} - \rho_1 \hat{q} \delta \rho_1,
\]  
(57)

which guarantees that \( \mathcal{C}_{\hat{u}} = 0 \) and \( \mathcal{C}_{\hat{\rho}_1} = F(\hat{\rho}) \).

Consider now the case \( \alpha = 0 \). We have \( \{ \hat{\rho}_2, \mathcal{C} \} = -\mathcal{C}_{\hat{u}} \nabla \rho_2 = 0 \), which is satisfied by \( \mathcal{C}_{\hat{u}} = -\hat{z} \times \nabla \mathcal{C}_{\hat{u}} \) for arbitrary \( F \) (the sign is arbitrary as is the sign of each of the terms in the Casimir listed in the second-to-bottom row of Table I). From \( \{ \mathcal{C}, \mathcal{C} \} = -\hat{q} \times \mathcal{C}_{\hat{u}} - \nabla \mathcal{C}_{\hat{u}} - \rho_1 \mathcal{C}_{\hat{\rho}_2} \nabla \hat{\rho}_2 = 0 \) it follows, on one hand, that \( \mathcal{C}_{\hat{\rho}_2} = -G(\hat{\rho}_2) \) for arbitrary \( G \) since \( \mathcal{C}_{\hat{u}} \perp \nabla \hat{\rho}_2 \). On the other hand, we have \( \mathcal{C}_{\hat{\rho}_2} \nabla \hat{\rho}_2 = \rho_1 (\hat{q} \nabla F + \nabla G) \), or, equivalently, \( \mathcal{C}_{\hat{\rho}_2} = \rho_1 (\hat{q} \tilde{F}' + \hat{G}') \) (modulo a constant). The Casimir follows upon integrating by parts with (57) in mind and by (5), which guarantees that \( F(\hat{\rho}_2)|_{\hat{D}} = \) const.

Finally, when \( \alpha = -1 \), from \( \{ \mathcal{C}, \mathcal{C} \} = -\hat{q} \mathcal{C}_{\hat{u}} - \nabla \mathcal{C}_{\hat{u}} = 0 \) one has \( \mathcal{C}_{\hat{u}} = q \hat{z} \times \nabla \mathcal{C}_{\hat{u}} \). Then, \( \{ \rho_1, \mathcal{C} \} = -\nabla \cdot \mathcal{C}_{\hat{u}} = 0 \) for any \( A, B \). While this specifies \( \mathcal{C}_{\hat{u}} \), obtaining \( \mathcal{C} \) from its functional derivatives as given seems not possible. So additional information is needed. This is given by \( \partial_1 \rho_1 F(\hat{q}) + \nabla \cdot \rho_1 F(\hat{q}) \hat{u} = 0 \), for any \( F \), because \( \hat{q} \) is transported by the flow of \( \hat{u} \) when \( n = 1 \) (36) and due to continuity of \( \rho_1 \). This already integrates to the Casimir in the second-to-top row of Table I, and thus suggests \( A = F \) and \( B = F' \). This is consistent with \( \mathcal{C}_{\hat{u}} = \nabla F' \times \hat{z} \). The Casimir follows upon invoking the admissibility condition, \( \mathcal{C}_{\hat{u}} \hat{u}|_{\hat{D}} = 0 \), which translates to \( F'(\hat{q})|_{\hat{D}} = \) const, with (57) in mind.

**Remark 4** The assumed iso-\( \hat{\rho}_{n_\alpha} \) nature of \( \partial \mathcal{D} \) (5), beyond guaranteeing conservation of Kelvin circulation(s) along (cf. IIIA 1), is needed to warrant the existence of Casimirs and hence possession of generalized Hamiltonian structure by the \( \mathbb{II}^{(0, \alpha)} + \mathbb{P} \) when \( \alpha = 0 \). Yet, this is no more than the admissibility condition a Casimir is assumed to satisfy, namely, \( \mathcal{C}_{\hat{u}} \hat{u}|_{\partial \mathcal{D}} = 0 \) (or \( \mathcal{C}_{\hat{m}} \hat{u}|_{\partial \mathcal{D}} = 0 \)).

### IV. THERMODYNAMICS

Having covered the geometry of the \( \mathbb{II}^{(0, \alpha)} + \mathbb{P} \) in detail, I proceed to interpret the model as a rotating shallow-water model with generalized thermodynamics. This is done by first showing that the rotating shallow-water model itself is a special case, living on an invariant subspace of the system. Two thermodynamically active submodels are then discussed, prior to introducing the most general class of thermodynamically active model(s). For each model family I present the PE model along with the corresponding QG version, a basic paradigm for the study of geophysical flow stability.\(^{20}\)

The rotating shallow-water model is in particular used to illustrate the procedure to derive in each case a QG approximation, and also to introduce generic geometric mechanics results that apply to the rest of the models. Some background physical setup is necessary, which I introduce first.

Let \( h(x, t) \) and \( H_r = \) const be instantaneous and reference (i.e., unperturbed) layer thickness, respectively. Let \( g_b = \) const be the buoyancy jump at the base of the layer in the reference state (I will focus on the reduced gravity case; the free-surface versions of the models to be discussed follow upon minimal reinterpretation of the parameters, such as reinterpretting \( g_b \) as \( g \), the acceleration of gravity). The instantaneous buoyancy field,
\[
\vartheta(x, z, t) := -g \frac{\rho_{\text{top}}(x, z, t) - \rho_{\text{bot}}}{\rho_0},
\]
(58)

where \( \rho_{\text{top}} \) is the density in top (active) layer, \( \rho_{\text{bot}} = \) const is the density in the bottom (inactive) layer, and \( \rho_0 = \) const is a reference density used in the Boussinesq approximation. Let \( N_r \) be the reference Brunt–Vaissala frequency, namely, the vertical derivative of the unperturbed form of (58). We will assume \( N_r = \) const, meaning that the reference state has uniform vertical stratification, i.e., the unperturbed form of (58) depends linearly in \( z \). In such a case, the stratification within the (active) layer is conveniently measured by the nondimensional parameter\(^{61}\)
\[
s := \frac{N_r^2 H_r}{g_b} > 0.
\]
(59)

The quantities
\[
R^2 := \frac{g_b H_r}{\bar{f}^2}, \quad L^2 := s R^2,
\]
(60)

where \( R \) is the external (equivalent barotropic) Rossby deformation radius and \( L \) is proportional to the-
nal Rossby deformation radius (for normal mode perturbations on a reference state with uniform stratification). When \( s \) is small, i.e., the stratification is weak, the above scales are well separated. Finally, consider the rescaled vertical coordinate\(^{56} \)

\[
\sigma := 1 + 2 \frac{z}{h(x,t)}, \tag{61}
\]

which allows one to compute a (vertical) average over \(-h(x,t) \leq z \leq 0\) simply as one half the integral over \(-1 \leq \sigma \leq 1\).

**A. The HL+ family**

1. **The HL+PE**

As noted above, the set \( M_\alpha := \{  \tilde{\rho}_\alpha = \text{const} \} \) is an invariant manifold of the \( IL^{(0,\alpha)+}\)PE system: once initialized on \( M_\alpha \), the system remains on \( M_\alpha \), for all \( t \). Identifying \( \rho_1 \) with layer thickness \( (h) \), on \( M_\alpha \), the dynamics are controlled by

\[
\partial_t \tilde{m} + \mathcal{L}_u \tilde{m} - h\nabla K + \nabla h^2 \varphi_{-1+}(h) = 0, \quad \partial_t h + \nabla \cdot h\tilde{u} = 0. \tag{62}
\]

This is system (8) for \( \alpha = -1 \), with the identification \( \rho_1 = h \) (so \( \tilde{m} = m(h + f) \) now, and on), and noting that \( \varphi_{-1} = \varphi_{-1+}(h) \). For arbitrary \( \varphi_{-1} \), the resulting system can be viewed as the equations for a shallow layer of constant density fluid on the \( \beta \)-plane with a generalized pressure “gradient” force, given by \(-h^{-1}\nabla h^2 \varphi_{-1+}(h) \). Making \( \varphi_{-1+} = \frac{1}{2} gh_0 h \), the standard rotating shallow-water equations follow. We refer to these models respectively as HL+PE and HLPE (instead of \( IL^{(0,-1)+}\)PE and \( IL^{(0,-1)}PE \)) where HL stands for homogeneous layer, to make explicit that they cannot accommodate thermodynamics as the fluid density is constant or, equivalently, \( \vartheta = g_0 \). Clearly, both model equations admit Euler–Poincare variational formulations and form Lie-Poisson systems, and thus the potential energy density choice \( \varphi_{-1+} = \frac{1}{2} gh_0 h + F(h) \) for some \( F \) enables the investigation of isothermal mixed-layer dynamics with forcing in a conservative context. This justifies the \( + \) sign notation, used to mean that the HL+PE adds the noted potentially additional feature to the HLPE. This notation is similarly adopted below.

2. **The HL+QG**

Let \( \varepsilon > 0 \) be a small parameter taken to represent a Rossby number, relating the ratio of inertial to Coriolis forces, e.g.,

\[
\varepsilon = \frac{U}{|f_0| R} \ll 1, \tag{63}
\]

where \( U \) is a characteristic velocity magnitude. In the QG scaling\(^{50} \)

\[
(|\tilde{u}|, h - H_t, \partial_t, \beta y) = O(\varepsilon U, \varepsilon R, \varepsilon f_0, \varepsilon f_0) \quad \text{and} \quad (\sigma, \xi, \xi_y, \xi_{yy}) = O(1, \varepsilon, \varepsilon^2, \varepsilon^2) \tag{64}
\]

Consistent with this, we write

\[
\begin{aligned}
\tilde{u}/U &= \hat{z} \times \nabla \tilde{\psi}/U + \ldots, \\
\tilde{h}/H_t &= 1 + \frac{\hat{z} \times \nabla \tilde{\psi} y}{f_0 R_t} + \ldots, \\
O : \varepsilon^2 &
\end{aligned}
\tag{65}
\]

where \( \tilde{\psi}(x,t) \) is a streamfunction and

\[
R_{-1+}^2 := \frac{2H_t \varphi'_{-1+}(H_t) + H_t^2 \varphi'^{2}_{-1+}(H_t)}{f_0^2}. \tag{66}
\]

The potential vorticity, \( \bar{q} = H_t^{-1}(\nabla^2 \bar{\psi} - R_{-1+}^2 \bar{\psi} + f) + O(\varepsilon^2) \). Since \( \bar{q} \) is transported for the HL class (the rhs of (36) vanishes), to lowest-order in \( \varepsilon \), i.e., \( O(\varepsilon^2) \), one has that the dynamics are controlled by

\[
\partial_t \bar{\xi} + \{ \bar{\psi}, \bar{\xi} \}_{xy} = 0 \tag{67a}
\]

where

\[
\nabla^2 \bar{\psi} - R_{-1+}^2 \bar{\psi} = \bar{\xi} - \beta y. \tag{67b}
\]

Equation (67a) with the invertibility principle (67b) forms the HL+QG model. The standard HLQG is recovered upon setting \( \varphi_{-1+} = \frac{1}{2} gh_0 H_t \), which gives \( R_{-1+} = R \).

**Remark 5** A peculiarity (cf., e.g., Shepherd\(^{66} \)) of (67) is that while \( \gamma = 0 \) (namely, constancy of the circulation of \( \bar{u} \) along \( \partial D \)) holds to lowest-order in \( \varepsilon \), it cannot be deduced from (67), unless \( R_{-1+} \uparrow \infty \) (i.e., the bottom of the layer is effectively rigid). In the general case with a soft bottom, i.e., when \( R_{-1+} \) is finite, the evolution equation (67a) must be complemented with \( \gamma = 0 \). More generally, when \( D \) includes islands, i.e., when \( D \) is multiply connected, constancy of \( \gamma \) along the boundary of each of these islands must be appended. If \( D \) is a zonal channel, constancy of \( \gamma \) along both coasts must be included.

a. **Hamiltonian structure**. The HLQG system is well-known\(^{17} \) to possess a generalized Hamiltonian structure. Such a structure is conveyed to (67) by the Hamiltonian (energy),

\[
\mathcal{H}[\xi] := \frac{1}{2} \int |\nabla \bar{\psi}|^2 + R_{-1+}^{-2} \bar{\psi}^2 = -\frac{1}{2} \int \bar{\psi} (\bar{\xi} - \beta y) \equiv -\frac{1}{2} \int (\bar{\xi} - \beta y) (\nabla^2 - R_{-1+}^{-2})^{-1} (\bar{\xi} - \beta y), \tag{68}
\]

by \( \hat{z} \times \nabla \bar{\psi} \cdot \tilde{u} \big|_{\partial D} = 0 \) and where \((\nabla^2 - R_{-1+}^{-2})^{-1}(\bar{\xi} - \beta y)\) represents a convolution of \( \bar{\xi} - \beta y \) with the Green’s function of the elliptic problem (67b), and the Poisson operator

\[
\mathcal{J} = -\{ \bar{\psi}, \cdot \}_{xy}, \tag{69}
\]

so the Poisson bracket

\[
\{ \bar{\psi}, \bar{\xi} \}_{xy} = \int \bar{\xi} \left\{ \frac{\delta \bar{\psi}}{\delta \xi}, \frac{\delta \bar{\psi}}{\delta \bar{\xi}} \right\}_{xy}, \tag{70}
\]
with the admissibility condition, \( \nabla \mathcal{H}_c \cdot \mathbf{n} |_{\partial D} = 0 \) for all functionals \( \mathcal{H} \) of the state variables, which above have been taken to be composed of simply \( \xi \) under the assumption that \( D \) does not include islands (or has the topology of a zonal channel). If this is not the case, the state variables should be augmented to \( \mu = (\xi, \gamma_1, \gamma_2, \ldots) \) with as many circulations as appropriate.\(^65\) We will ignore this technicality here and below, which does not have an immediate consequence for our purposes (stability analyses\(^63\) do require one to account for it, though). Clearly, (67a) follows from \( \partial_t \xi = \{ \xi, \mathcal{H} \} = \mathcal{J} \mathcal{H}_\xi \) since \( \mathcal{H}_\xi = -\psi \).

For future reference, we note that (70) is a special case of Lie–Poisson brackets of the general form\(^67\)

\[
\{ \mathcal{W}, \mathcal{V} \}_\mu = W_{ab}^c \int \mu^c \left( \frac{\delta \mathcal{W}}{\delta \mu^a}, \frac{\delta \mathcal{V}}{\delta \mu^b} \right)_{xy}, \tag{71}
\]

where the constants \( W_{ab}^c \) transform like the components of a (2,1)-tensor under linear transformations of \( \mu \). The tensor \( W \) is symmetric in its upper indices, viz.,

\[
W_{ab}^c = W_{ba}^c, \tag{72a}
\]

so the bracket (71) is antisymmetric. Furthermore, \( W_{ab}^c W_{'c}^{ba'} = W_{ac}^b W_{ba}^c \), which guarantees it satisfies the Jacobi identity. If \( W \) is viewed as a collection of matrices \( W(b) \) with \((c, a)\)-th entry given by \( W_{ac}^b \), the latter property means that these matrices commute, namely

\[
W(a)W(b) = W(b)W(a), \quad a \neq b. \tag{72b}
\]

In (70), \( W = 1 \), simply. (A deeper geometric interpretation of the generalized Hamiltonian formulation of QG systems, which is different than that of PE systems, is outlined in App. C.)

**Remark 6** Generalized Hamiltonian systems with Lie–Poisson brackets of the form seem (71) do not seem possible to be obtained via a Legendre transformation as are systems with brackets of the form (45). Thus Euler–Poincare variational formulations may not exist or at least are difficult to be derived for QG-type systems. However, ad-hoc variational formulations have been proposed in the literature.\(^17,42,68\)

**b. Conservation laws.** Symmetry-related integrals of motion of (67) are the energy (\( \mathcal{H} \)), the zonal momentum

\[
\mathcal{H}^z := \int y \xi \tag{73}
\]

(with \( D \) is \( x \)-symmetric), and the angular momentum

\[
\mathcal{H}^\theta := -\int r \xi \tag{74}
\]

(on an \( f \)-plane when \( D \) is an axisymmetric domain).\(^66\) The Casimirs of the bracket (71) are \( \mathcal{E} = \int F(\xi) \) for any \( F \) (e.g., Morrison\(^37\)).

**B. The IL\(^0+\) family**

1. **The IL\(^0+\) PE**

We are now ready to introduce the first family of shallow-water models with thermodynamics. This follows from the IL\(^0+\) model (8) upon setting \( \alpha = 0 \). Making \( \rho_1 = h \), as before, and \( \rho_2 = \bar{\theta} \), viz., the vertical average of the buoyancy field (58) across the layer extent, in (8) we obtain

\[
\begin{align*}
\partial_t \bar{\mathbf{m}} + \mathcal{L}_{\mathbf{u}} \bar{\mathbf{m}} - h \nabla K + \nabla h^2 \partial_h \varphi_0 + (h, \bar{\theta}) &= 0, \\
\partial_t h + \nabla \cdot \mathbf{u} &= 0, \\
\partial_t \bar{\theta} + \mathbf{u} \cdot \nabla \bar{\theta} &= 0.
\end{align*} \tag{75}
\]

Choosing the potential energy density as

\[
\varphi_{0+} := \varphi_0 = \frac{1}{2} h \bar{\theta} \tag{76}
\]

reduces (75) to the IL\(^0\)PE model,\(^55\) which is the rotating shallow-water model with buoyancy varying in horizontal position and time, i.e., thermodynamically active, that we sought to extend. No variation in the vertical is allowed for the dynamical fields, which justifies the superscript in IL\(^0\)PE. The IL\(^0\)PE formally follows from the IL\(^\infty\)PE by replacing the horizontal velocity and buoyancy in the model by their vertical averages, \( \bar{\mathbf{u}} \) and \( \bar{\theta} \), respectively, and further by vertically averaging the resulting pressure gradient, namely,

\[
\nabla p = \nabla (z + h) \bar{\theta}, \tag{77}
\]

which gives

\[
\nabla p = \frac{1}{2} h^{-1} \nabla h^2 \bar{\theta} \equiv h^{-1} \nabla h^2 \partial_h \varphi_0 \tag{78}
\]

with \( \varphi_0 \) as in (76). It must be noted, however, that while the horizontal velocity is set to \( \bar{\mathbf{u}} \), because \( \nabla \partial_z p = \nabla \bar{\theta} \), the velocity includes\(^57\) a linear vertical shear, implicitly, as it follows from the thermal–wind balance, which dominates at low frequency; we will return to this in the section that follows. Clearly, (78) is a special case of\(^57\) \( h^{-1} \nabla h^2 \partial_h \varphi_{0+} \) (for arbitrary \( \varphi_{0+} \)), which extends the IL\(^0\)PE to (75), referred here to as the IL\(^0\)PE. With a generalized pressure “gradient” force, the IL\(^0\)PE model has the potential of expanding the realm of applicability of the IL\(^0\)PE system. A suitable choice of \( \varphi_{0+} \) can allow one to study forced–dissipative mixed-layer hydrodynamics with thermodynamics in a conservative setting, as the IL\(^0\)PE admits an Euler–Poincare formulation and has a generalized Hamiltonian structure.

a. **Connection with Morrison and Greene\(^40\).** The nonrotating form of the IL\(^0\)PE, i.e., with \( f = 0 \), follows from the magnetohydrodynamics (MHD) model considered by Morrison and Greene\(^40\) upon neglecting the magnetic field, i.e., \( \mathbf{B} = 0 \) in the notation of that paper, particularizing the resulting system to two-space dimensions, and interpreting the mass density \( (\rho, \text{in the notation of Morrison and Greene} \text{\cite{MorrisonGreene}}) \) and entropy per unit mass (\( s \), in
their notation) variables as layer thickness \( h \) and vertically averaged buoyancy \( \tilde{\theta} \), respectively, in the internal energy per unit mass, \( U(\rho, s) \). That the (nonrotating) IL\(^0\)PE is thus equivalent to Morrison and Greene’s hydrodynamics system with \( U(\rho, s) = \frac{1}{2} \rho s \) had remained elusive until present.\(^{48}\)

2. The IL\(^0+\) QG

Consistent with the QG scaling \((64)\), consider

\[
\begin{align*}
\frac{\dot{\mathbf{u}}}{U} &= \frac{\dot{\mathbf{z}} \times \nabla \tilde{\psi}}{U} + \ldots, \\
h/H_t &= 1 + \frac{\tilde{\theta} - 2\partial H_t}{\tilde{\eta} R_{0+}} + \ldots, \\
\frac{\tilde{\theta}}{g_b} &= 1 + \frac{2}{\tilde{\eta} R_{0+}} \tilde{\psi} + \ldots, \\
O : 1 &= \varepsilon^2
\end{align*}
\]

(79)

where

\[
R_{0+}^2 := \frac{2H_t \partial H_t, \Phi_{0+} + H_t^2 \partial^2 H_t, \Phi_{0+}}{\tilde{\eta}^2},
\]

(80)

and the shorthand notation \( \Phi_{0+} \) for \( \varphi_{0+}(H_t, g_b) \) was introduced. The notation \( \psi_{\sigma} \) is clarified by noting that

\[
\partial_z \mathbf{u} = \frac{2g_b}{\tilde{\eta} R_{0+}^2} \mathbf{z} \times \nabla \psi_{\sigma} + O(\varepsilon^2)
\]

(81)

is the vertical shear that the horizontal velocity implicitly has by the thermal-wind balance. When \( \Phi_{0+} = \frac{1}{2} H_t g_b \), the streamfunction would read (with \((60)\) in mind) as

\[
\psi = \tilde{\psi} + (1 + 2 \frac{H_t}{\tilde{\eta}^2}) \psi_{\sigma},
\]

(82)

which better justifies the notation (note that \( \sigma = 1 + 2 \frac{H_t}{\tilde{\eta}^2} + O(\varepsilon) \) by the QG scaling). Plugging the expansions \((79)\) in the potential vorticity equation \((36)\) (with \( \rho_1 = h \) and \( \rho_2 = \tilde{\theta} \)) and generic potential energy density \( \varphi_{0+} \), and the equation for \( \rho_3(= \tilde{\theta}) \) in \((75)\), we obtain, to \( O(\varepsilon^2) \), i.e., to lowest-order in \( \varepsilon \), the following set:

\[
\begin{align*}
\partial_t \xi + \{\xi, \tilde{\psi}\}_{xy} - R_{0+}^{-2} \{\tilde{\psi}, \psi_{\sigma}\}_{xy} &= 0, \\
\partial_t \psi_{\sigma} + \{\tilde{\psi}, \psi_{\sigma}\}_{xy} &= 0,
\end{align*}
\]

(83a)

where \( R_{0+}^2 := (2\partial H_t, \Phi_{0+})^{-1} R_{0+}^2 \) and

\[
\nabla^2 \tilde{\psi} - R_{0+}^{-2} \tilde{\psi} = \xi - R_{0+}^{-2} \psi_{\sigma} - \beta \psi =: H_{0+}(\xi, \psi_{\sigma}).
\]

(83b)

Equations \((83a)\) with the invertibility principle \((83b)\) form the IL\(^0+\) QG system. The IL\(^0\)QG model\(^{10}\) follows as the special case \( \Phi_{0+} = \frac{1}{2} g_b H_t \), for which \( R_{0+} = R_{0+} = R \).

a. Hamiltonian structure. The IL\(^0+\) QG system \((83)\) possesses a generalized Hamiltonian structure endowed by the Hamiltonian

\[
\mathcal{H}[\xi, \psi_{\sigma}] := \frac{1}{2} \int |\nabla \tilde{\psi}|^2 + R_{0+}^{-2} \tilde{\psi}^2 = -\frac{1}{2} \int \tilde{\psi} H_{0+}(\xi, \psi_{\sigma})
\]

\[
= -\frac{1}{2} \int H_{0+}(\xi, \psi_{\sigma})(\nabla^2 - R_{0+}^{-2})^{-1} H_{0+}(\xi, \psi_{\sigma}),
\]

(84)

since \( \mathbf{z} \times \nabla \psi \cdot \mathbf{n}|_{\partial D} = 0 \) and where \( (\nabla^2 - R_{0+}^{-2})^{-1} H_{0+}(\xi, \psi_{\sigma}) \) represents a convolution of \( H_{0+}(\xi, \psi_{\sigma}) \) with the Green’s function of the elliptic problem \((83b)\), and the Poisson operator

\[
\mathcal{J} = -\left( \begin{array}{cc}
\{\xi, \cdot\}_{xy} & \{\psi_{\sigma}, \cdot\}_{xy} \\
\{\psi_{\sigma}, \cdot\}_{xy} & 0
\end{array} \right),
\]

(85)

which leads to the Lie–Poisson bracket \((71)\) on \( \mu = (\xi, \psi_{\sigma}) \) with \( W_1^{11} = W_2^{12} = W_2^{21} = 1 \) and zero otherwise (note that \( W \) satisfies the required symmetry property \((72a)\) and further the resulting \( W \)'s satisfy \((72b)\) since \( W_1^{(1)} = \text{Id}^2 \times^2 \) and

\[
W^{(2)} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},
\]

(86)

which commute). The first and second equations in \((83a)\) follow from \( \partial_t \xi = \{\xi, \tilde{\psi}\} \) and \( \partial_t \psi_{\sigma} = \{\psi_{\sigma}, \tilde{\psi}\} \), respectively, upon noting that

\[
\frac{\delta \mathcal{H}}{\delta \xi} = -\tilde{\psi}, \quad \frac{\delta \mathcal{H}}{\delta \psi_{\sigma}} = \tilde{R}_{0+}^{-2} \tilde{\psi}.
\]

(87)

Remark 7 The Poisson operator \((85)\) and corresponding Lie–Poisson bracket turn out to be the same as those for “low–β” reduced magnetohydrodynamics\(^{41}\) and incompressible, nonhydrostatic, Boussinesq fluid dynamics on a vertical plane.\(^{39}\)

b. Conservation laws. The Hamiltonian \((\mathcal{H})\) of the IL\(^0+\) QG \((84)\) is invariant under time shifts. As the generator of infinitesimal such transformations, by Noether’s theorem \( \mathcal{H} \) (i.e., the energy) is preserved under dynamics of \((83)\). In a zonally symmetric domain \( D \) of the \( β \)-plane, \( \mathcal{H} \) is invariant under \( x \)-translations. The corresponding generator \((\mathcal{M}^x)\) must be conserved. Since \( \mathcal{M}^x \) must satisfy \( \{\psi_{\sigma}, \mathcal{M}^x\}_{xy} = \partial_x \psi_{\sigma} \), it follows that \( \mathcal{M}^x \) is given by \((73)\), just as in the HL+QG model, sufficiently and necessarily: given any two vectors \( \mathbf{a}, \mathbf{b} \) on \( D, \mathbf{a} \times \mathbf{b} = a_1 \mathbf{z} \) iff \( b = (1, 0) \). This immediately gives \( \{\xi, \mathcal{M}^x\}_{xy} + \{\psi_{\sigma}, \mathcal{M}^x\}_{xy} = \partial_x \xi \). The conservation of \( \mathcal{M}^x \) can be verified directly from \((83)\) with a careful examination of the boundary terms. On an \( f \)-plane the Hamiltonian is invariant under rotations in an axisymmetric domain. The generator \( \mathcal{M}^\theta \) of infinitesimal rotations is conserved by Noether’s theorem. This must satisfy \( \{\psi_{\sigma}, \mathcal{M}^\theta\}_{r\phi} = r \partial_{\phi} \psi_{\sigma} \), which holds iff \( \mathcal{M}^\phi \) is given by \((74)\), as for the HL+QG. This immediately makes \( \{\xi, \mathcal{M}^\theta\}_{r\phi} + \{\psi_{\sigma}, \mathcal{M}^\theta\}_{r\phi} = r \partial_{\phi} \xi \). Finally, the Casimirs of \((83)\) are given by \( \mathcal{C} = \int \xi F(\psi_{\sigma}) + G(\psi_{\sigma}) \) where \( F, G \) are arbitrary, which have been known for a long time.\(^{34,41,56}\)

C. The IL\(^0(0,1)^+\) family

1. The IL\(^0(0,1)^+\) PE

One step above the IL\(^0\) class in dynamical richness is the family of shallow-water models with thermodynamics
where $\varphi_3$. cf. (90).

The result is dient, viz., Vertically average the ensuing horizontal pressure gra-

age, $\bar{}$ series of steps as follows.

layer model to study the interplay between hydrodynamics

interfaces, a model similar to the IL

Remark 8 has quadratic vertical shear.

vary, albeit linearly, in the vertical; cf. (88). Implicitly by

follows from the freedom of the buoyancy in the IL

restratification mixed-layer early with it. Important processes such as

ordinate, while the second slot that buoyancy varies lin-

means that velocity does not vary with the vertical co-

vanishes for vertically averaged horizontal velocity, $\bar{u} = \bar{u}$, and $z$-vertical vector $w = w_0$, i.e., as given in (93).

Clearly, (92) represents a particular case of

$h^{-1}\nabla h^2 \varphi_{(0,1)}(\bar{h}, \bar{\varphi}, \varphi_\sigma)$ for arbitrary $\varphi_{(0,1)+}$, which extends the IL$^{(0,1)}$PE to (89). We refer to (89) as the IL$^{(0,1)+}$PE, which has the power of expanding the domain of applicability of the IL$^{(0,1)}$PE. As noted for the IL$^0$ class, an appropriate choice of $\varphi_{(0,1)+}$ enables investigation of mixed-layer processes, such as restratification by baroclinic instability, with forcing/dissipation, but in a Hamiltonian (i.e., conservative) setting.

2. The IL$^{(0,1)}$ QG

Next I will present a derivation of the QG approximation to the IL$^{(0,1)}$PE. Working with a generic energy potential density $\varphi_{1+}$ complicates the algebra without shedding much light on the problem. If a specific choice of $\varphi_{1+}$ turns out to be relevant, then steps similar to those below can be taken to derive the corresponding QG set of equations.

With the above in mind, consistent with the QG scaling (64), the following expansions are proposed:

\begin{equation}
\begin{aligned}
\frac{\bar{u}}{U} &= \frac{\bar{z} \times \nabla \bar{\psi}}{U} + \ldots, \\
\frac{h}{H_t} &= 1 + \frac{\psi - \psi_a + \frac{4}{3} \psi_{\sigma}}{f_0 R_1^0} + \ldots, \\
\frac{\bar{\varphi}}{g_0} &= 1 + \frac{2}{f_0 R_1^0} \psi_{\sigma} + \ldots, \\
\varphi_{\sigma} / \frac{1}{2} N_2^2 H_t &= 1 + \frac{8}{f_0 s^2} \psi_{\sigma^2} + \ldots, \\
O &:= 1 - \frac{\varepsilon}{\bar{z}^2}
\end{aligned}
\end{equation}

where $R_1^0 := (1 - \frac{1}{3} s) R^2$. The origin of $\psi_{\sigma}(x, t)$ and $\psi_{\sigma^2}(x, t)$ is clarified upon realizing that

\begin{equation}
\partial_z \tilde{u} = \frac{2 g_0}{f_0^2 R^2} \frac{\bar{z} \times \nabla \psi_{\sigma}}{2} + 4 N_2^2 H_t \frac{1 + 2 \bar{\varphi}_\sigma}{f_0 s R^2} + O(\varepsilon^2)
\end{equation}

(96) is the vertical shear that the horizontal velocity implicitly has by the thermal-wind balance. The streamfunction would then read (with (59)–(60) in mind) as

\begin{equation}
\psi = \bar{\psi} + \left(1 + 2 \psi_{\sigma} \frac{2}{s R_t} \right) \psi_{\sigma} + \left(1 + 2 \frac{\bar{\varphi}_\sigma}{s R_t} \right)^2 \psi_{\sigma^2}.
\end{equation}

Plugging the expansions (95) in the potential vorticity equation (36) (with $\rho_1 = h$, $\rho_2 = \bar{\varphi}$, and $\rho_3 = \varphi_\sigma$, the

arise from (8) with $\alpha = 1$, assuming that the buoyancy field not only varies in the horizontal and time, but also in the vertical, linearly. Namely,

\begin{equation}
\partial_t (x, \sigma, t) = \partial_t (x, t) + \sigma \partial_t (x, t).
\end{equation}

Making $\rho_1 = h$ and $\rho_2 = \bar{\varphi}$ as in the previous section, and further setting $\rho_3 = \varphi_\sigma$, the equations of the model read

\begin{equation}
\begin{aligned}
\partial_t \bar{m} + \mathcal{L}\bar{m} - h \nabla K + \nabla h^2 \partial_t \varphi_{1+}(h, \bar{\varphi}, \varphi_\sigma) = 0, \\
\partial_t \bar{h} + \nabla \cdot \bar{u} = 0, \\
\partial_t \bar{\varphi} + \bar{u} \cdot \nabla \bar{\varphi} = 0, \\
\partial_t \varphi_\sigma + \bar{u} \cdot \nabla \varphi_\sigma = 0,
\end{aligned}
\end{equation}

where $\varphi_{1+}$ is arbitrary.

The specific potential energy density choice

\begin{equation}
\varphi_{1+} = \varphi_1 := \frac{1}{3} \left( \bar{\varphi} - \frac{2}{3} \varphi_\sigma \right)
\end{equation}

leads to a model, which will be called the IL$^{(0,1)}$PE, that offers a better representation of thermodynamics than the IL$^0$ class. The first slot in the superscript in IL$^{(0,1)}$PE means that velocity does not vary with the vertical coordinate, while the second slot that buoyancy varies linearly with it. Important processes such as mixed-layer restratification, which can be expected by baroclinic instability, can now be represented. This follows from the freedom of the buoyancy in the IL$^{(0,1)}$PE to vary, albeit linearly, in the vertical; cf. (88). Implicitly by the thermal–wind balance, the velocity in the IL$^{(0,1)}$PE has quadratic vertical shear.

Remark 8 Ignoring forcing and fluid exchanges across interfaces, a model similar to the IL$^{(0,1)}$PE was used by Schopf and Cane in an intermediate layer of a multi-layer model to study the interplay between hydrodynamics and thermodynamics in the equatorial near-surface ocean.

The IL$^{(0,1)}$PE can be obtained from the IL$^\infty$PE in a series of steps as follows.

1. Replace the horizontal velocity with its vertical average, $\bar{u}$, and the buoyancy field by (88).

2. Vertically average the ensuing horizontal pressure gradient, viz.,

\begin{equation}
\nabla \tilde{p} = (z + h) \bar{\varphi} + z \left(1 + \frac{\bar{\varphi}}{h} \right) \nabla \varphi_\sigma + \left(\bar{\varphi} - \frac{z^2}{h^2} \varphi_\sigma \right) \nabla h.
\end{equation}

The result is

\begin{equation}
\nabla \tilde{p} = \frac{1}{2} h^{-1} \nabla h^2 \left(\bar{\varphi} - \frac{2}{3} \varphi_\sigma \right) \equiv \nabla h^{-1} \nabla h^2 \partial_t \varphi_1;
\end{equation}

cf. (90).

3. Note that consistent with the horizontal velocity being independent of the vertical coordinate is the vertical velocity depending linearly in it:

\begin{equation}
w_{\tilde{u}} := \frac{\sigma - 1}{2} \frac{D h}{D t}.
\end{equation}

4. Realize that transportation under the flow of $\bar{\varphi}$ of $\bar{\varphi}$ and $\varphi_\sigma$ implies transportation of (88) under the flow of $(\bar{u}, w_0)$.

To see the last step, it is best to write the three-dimensional material derivative in the IL$^\infty$PE as $\partial_t |_\sigma + \bar{u} \cdot \nabla |_\sigma + w_\sigma \partial_\sigma$, where the $\sigma$-vertical velocity

\begin{equation}
w_\sigma := \frac{2}{h} \left(1 - \frac{\sigma}{h} \left(\partial_t h + \bar{u} \cdot \nabla h + w\right)\right)
\end{equation}

vanishes for vertically averaged horizontal velocity, $\bar{u} = \bar{u}$, and $z$-vertical vector $w = w_0$, i.e., as given in (93).
potential energy density (90), specifically) and the equations for \( \hat{\rho}_2(= \hat{\nu}) \) and \( \hat{\rho}_3(= \hat{\sigma}_s) \) and in (75), we obtain, to \( O(\varepsilon^2) \), the following system:

\[
\begin{align*}
\partial_t \tilde{\xi} + \{\tilde{\psi}, \xi\}_{xy} - R_1^{-2}\{\tilde{\psi}, \psi_s - \frac{2}{3}\psi_s^2\}_{xy} &= 0, \\
\partial_t \psi_s + \{\tilde{\psi}, \psi_s\}_{xy} &= 0, \\
\partial_t \psi_{s_2} + \{\tilde{\psi}, \psi_{s_2}\}_{xy} &= 0,
\end{align*}
\]  

(98a)

where

\[
\nabla^2 \tilde{\psi} - R_1^{-2} \tilde{\psi} = \tilde{\xi} - R_1^{-2}(\psi_s - \frac{2}{3}\psi_s^2) - \beta y \\
=: H_1(\tilde{\xi}, \psi_s, \psi_{s_2}).
\]  

(98b)

Equations (98a) with the invertibility principle (98b) form the IL\(^{(0,1)}\)QG model. The IL\(^{0}\)QG, i.e., (83a) with \( \tilde{R}_{0+} = R_{0+} = R \), is recovered upon ignoring \( \psi_{s_2} \) and making \( s = 0 \).

**a. Hamiltonian structure.** The IL\(^{(0,1)}\)QG system (98) possesses a generalized Hamiltonian structure conveyed by the Hamiltonian

\[
\mathcal{H}(\tilde{\xi}, \psi_s, \psi_{s_2}) := \frac{1}{2} \int |\nabla \tilde{\psi}|^2 + R_1^{-2} \tilde{\psi}^2 \\
= -\frac{1}{2} \int \tilde{\psi} H_1(\tilde{\xi}, \psi_s, \psi_{s_2}) \\
= -\frac{1}{2} \int H_1(\nabla^2 - R_1^{-2})^{-1} H_1, \quad (99)
\]

where \( \hat{z} \times \nabla \tilde{\psi} \cdot \hat{n}|_{\partial D} = 0 \) was used and \( (\nabla^2 - R_1^{-2})^{-1} H_1(\tilde{\xi}, \psi_s, \psi_{s_2}) \) is a convolution of \( H_1(\tilde{\xi}, \psi_s, \psi_{s_2}) \) with the Green’s function of the elliptic problem (83b), and the Poisson operator

\[
\mathcal{J} = - \begin{pmatrix}
\{\tilde{\xi},\cdot\}_{xy} & \{\psi_s,\cdot\}_{xy} & \{\psi_{s_2},\cdot\}_{xy} \\
\{\psi_s,\cdot\}_{xy} & 0 & 0 \\
\{\psi_{s_2},\cdot\}_{xy} & 0 & 0
\end{pmatrix},
\]

(100)

which leads to the Lie–Poisson bracket (71) on \( \mu = (\tilde{\xi}, \psi_s, \psi_{s_2}) \) with \( W_{11}^1 = W_{22}^2 = W_{33}^3 = W_{12}^2 = W_{13}^3 = W_{23}^3 = 1 \) and zero otherwise. Note that \( W \) indeed satisfies the required symmetry property (72a) and further the \( W \)'s satisfy (72b) since \( W^{(1)} = \text{Id}^{3 \times 3} \),

\[
W^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad W^{(3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},
\]

(101)

which commute.

Equations in (98a) follow, in order, from \( \partial_t \tilde{\xi} = \{\tilde{\xi}, \mathcal{H}\}_{xy}, \partial_t \psi_s = \{\psi_s, \mathcal{H}\}_{xy}, \) and \( \partial_t \psi_{s_2} = \{\psi_{s_2}, \mathcal{H}\}_{xy} \),

(102)

b. Conservation laws. The integrand of the Hamiltonian (\( \mathcal{H} \)) of the IL\(^{(0,1)}\)QG (99) does not depend explicitly on \( t \). So \( \mathcal{H} \) is invariant under time shifts. Being the generator of infinitesimal such transformations, by Noether’s theorem \( \mathcal{H} \) (i.e., the energy) is preserved under dynamics of (98). In a zonally symmetric domain \( D \) of the \( \beta \)-plane, \( \mathcal{H} \) is invariant under \( x \)-translations. The corresponding generator \( (\mathcal{M}^x) \) must be conserved. As for the HL+QG and IL\(^{0+}\)QG models, this turns out to be given by (73) as shown next. Since \( \mathcal{M}^x \) must satisfy \( \{\psi_s, \mathcal{M}^x\}_{xy} = \partial_x \psi_s \) and \( \{\psi_{s_2}, \mathcal{M}^x\}_{xy} = \partial_x \psi_{s_2} \), it follows that \( \mathcal{M}^x = \int y \xi, sufficiently and necessarily by the same argument given above: any given two vectors \( a, b \) on \( D, \ a \times b = a_1 \hat{z} \) iff \( b = (1, 0, 0) \). This immediately gives \( \{\xi, \mathcal{M}^x\}_{xy} + \{\psi_s, \mathcal{M}^x\}_{xy} + \{\psi_{s_2}, \mathcal{M}^x\}_{xy} = \partial_x \xi \). On an \( f \) -plane the Hamiltonian is invariant under rotations in an axisymmetric domain. The generator \( \mathcal{M}^\phi \) of infinitesimal rotations is conserved by Noether’s theorem. This must satisfy \( \{\psi_s, \mathcal{M}^\phi\}_{r\phi} = r \partial_r \psi_s \) and \( \{\psi_{s_2}, \mathcal{M}^\phi\}_{r\phi} = r \partial_r \psi_{s_2} \), which holds if and only if \( \mathcal{M}^\phi \) is given by (74), just as for the HL+QG and IL\(^{0+}\)QG. This immediately makes \( \{\xi, \mathcal{M}^\phi\}_{r\phi} + \{\psi_s, \mathcal{M}^\phi\}_{r\phi} + \{\psi_{s_2}, \mathcal{M}^\phi\}_{r\phi} = r \partial_r \xi \). Finally, the Casimirs of (100) are given by

\[
\mathcal{C} = \int \xi + F(\psi_s, \psi_{s_2})
\]

(103)

where \( F \) is an arbitrary function. Indeed, \( \{\psi_s, \mathcal{C}\}_{xy} = 0 = \{\psi_{s_2}, \mathcal{C}\}_{xy} \) is satisfied iff \( \mathcal{C} = \text{const} \). On the other hand, \( \{\xi, \mathcal{C}\}_{xy} + \{\psi_s, \mathcal{C}\}_{xy} + \{\psi_{s_2}, \mathcal{C}\}_{xy} = 0 \) is satisfied iff \( \partial_x \mathcal{C} = \partial_y \mathcal{C} = \partial_x \mathcal{C} \). Thus (103) follows. This Casimir should be possible to be derived using the method developed by Thiffeault and Morrison.

D. The IL\(^{(0,\alpha)}\) family

1. The IL\(^{(0,\alpha)}\) PE

Finally, a general class of rotating shallow-water models with thermodynamics is the IL\(^{(0,\alpha)}\)PE itself with the identifications \( \rho_1 = \hat{h}, \rho_2 = \hat{\nu}, \rho_3 = \hat{T}, \rho_4 = \hat{\theta}, \hat{\rho}_3 = \hat{\theta}_s, \) etc., in (8), and the interpretation of the buoyancy field as the polynomial expansion (no claim on convergence is made or implied)

\[
\bar{\theta}(x, \sigma, t) = \hat{\theta}(x, t) + \sum_{\alpha} (\sigma^n - \overline{\sigma^n}) \bar{\theta}_\alpha(x, t),
\]

(104)

where

\[
\overline{\sigma^n} = \begin{cases}
0 & n : \text{odd}, \\
\frac{1}{n+1} & n : \text{even}.
\end{cases}
\]

(105)

The full (i.e., in the IL\(^\infty\)PE model) hydrostatic pressure field produced by (104),

\[
p = \frac{1}{2} \hat{h} \bar{\theta} + \frac{1}{2} \sum_{\alpha} (\sigma^{n+1} - (\sigma + 1)^{n+1}) \bar{\theta}_\alpha(n+1)_{\hat{h} \bar{\theta}_\alpha}.
\]

(106)
TABLE II. Casimir densities depending on submodel class (as defined by the potential energy density choice) and type (relative to the Rossby number (ε) size, finite or infinitesimally small). The functions $F$ and $G$ are arbitrary. The symbol ~ is used to mean proportional to, asymptotically as $\varepsilon \to 0$.

\[
\begin{array}{cccc}
\text{Class} & \text{Type} & \varphi_{\alpha} & \text{N.B.} \\
\hline
\text{HL} & 1/2 \varphi h & hF(\bar{q}) & q - f_0/H_t \sim \bar{q} \\
\text{IL}^0 & 1/2 \varphi h & h\bar{q}F(\bar{d}) + hG(\bar{d}) & \xi F(\psi_\sigma) + G(\psi_\sigma) \\
\text{IL}^{(0,1)} & 1/2 (\bar{d} - \frac{1}{2} \varphi h) & h\bar{q} + hF(\bar{d}, \varphi_\sigma) & \xi + F(\psi_\sigma, \psi_{\sigma+1}) \\
\text{IL}^{(0,\alpha)} & 1/2 (\bar{d} - \sum_{1}^{\sigma n+1} \varphi_{\sigma n}) h & h\bar{q} + hF(\bar{d}, \varphi_\sigma, \ldots, \varphi_{\sigma n}) & \xi + F(\psi_\sigma, \ldots, \psi_{\sigma n+1}) \\
\end{array}
\]

Noting that $\nabla p = \nabla (1/2 \varphi(1 - \sigma) \nabla h)$, one finds

\[
\nabla p = \frac{1}{2} h \left( \nabla \bar{d} - \sum_{1}^{\sigma n+1} \varphi_{\sigma n} \right) + \left( \bar{d} - \sum_{1}^{\sigma n+1} \varphi_{\sigma n} \right) \nabla h. \tag{107}
\]

Under the same considerations as for the derivation of the $\text{IL}^{(0,1)} \text{PE}$ (steps 1–4 in Sec. IV.C), the $\text{IL}^{(0,\alpha)} \text{PE}$ with the potential energy density choice

\[
\varphi_{\alpha} = \varphi_\alpha := \frac{1}{2} h \left( \bar{d} - \sum_{1}^{\sigma n+1} \varphi_{\sigma n} \right) \tag{108}
\]

gives the dynamical equations consistent with the buoyancy field representation (104). We can refer to the resulting model as the $\text{IL}^{(0,\alpha)} \text{PE}$. The above clarifies the superscript in $\text{IL}^{(0,\alpha)} \text{PE}$: no vertical variation is allowed for the horizontal velocity, while polynomial vertical variation is permitted for the buoyancy field up to an order degree. The model velocity has implicit vertical shear, with dependence in the vertical coordinate being polynomial up to a degree exceeding in one unit that of the buoyancy field. Finally, potential energy density choices more general than $\varphi_{\alpha}$, allowed by the $\text{IL}^{(0,\alpha)} \text{PE}$, can expand the realm of applicability of the model; also, an appropriate choice can enable the study of the response to forcing and/or dissipation in a Hamiltonian context.

2. The $\text{IL}^{(0,\alpha)} \text{QG}$

The $\text{IL}^{(0,\alpha)} \text{QG}$ is obtained from the $\text{IL}^{(0,\alpha)} \text{PE}$ by considering the expansions in (95), but with that for $h$ replaced by

\[
h/H_t = 1 + \bar{\psi} - \psi_\sigma + \sum_{1}^{n+1} \frac{(n + 1)}{f_0 R_\alpha^2} \psi_{\sigma n+1} + \ldots, \tag{109}
\]

where $R_\alpha^2 := (1 - \frac{1}{2} \sum_{1}^{\sigma n+1} \bar{\psi}_{\sigma n}) t$ and further proposing

\[
\varphi_{\sigma n} \sim \frac{4(n+1)}{f_0 R_\alpha^2} \psi_{\sigma n+1} + \ldots \quad \varepsilon \quad \varepsilon^2\tag{110}
\]

$n = 2, \ldots, \alpha$, consistent with a uniform reference density stratification (recall that the QG fields are $O(\varepsilon)$ perturbations on such a reference state). The $\text{IL}^{(0,\alpha)} \text{QG}$ is given by

\[
\begin{aligned}
\partial_t \bar{\xi} + \{ \bar{\psi}, \xi \}_{xy} - R_\alpha^{-2} \{ \bar{\psi}, \nu(\psi_\sigma, \ldots, \psi_{\sigma n+1}) \}_{xy} & = 0, \\
\partial_t \psi_{\sigma n} + \{ \psi, \psi_{\sigma n} \}_{xy} & = 0.
\end{aligned}\tag{111a}
\]

This system possesses a generalized Hamiltonian structure with Hamiltonian given by

\[
\mathcal{H}[\bar{\xi}, \psi_\sigma, \ldots, \psi_{\sigma n}] := \frac{1}{2} \int \frac{\nabla \bar{\psi}}{2} + R_\alpha^{-2} \bar{\psi}^2 \tag{112}
\]

and Lie–Poisson bracket (71) on $\mu = (\bar{\xi}, \psi_\sigma, \ldots, \psi_{\sigma n})$ with $W_1 = W_{1n} = 1, n = 1, \ldots, \alpha + 1$, and zero otherwise. Conservation of energy (112), zonal momentum (73), and angular momentum (74) are related by Noether’s theorem with symmetry of (111) under time shifts, $x$-translations (in a zonal domain), and rotations (in an axisymmetric domain of the $f$ plane), respectively. Finally, the corresponding infinite of Casimirs, which do not generate any explicit symmetry, is given by

\[
\mathcal{C} = \int \bar{\xi} + F(\psi_\sigma, \ldots, \psi_{\sigma n+1}), \tag{113}
\]

where $F$ is arbitrary. In Table II the Casimirs of all submodels discussed are compared.

V. STRATIFICATION EFFECTS

I had opened this paper showing in Fig. 1 a snapshot at $t f_0 = 14$ (for additional snapshots, cf. Fig. 3, right panel, multimedia view) of the buoyancy ($\bar{d}$) from
a simulation of the IL\(^0\)QG in a doubly periodic domain of the \(f\)-plane, and I close it by showing in Fig. 2 the same but based on a simulation of the IL\(^{(0,1)}\)QG, which has, in addition to lateral density inhomogeneity, linear vertical stratification. The initial conditions are the same, viz., zero QG potential vorticity (\(\xi\)) and linear \(\nabla\). As in Holm, Luesink, and Pan\(^{20}\), topographic forcing to the \(\xi\)-equation, of the form \(R^{-2}\{\psi_{\xi},\psi_0\}_{xy}\) (which does not spoil the Hamiltonian structure of the system) where \(\psi_0(x)\) is a prescribed small-amplitude sinusoid, was added to speedup the development of small-scale (compared to \(R\), the equivalent-barotropic Rossby radius of deformation) Kelvin–Helmholtz-like rollup filaments. Note that the filamentation is much less intense in the IL\(^{(0,1)}\)QG than in the IL\(^0\)QG. It has been argued\(^{20}\) that the lack of Kelvin–Helmholtz circulation conservation (more precisely, its creation and possibly ensuing inverse-energy cascade suppression) in the IL\(^0\) model plays a role in the formation of small-scale rollup filaments. The IL\(^{(0,1)}\) does not conserve circulation either, but the filamentation is much weaker, as noted. The inclusion of stratification seems responsible for halting its development, possibly by introducing a high-wavenumber instability cutoff, lacking in the IL\(^0\)\(^{5,7}\). This might be expected because the IL\(^{(0,1)}\) can represent baroclinic instability at long scales (i.e., of the order of \(R\)) as well as at short scales (i.e., of the order of \(sR\), when the stratification measure \(s\), defined in (59), is small).\(^{7,60}\) This demands a thorough investigation of the stability properties of the IL\(^{(0,1)}\) (and more generally the IL\(^{(0,\alpha)}\)), exploiting the symmetry-related conservation laws of the system(s), and its consequences, which is reserved for future research.

VI. SUMMARY, DISCUSSION, AND OUTLOOK

The renewed interest in thermodynamically-active rotating shallow-water modeling\(^{5,6,12,14,20,35,69–72}\) motivated this work, which presented extended rotating shallow-water theories with thermodynamics and geometry. With a focus on the ocean mixed-layer, the topmost part of the ocean, a general model was introduced and interpreted as such, considering in addition several submodels, ranging from the shallow-water equations themselves. Unlike the latter, all other models discussed have buoyancy varying arbitrarily in horizontal position and time, and possibly also in the vertical in polynomial form up to an arbitrary degree. Unlike the now classical “thermal” rotating shallow-water model, referred to as Ripa’s model by many authors due to Pedro Ripa’s contribution to its study,\(^{55–58,60}\) the “thermal” models presented here are capable of representing important mixed-layer ocean processes like restratification by baroclinic instability.\(^{15}\) All models discussed were shown to admit Euler–Poincare variational formulation\(^{21}\) and to possess generalized Hamiltonian structure.\(^{40}\) This is important because structure-preserving algorithms\(^{35}\) can be applied in simulations, preventing numerical artifacts. Moreover, flow-topology-preserving techniques to build parametrizations of typically unresolved scales can be applied to investigate the effects of these on transport at resolvable scales.\(^{8,18}\)

Allowing variable density, the acronym IL, standing for “inhomogeneous layer,” was used to denote the models by appending to it superscripts that indicate the amount of vertical variation allowed in the horizontal velocity and buoyancy fields. The most general model, denoted IL\(^{(0,\alpha)}\), allows buoyancy to vary polynomially up to degree \(\alpha\) while velocity is kept constant in the vertical. The + sign indicates that the model enables, through an appropriate choice of potential energy density, investigation of forced/dissipative nonisothermal mixed-layer dynamics in a conservation (Hamiltonian) setting. Otherwise the IL\(^{(0,\alpha)}\) has a potential energy density strictly consistent with the buoyancy vertical structure in the model. Quasigeostrophic (QG) versions of the corresponding primitive equations (PE) were derived and their geometric structure discussed in detail. In particular, the IL\(^{(0,0)}\)PE, or simply IL\(^0\)PE, which corresponds to Ripa’s model, was explicitly connected with earlier work by Morrison and Greene\(^{40}\): it is a special case of the hydrodynamics form of the model discussed by these authors.

For future research I reserve a thorough investigation of the stability of stratified steady-flow solutions of the model derived here. This should lead to insight into the reported tendency of stratification to halt the development of rollup filaments when buoyancy is kept uniform in the vertical. Also left for future research is including
vertical shear in the velocity, as devised in Ripa\textsuperscript{56} but in a geometry-preserving manner. This is hoped to be achieved by performing truncations in the amount of vertical variation allowed in the dynamical fields directly in the Lagrangian,\textsuperscript{19,62} or the Hamiltonian,\textsuperscript{45} of the IL\textsuperscript{∞}PE, i.e., exact, model. The multilayer form\textsuperscript{5} of the IL\textsuperscript{1} model of Ripa\textsuperscript{56} was shown to behave in the ageostrophic baroclinic instability problem quite accurately compared to the IL\textsuperscript{∞}PE model. Yet issues with the numerical implementation of the IL\textsuperscript{∞}PE, and thus with its practical use, were noted,\textsuperscript{11} possibly in connection with the seeming lack of geometric structure of the model.

**AUTHOR’S CONTRIBUTIONS**

This paper is authored by a single individual who entirely carried out the work.

**DATA AVAILABILITY**

This paper does not involve the use of data.

**ACKNOWLEDGMENTS**

The comments by an anonymous reviewer led to improvements to this paper. I thank Prof. Philip J. Morrison for calling my attention to Morrison and Greene\textsuperscript{40} and for lively discussions on Poisson brackets during the 2021 Aspen Center for Physics workshop “Transport and Mixing of Tracers in Geophysics and Astrophysics,” where this and additional ongoing work was initiated. Further stimulating discussions with Prof. Darryl D. Holm on “thermal fluids” are acknowledged. The Aspen Center for Physics is supported by NSF grant PHY-1607611.

**Appendix A: Proof for the Jacobi identity of (45)**

The commutator in (28) is antisymmetric (i.e., $[a, b] = -[b, a]$) for any vectors $a, b$ and satisfies the Jacobi identity (viz., $[[a, b], c] + \bigcirc = 0$ for all vectors $a, b, c$). The first property is obvious; the second one involves some algebra but is otherwise quite straightforward to verify:

$$[a, b, c] = \left( ((a \cdot \nabla) b) \cdot \nabla \right) c + \left( ((b \cdot \nabla) a) \cdot \nabla \right) c - \left( ((b \cdot \nabla) a) \cdot \nabla \right) c + \left( (b \cdot \nabla) (a \cdot \nabla) \right) c - \left( c \cdot \nabla \right) (a \cdot \nabla) b + \left( (a \cdot \nabla) (b \cdot \nabla) \right) c + (bc \cdot \nabla) a + (cb \cdot \nabla) a + (ac \cdot \nabla) b + (ca \cdot \nabla) b + (ba \cdot \nabla) c + (ab \cdot \nabla) c$$

(A1)

$$[[b, c], a] = \left( (b \cdot \nabla) (c \cdot \nabla) a \right) - \left( (c \cdot \nabla) (b \cdot \nabla) a \right) - \left( a \cdot \nabla \right) (b \cdot \nabla) c + \left( a \cdot \nabla \right) (c \cdot \nabla) b + \left( ab \cdot \nabla \right) c + \left( ac \cdot \nabla \right) b$$

(A2)

$$[[a, b], c] = \left( (a \cdot \nabla) (b \cdot \nabla) c \right) - \left( (b \cdot \nabla) (a \cdot \nabla) c \right) - \left( c \cdot \nabla \right) (a \cdot \nabla) b + \left( (a \cdot \nabla) (b \cdot \nabla) c \right) + \left( (b \cdot \nabla) (a \cdot \nabla) c \right)$$

(A3)

where $ab : cd = d^j c^j d_j$; adding (A1)–(A3) with $ab = ba$ in mind proves the Jacobi identity for $[,]$. Now,

$$\{\mathcal{U}, \mathcal{V}\}^m = -\{\mathcal{V}, \mathcal{U}\}^m, \quad \{\mathcal{U}, \mathcal{V}\}^{\rho_{1\alpha}} = -\{\mathcal{V}, \mathcal{U}\}^{\rho_{1\alpha}},$$

(A4)

manifestly for any functionals $\mathcal{U}, \mathcal{V}$. Our goal is to demonstrate that

$$\{\mathcal{U}, \mathcal{V}\} := \{\mathcal{U}, \mathcal{V}\}^m + \sum \{\mathcal{U}, \mathcal{V}\}^{\rho_{1\alpha}}$$

(A5)

satisfies $\{\{\mathcal{U}, \mathcal{V}\}, \mathcal{W}\} + \bigcirc = 0$ for all functionals $\mathcal{U}, \mathcal{V}, \mathcal{W}$. More precisely, we seek to show that

$$\{\{\mathcal{U}, \mathcal{V}\}, \mathcal{W}\} = \{\{\mathcal{U}, \mathcal{V}\}, \mathcal{W}\}^m + \sum \{\{\mathcal{U}, \mathcal{V}\}, \mathcal{W}\}^{\rho_{1\alpha}}$$

(A6)

vanishes upon $\bigcirc$. To do it, we consider the $\mathbf{m}$ bracket and the sum of $\rho_{1\alpha}$ brackets in (A6) separately, with the following in mind:

$$\{\mathcal{U}, \mathcal{V}\}_m = -[\mathcal{U}_m, \mathcal{V}_m], \quad \{\mathcal{U}, \mathcal{V}\}^{\rho_{1\alpha}} = -[\mathcal{U}_m, \nabla \mathcal{V}^{\rho_{1\alpha}} - \mathcal{V}_m, \nabla \mathcal{U}^{\rho_{1\alpha}}],$$

(A7)

(A8)

where terms involving second-order functional derivatives of $\mathcal{U}$ and $\mathcal{V}$ have been omitted. This can be done with no harm because the (manifestly) skew-adjointness of the Poisson operator $\mathcal{J}^\dagger$, defined in (44), accounts\textsuperscript{36} for the lack of contribution of these terms to $\{\{\mathcal{U}, \mathcal{V}\}, \mathcal{W}\} + \bigcirc$. For completeness, I give a quick proof here. First note that, for $\mathcal{U}[\mu] = \int U(x; \mu, \nabla \mu, \ldots)$, the second variational derivative is defined as the unique element $\delta^2 \mathcal{U} / \delta \mu^2$.
FIG. 3. Evolution of \( \bar{\vartheta} \) (vertically averaged buoyancy) in the IL\(^0\)QG (left) and IL\(^{(0,1)}\)QG (right) models. The initial condition has uniform QG potential vorticity (\( \bar{\xi} \)) and linear \( \bar{\vartheta} \). The domain of integration is doubly periodic, and lies on the \( f \)-plane. Length is scaled by the external Rossby radius of deformation. Time is scaled by the mean Coriolis parameter. Topographic forcing to the \( \xi \)-equation, of the form \( \{ \bar{\vartheta}, \psi \}_x \), where \( \psi_0(\mathbf{x}) \) is a prescribed small-amplitude sinusoid, is included. A pseudospectral code with a small amount of hyperviscosity and a fourth-order Runge–Kutta time stepper is used. The grid resolution is 512 \( \times \) 512.

(Multimedia view.)

satisfying

\[
\frac{d^2}{dz^2} \mathcal{W}[\mu + \varepsilon \delta \mu] = \int \delta \mu \cdot \frac{\delta^2 \mathcal{W}}{\delta \mu^2} \delta \mu. \tag{A9}
\]

Since \( \mathcal{W}_\mu = \partial_\mu U \), it is clear then that \( \mathcal{W}_\mu := \frac{\delta^2 \mathcal{W}}{\delta \mu^2} = \partial_\mu \mathcal{W}_\mu \). Skew-adjointness of \( \mathcal{J} \) means \( \mathcal{J} \mathcal{W}_\mu \cdot \mathcal{J} \mathcal{W}_\nu = - \mathcal{J} \mathcal{W}_\nu \cdot \mathcal{J} \mathcal{W}_\mu \), where \( \mu = (\mathbf{m}, \rho) \). The ignored terms in \( \{ \mathcal{W}, \mathcal{V} \}_\mu \) are \( \mathcal{W}_\mu \cdot \mathcal{J} \mathcal{V}_\mu - \mathcal{V}_\mu \cdot \mathcal{J} \mathcal{W}_\mu \), where skew-adjointness of \( \mathcal{J} \) was used and \( \mathcal{W}_\mu \cdot \mathcal{J} \mathcal{V}_\mu = (\mathcal{J} \mathcal{V}_\mu)^* \mathcal{W}_\mu^* \). Then we have, including these terms only, \( \{ \mathcal{W}, \mathcal{V} \}_\mu = (\mathcal{W}_\mu \cdot \mathcal{J} \mathcal{V}_\mu) \cdot \mathcal{J} \mathcal{W}_\mu \). To see that \( \{ \mathcal{W}, \mathcal{V} \}_\mu + \circ = 0 \) (when the terms in question are included only), it is enough to realize that \( (\mathcal{W}_\mu \cdot \mathcal{J} \mathcal{V}_\mu) \cdot \mathcal{J} \mathcal{W}_\mu \equiv (\mathcal{W}_\mu + \mathcal{W}_\mu^*) \cdot \mathcal{J} \mathcal{W}_\mu \), noting that \( \mathcal{W}_\mu + \mathcal{W}_\mu^* = \mathcal{W}_\mu^* \).

Let us start with the \( \mathbf{m} \) bracket, which, using (A7), reads

\[
\{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}_\mathbf{m} = - \int \mathbf{m} \cdot [\{ \mathcal{W}, \mathcal{V} \}_\mathbf{m}, \mathcal{W}_\mathbf{m}] \\
+ \int \mathbf{m} \cdot [\mathcal{W}_\mathbf{m}, [\mathcal{W}_\mathbf{m}, \mathcal{W}_\mathbf{m}]] \tag{A10}
\]

Since \([,] \) satisfies the Jacobi identity, we readily find

\[
\{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}_\mathbf{m} + \circ = 0. \tag{A11}
\]

We now turn to the sum of \( \rho_{\alpha} \) brackets in (A6), which require more elaboration. It is enough to consider one term only, though. More precisely, \( \{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}^\rho_{\alpha} \)

\[
= - \int \rho_{\alpha} (\{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}} - \mathcal{W}_\mathbf{m} \cdot \nabla \{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}_{\rho_{\alpha}}) \\
= + \int \rho_{\alpha} (\{ \{ \mathcal{W}, \mathcal{V} \}, \mathcal{W} \}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}} \\
- \mathcal{W}_\mathbf{m} \cdot \nabla (\mathcal{W}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}} - \mathcal{W}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}})) \\
= + \int \rho_{\alpha} (\{ \mathcal{W}, \mathcal{V} \}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}} \\
- (\mathcal{W}_\mathbf{m} \cdot \nabla \mathcal{W}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}} - \mathcal{W}_\mathbf{m} \cdot \nabla \mathcal{W}_{\rho_{\alpha}}) \cdot \nabla \mathcal{W}_{\rho_{\alpha}}) \tag{A12}
\]

where, in order, we took into account (A7)–(A8) and

\[
(\mathbf{a} \cdot \nabla)\mathbf{b} \cdot \mathbf{c} = ((\mathbf{a} \cdot \nabla)\mathbf{b}) \cdot \mathbf{c} + \mathbf{a} \cdot \mathbf{b} : \nabla \mathbf{c} \tag{A13}
\]

(recalling that \( \mathbf{ab} = \mathbf{ba} \)). More explicitly, we have:
\begin{align}
\{\{\mathcal{W}, \mathcal{W}\}, \mathcal{W}\}^\rho_{na} &= \int \rho_{(n_a)} \\
&= \left(\left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)} - \left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)}\right) \\
&\quad - \left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)} + \left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)}\right). \\
&= \left(\left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)} - \left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)}\right)
\end{align}
(A14)

Similarly, \(\{\{\mathcal{W}, \mathcal{W}\}, \mathcal{W}\}^\rho_{na} = \int \rho_{(n_a)}\)

\begin{align}
&= \left(\left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)} - \left((\mathcal{W}_m \cdot \nabla)\mathcal{W}_m\right) \cdot \nabla \rho_{(n_a)}\right)
\end{align}
(A15)

and \(\{\{\mathcal{W}, \mathcal{W}\}, \mathcal{W}\}^\rho_{na} + \circ = 0. \tag{A17}\)

Thus, (A11) and (A17) together produce

\begin{align}
\{\{\mathcal{W}, \mathcal{W}\}, \mathcal{W}\} + \circ = 0, \tag{A18}\end{align}

and the proof is completed. \(\square\)

Appendix B: Clarification of the constraints (26)-(27)

The constraints (26)-(27) allow one to “skip one step” and obtain the equations of motion directly in Eulerian (spatial) variables, i.e., without having to transform back to these variables after extremizing an action under variations of particle paths at fixed Lagrangian (material) labels and time. Indeed, the constraints represent Eulerian variable variations induced by such path variations. Explicitly, \(\eta(x, t) := \delta q_t \circ q_{-t}(x)\), where \(q_t:l \mapsto x\) is the path in \(D \subseteq \mathbb{R}^2\) of the fluid particle marked with label \(l\), taken as position in \(D\) at \(t = 0\), which is viewed as the reference configuration of the fluid. The Lagrangian-to-Eulerian coordinates map \(q_t:l \mapsto x\) is assumed to be smooth as well as its inverse \(q_t^{-1} = q_{-t}\). The relationship between \(q_t\) and \(F_t\), i.e., the \((t_0, t)-flow\) map of \(\tilde{u}(x, t)\), obtained by solving \(\dot{x} = \tilde{u}(x, t)\) with initial condition \(x_0 = x(t_0)\), is \(F_t = q_t \circ q_{-t}\).

Noting that \(\tilde{u}(x, t) = \partial_t q_t \circ q_{-t}(x) = \dot{x}\), upon taking its variation \(\delta q_t \circ q_{-t}(x) = \delta x\) and the time derivative of \(\eta(\partial_t \eta + (\tilde{u} \cdot \nabla)\eta) = \delta \tilde{x}\), constraint (26) follows.

Constraint (27) follows upon noting first that (2) is equivalent to \(\rho_{na}(x, t) = \rho_{na}\int_0^t \alpha_\eta \circ q_{-t}(x)\), where \(\rho_{na}(l)\) is the amount of “density mass” carried by fluid particle \(l\), which is preserved in time, and \(J_t(l) := \partial_l (q_t(l)/\partial l)\) is the Jacobian of the map \(q_t\). The variation of \(\rho_{na}\), \(\delta l_t \rho_{na} + \eta \cdot \nabla \rho_{na} = -\rho_{na} \cdot \eta\), by invertibility of \(q_t\), which leads to (27).

Appendix C: Base Lie algebras, extensions, and realization envelopings

We recall a few abstract results that are needed to make formal statements on the geometry of results presented in this paper. The set \(\{q_t\}\) forms a one-parameter group under composition. As \(q_t\) is a diffeomorphism, i.e., it is smoothly invertible, \(\{q_t\}\) has the structure of a differentiable manifold, and, hence, \((\{q_t\}, \circ)\) represents a Lie group, denoted \(\text{Diff}(D)\), which provides a representation for the fluid configuration space: knowing \(q_t\) tells one where a fluid particle goes.

The velocity \(\tilde{u}(x, t)\) is an element of the tangent space to \(\text{Diff}(D)\) at \(x\), \(T_x \text{Diff}(D)\). Since \(x = q_k(l)\), the tangent space to \(\text{Diff}(D)\) at the identity \(q_0(l) = 1\), namely, \(T_1 \text{Diff}(D)\), uniquely determines all other tangent spaces to \(\text{Diff}(D)\). The tangent space has the structure of a vector space. The elements of \(T_1 \text{Diff}(D)\) are invariant under composition of \(q_t\) by the particle re-labelling \(r(l)\) on the right. Indeed, under \(l \mapsto r(l)\) it follows that \(\tilde{u} \mapsto \partial_t q_t \circ r^{-1} \circ q_t \circ r^{-1} = \partial_t q_t \circ q_{-t} \circ (r \circ r^{-1}) \equiv \tilde{u}\). The vector space of right-invariant vectors of a Lie group forms a Lie algebra. For \(\text{Diff}(D)\), this is denoted \(\mathfrak{X}\), which is isomorphic to \(T_1 \text{Diff}(D)\).

In addition to its vector space structure, \(\mathfrak{X}\) is equipped with a bilinear, right-invariant product \([\cdot, \cdot]: \mathfrak{X} \times \mathfrak{X} \rightarrow \mathfrak{X}\), known as the Lie bracket, which satisfies \([a, b]_X = -[b, a]_X\) (antisymmetry) and \([[a, b]]_X + \circ = 0\) (Jacobi identity) for all \(a, b, c \in \mathfrak{X}\). The bracket is given by minus the commutator of vector fields (28), namely, the Lie derivative of a vector field in \(\mathfrak{X}\) along the flow of another one in \(\mathfrak{X}\), which expresses how the elements of \(\mathfrak{X}\) act on themselves.

When \(J_t(l) = 1\), i.e., \(q_t\) is an area preserving diffeomorphism, then \(\tilde{u}\) is divergence free. Namely, it is determined by a streamfunction \(\psi(x, t)\), i.e., \(\tilde{u} = \tilde{z} \times \nabla \psi\), which \(\hat{\psi}\) is viewed as an element of the vector space \(\mathcal{F}\) of smooth \((C^\infty)\) time-dependent functions (in \(D\)). The Lie bracket of the corresponding Lie algebra, \(\mathfrak{X}_{\text{area}} \cong T_1 \text{Diff}_{\text{area}}(D)\), is \([[\tilde{u}_1, \tilde{u}_2]]_{\text{area}} = -[[\tilde{u}_1, \tilde{u}_2]] = -[\tilde{z} \times \nabla \psi_{\tilde{1}}, \tilde{z} \times \nabla \psi_{\tilde{2}}] = \tilde{z} \times \nabla \{\psi_{\tilde{1}}, \psi_{\tilde{2}}\}_{xy}\), where \(\{,\}_{xy}\) is the canonical Poisson
bracket \( (34) \). This bracket is manifestly antisymmetric and satisfies the Jacobi identity. Furthermore, it is a derivation in each of its arguments; thus it satisfies \( \{ U, V \}, W \}_x = U \{ V, W \}_x + \{ U, W \}_x V \) (Leibniz rule).

The vector space \( \mathcal{F} \) together with the canonical bracket forms a Lie enveloping algebra,\(^{38}\) which we denote \( \tilde{\mathfrak{g}} \).

1. PE systems

The proof in App. A expands Example 5.B of Marsden, Ratiu, and Weinstein\(^{30}\) as follows. Let \( \mathcal{F}^n \) represent the \( n \)-dimensional vector space given by \( n \) copies of \( \mathcal{F} \). The vector space \( \mathcal{X} \times \mathcal{F}^{n+2} \) with the bracket given by \( [\cdot, \cdot] \) on \( \mathcal{X} \times \mathcal{F}^{n+2} \) with the bracket given by \( \{ \cdot, \cdot \} \) on \( \mathcal{X} \times \mathcal{F}^{n+2} \)

\[ \text{Remark 9} \quad \text{While the Lie-Poisson system \( (41) \) is defined on} \quad \mathcal{X} \times \mathcal{F}^{n+2}, \quad \text{its Euler-Poincare counterpart,} \quad (24) \quad \text{with the second equation of \( (8) \), is defined on} \quad \mathcal{X} \times \mathcal{F}^{n+2}. \quad \text{The connection between the two, equivalent formulations is provided by the partial Legendre transformation}^{21} \quad (\hat{u}, \rho) \quad \text{by \( (37) \), viz.,} \quad \mathcal{H} \{ \hat{u}, \rho \} = \int \hat{u} \cdot \hat{u} - \mathcal{L}(\hat{u}, \rho). \]

2. QG systems

The bracket \( [\cdot, \cdot]_{\tilde{\mathfrak{g}}} \) of \( \tilde{\mathfrak{g}} \) is a Poisson manifold and is a product for a realization of a QG system. Namely, \( \{ \cdot, \cdot \} \) on \( \mathcal{X} \times \mathcal{F}^{n+2} \)

\[ \text{Remark 10} \quad \text{For} \quad \mathcal{H} \{ \mu \} \quad \text{given, the motion equation} \quad \partial_t \mu = \{ \mu, \mathcal{H} \} \quad \text{does not seem possible to be obtained through Legendre transformation as is} \quad \partial_t \mu = \{ \mu, \mathcal{H} \}. \quad \text{Thus an Euler-Poincare variational formulation may not exist or is quite difficult to be derived for QG-type systems. Instead, ad-hoc variational formulations have been proposed in the literature.}^{17,42,68} \]

With the product \( \{ U, V \}, \tilde{\mathfrak{g}} \) \( = \{ U_1, V_1 \}, \tilde{\mathfrak{g}}_1 \) \( = \{ U_1, V_1 \}, \mathfrak{g}_1 \), \( \tilde{\mathfrak{g}} \) \( \text{is a Poisson manifold} \quad \text{and is a product for a realization of a QG system considered here represent particular cases.} \]

**AUTHOR DECLARATIONS**

**Conflict of interest**

The author has no conflicts to disclose.

1. Arnold, V., “Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l’hydrodynamique des fluides parfaits,” Annales de l’Institut Fourier 16, 319–361 (1966).
2. Beier, E., “A numerical investigation of the annual variability in the Gulf of California,” J. Phys. Oceanogr. 27, 615–632 (1997).
3. Benjamin, T., “Impulse, flow force and variational principles,” IMA J. Appl. Math. 32, 3–68 (1984).
4. Beron-Vera, F. J., “Constrained-Hamiltonian shallow-water dynamics on the sphere,” in Nonlinear Processes in Geophysical Fluid Dynamics: A Tribute to the Scientific Work of Pedro Ripa, edited by O. U. Velasco-Fuentes, J. Sheinbaum, and J. Ochoa (Kluwer, 2003) pp. 29–51.
5. Beron-Vera, F. J., “Multilayer shallow-water model with stratification and shear,” Rev. Mex. Fis. 67, 351–364 (2021).
6. Beron-Vera, F. J., “Nonlinear saturation of thermal instabilities,” Phys. Fluids 33, 036608 (2021).
7. Beron-Vera, F. J. and Ripa, P., “Free boundary effects on baroclinic instability,” J. Fluid Mech. 352, 245–264 (1997).
8. Cotter, C., Crisan, D., Holm, D., Pan, W., and Shevchenko, I., “Data assimilation for a quasi-geostrophic model with circulation-preserving stochastic transport noise,” J. Stat. Phys. 179, 1186 – 1221 (2020).
9. Dellar, P. J., “Common Hamiltonian structure of the shallow water equations with horizontal temperature gradients and magnetic fields,” Phys. Fluids 15, 292–297 (2003).
10. Dronkers, J., “Tidal computations in rivers, coastal areas and seas,” J. Hydraul. Div. 95, 44–77 (1969).
11. Eldevik, T., “On front dynamics in two model oceans,” J. Phys. Oceanogr. 32, 2,915–2,925 (2002).
12. Eldred, C., Dubos, T., and Kritsikis, E., “A quasi-hamiltonian discretization of the thermal shallow water equations,” Journal of Computational Physics 379, 1–31 (2019).
13. Gawlik, E. S.and Gay-Balmaz, F., “A structure-preserving finite element Mmethod for compressible ideal and resistive MHD,” Journal of Plasma Physics , in press (2021).
14. Gouzien, E., Lahaye, N., Zeitlin, V., and Dubos, T., “Thermal instability in rotating shallow water with horizontal temperature/density gradients,” Physics of Fluids 29, 101702 (2017).
15. Haine, T. W.and Marshall, J., “Gravitational, symmetric and baroclinic instability of the ocean mixed layer,” J. Phys. Oceanogr. 28, 634–658 (1998).
Holm, D., Marsden, J., Ratiu, T., and Weinstein, A., “Nonlinear stability conditions and a priori estimates for barotropic hydrodynamics,” Phys. Lett. A 98, 15–21 (1983).

Holm, D. and Zeitlin, V., “Hamilton’s principle for quasi-geostrophic motion,” Phys. Fluids 10, 800–806 (1998).

Holm, D. D., “Variational principles for stochastic fluid dynamics,” Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 471, 20140963 (2015).

Holm, D. D. and Luesink, E., “Stochastic wave–current interaction in thermal shallow water dynamics,” J. Nonlinear Sci. 31, 29 (2021).

Holm, D. D., Luesink, E., and Pan, W., “Stochastic mesoscale circulation dynamics in the thermal ocean,” Phys. Fluids 33, 046603 (2020).

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

Holm, D. D., Marsden, J. E., and Ratiu, T. S., “The Euler-Poincaré equations in geophysical fluid dynamics,” in Large-Scale Atmosphere-Ocean Dynamics II: Geometric Methods and Models, edited by J. Norbury and I. Roulstone (Cambridge University Press, 2002) Chap. 7, pp. 251–299.

Kraus, M., Tassi, E., and Grasso, D., “Variational integrators for reduced magnetohydrodynamics,” Journal of Computational Physics 321, 435–458 (2016).

Lahaye, N., Zeitlin, V., and Dubos, T., “Coherent dipoles in a mixed layer with variable buoyancy: Theory compared to observations,” Ocean Modelling 153, 101673 (2020).

Laviole, R., “A mesoscale numerical model of lake-effect storms,” J. Atmos. Sci. 29, 1025 – 1040 (1972).

Lewis, D., Marsden, J., and Montgomery, R., “The Hamiltonian structure for dynamic free boundary problem,” Physica D 18, 391–404 (1986).

Marle, C.-M., “On Henri Poincaré’s Note ‘Sur une forme nouvelle des équations de la Mécanique’,” Journal of Geometry and Symmetry in Physics 29, 1–38 (2013).

Marsden, J. and Weinstein, A., “Coadjoint orbits, vortices and clebsch variables for incompressible flows,” Physica D 7, 305–323 (1983).

Marsden, J. E. and Ratiu, T., Introduction to Mechanics and Symmetry, 2nd ed., Texts in Applied Mathematics, Vol. 17 (Springer, 1999).

Marsden, J. E., Ratiu, T., and Weinstein, A., “Semidirect products and reduction in mechanics,” Transactions of the American Mathematical Society 281, 147–177 (1984).

Marsden, J. E. and Weinstein, A., “The Hamiltonian structure of the Maxwell-Vlasov equations,” Physica D 4, 349–406 (1982).

McCreary, J. P., Kundu, P., and Molinari, R., “A numerical investigation of dynamics, thermodynamics and mixed-layer processes in the Indian Ocean,” Prog. Oceanogr. 31, 181–244 (1993).

McIntyre, M. E. and Shepherd, T. G., “An exact local conservation theorem for finite-amplitude disturbances to non-parallel shear flows, with remarks on Hamiltonian structure and on Arnold’s stability theorems,” J. Fluid Mech. 181, 527–565 (1987).

McWilliams, J. C., “Submesoscale currents in the ocean,” Proc R Soc A 472, 20160117 (2016).

Möller, H., Hidalgo, J., Martínez-Lopez, B., and Ruíz-Angulo, A., “Influence of stratification and Yucatan Current transport on the Loop Current Eddy shedding process,” Journal of Geophysical Research: Oceans 126, e2020JC016315 (2021).

Morrison, P. J., “Poisson brackets for fluids and plasmas,” in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, edited by M. Tabor and Y. Tres (Institute of Physics Conference Proceedings 88, 1982) pp. 13–46.

Morrison, P. J., “Hamiltonian description of the ideal fluid,” Rev. Mod. Phys. 70, 467–521 (1998).

Morrison, P. J., “Hamiltonian fluid dynamics,” in Encyclopedia of Mathematical Physics, edited by J.-P. Francoise, G. L. Naber, and T. S. Tsun (Academic Press, Oxford, 2006) pp. 593–600.

Morrison, P. J., “Structure and structure-preserving algorithms for plasma physics,” Physics of Plasmas 24, 055502 (2017).

Morrison, P. J. and Greene, J. M., “Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics,” Phys. Rev. Lett. 45, 790–794 (1980).

Morrison, P. J. and Hazeltine, R. D., “Hamiltonian formulation of reduced magnetohydrodynamics,” Phys. Fluids 27, 886–897 (1984).

Morrison, P. J., Lingam, M., and Acevedo, R., “Hamiltonian and action formalisms for two-dimensional gyroviscous magnetohydrodynamics,” Physics of Plasmas 21, 082102 (2014).

Mungkasi, S. and Roberts, S. G., “A smoothness indicator for numerical solutions to the Ripa model,” J. Phys.: Conf. Ser. 693, 012011 (2016).

Newcomb, W. A., “Lagrangian and Hamiltonian methods in magnetohydrodynamics,” Nuclear Fusion: Supplement Part 2 (1962).

Nore, C. and Shepherd, T. G., “A Hamiltonian weak-wave model for shallow-water flow,” Proc. R. Soc. Lond. A 453, 563–580 (1997).

This object can be interpreted as a Lie derivative upon appropriate interpretation of a and b, which I intentionally omit as the algebraic approach taken serves my purposes. Appendices B and C do discuss some differential geometry aspects which are needed to provide a deeper geometric interpretation of some of the results of the paper, but these can be safely ignored by the reader if not interested.

The Casimirs are related to the particle relabelling symmetry of fluid dynamics, which permits one to switch between the Lagrangian and Eulerian descriptions.

Dellar’s ”writes the Lie-Poisson bracket for the ILPME, viz., (45) for a = 0, referring to ? , where a planar compressible MHD system is discussed in relation with that of Morrison and Greene.

O’Brien, J. J. and Reid, R. O., “The non-linear response of a two-layer, baroclinic ocean to a stationary, axially-symmetric hurricane: Part I: Upwelling induced by momentum transfer,” J. Atmos. Sci. 24, 197–207 (1967).

Pedlosky, J., Geophysical Fluid Dynamics, 2nd ed. (Springer, 1987) p. 624 pp.

Rehman, A., Ali, I., and Qamar, S., “Exact riemann solutions of the ripa model for flat and non-flat bottom topographies,” Results in Physics 8, 104 – 113 (2018).

Ripa, P., “General stability conditions for a multi-layer model,” J. Fluid Mech. 222, 119–137 (1991).

Ripa, P., “Sistemas Hamiltonianos singulares. I: Planteamiento del caso discreto, Teorema de Noether,” Rev. Mex. Fís. 38, 984–1004 (1992).

Ripa, P., “Arnold’s second stability theorem for the equivalent barotropic model,” J. Fluid Mech. 257, 597–605 (1993).

Ripa, P., “Conservation laws for primitive equations models with inhomogeneous layers,” Geophys. Astrophys. Fluid Dyn. 70, 85–111 (1993).

Ripa, P., “On improving a one-layer ocean model with thermodynamics,” J. Fluid Mech. 305, 169–201 (1995).

Ripa, P., “Linear waves in a one-layer ocean model with thermodynamics,” J. Geophys. Res. C 101, 1233–1245 (1996).

Ripa, P., “Low frequency approximation of a vertically integrated ocean model with thermodynamics,” Rev. Mex. Fís. 42, 117–135 (1996).

Ripa, P., “‘Inertial’ oscillations and the β-plane approximation(s),” J. Phys. Oceanogr. 27, 633–647 (1997).

Ripa, P., “On the validity of layered models of ocean dynamics and thermodynamics with reduced vertical resolution,” Dyn. Atmos. Oceans 29, 1–40 (1999).

Ripa, P., “Baroclinic instability in a reduced gravity, three-dimensional, quasi-geostrophic model,” J. Fluid Mech. 403, 1–22 (2000).

Salmon, R., “Practical use of Hamilton’s principle,” J. Fluid Mech. 132, 431–444 (1983).
Sanchez-Linares, C., de Luna, T. M., and Castro Diaz, M. J., "A HLLC scheme for Ripa model," Applied Mathematics and Computation 272, 369–384 (2016).

Schopf, P. and Cane, M., “On equatorial dynamics, mixed layer physics and sea surface temperature,” J. Phys. Oceanogr. 13, 917–935 (1983).

Shepherd, T., “Rigorous bounds on the nonlinear saturation of instabilities to parallel shear flows,” J. Fluid Mech. 196, 291–322 (1988).

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

Thiffeault, J.-L. and Morrison, P. J., “Classification and Casimir invariants of Lie–Poisson brackets,” Physica D 136, 205–244 (2000).

Virasoro, M. A., “Variational principle for two-dimensional incompressible hydrodynamics and quasigeostrophic flows,” Phys. Rev. Lett. 47, 1,181–1,183 (1981).

Warneford, E. S. and Dellar, P. J., “The quasi-geostrophic theory of the thermal shallow water equations,” J. Fluid Mech. 723, 374–403 (2013).

Warneford, E. S. and Dellar, P. J., “Thermal shallow water models of geostrophic turbulence in Jovian atmospheres,” Physics of Fluids 26, 016603 (2014).

Warneford, E. S. and Dellar, P. J., “Super- and sub-rotating equatorial jets in shallow water models of Jovian atmospheres: Newtonian cooling versus Rayleigh friction,” Journal of Fluid Mechanics 822, 484–511 (2017).

Zeitlin, V., Geophysical Fluid Dynamics: Understanding (Almost) Everything with Rotating Shallow Water Models (Oxford University Press, 2018).