J. Fluid Mech. (2010), *vol.* 651, *pp.* 387–413.
doi:10.1017/S0022112009993922

Multilayer shallow water equations with complete Coriolis force. Part 1. Derivation on a non-traditional beta-plane

ANDREW L. STEWART AND PAUL J. DELLAR†

OCIAM, Mathematical Institute, 24–29 St Giles', Oxford, OX1 3LB, UK

(Received 22 May 2009; revised 9 December 2009; accepted 9 December 2009; first published online 24 March 2010)

We derive equations to describe the flow of multiple superposed layers of inviscid, incompressible fluids with constant densities over prescribed topography in a rotating frame. Motivated by geophysical applications, these equations incorporate the complete Coriolis force. We do not make the widely used 'traditional approximation' that omits the contribution to the Coriolis force from the locally horizontal part of the rotation vector. Our derivation is performed by averaging the governing Euler equations over each layer, and from two different forms of Hamilton's variational principle that differ in their treatment of the coupling between layers. The coupling may be included implicitly through the map from Lagrangian particle labels to particle coordinates, or explicitly by adding terms representing the work done on each layer by the pressure exerted by the layers above. The latter approach requires additional terms in the Lagrangian, but extends more easily to many layers. We show that our equations obey the expected conservation laws for energy, momentum and potential vorticity. The conserved momentum and potential vorticity are modified by non-traditional effects. The vertical component of the rotation vector that appears in the potential vorticity for each layer under the traditional approximation is replaced by the component perpendicular to the layer's midsurface. The momentum includes an additional contribution that reflects changes in angular momentum caused by changes in a fluid element's distance from the rotation axis as it is displaced vertically. Again, this effect is absent in the traditional approximation.

1. Introduction

Geophysical fluid dynamics, the study of the motion of the Earth's atmosphere and oceans, is concerned with the behaviour of rotating, stratified fluids, over wide ranges of length and time scales, and often in complex geometries. Simplified and approximate models therefore play a very important role in providing insight into processes that occur in the full equations. Shallow water equations are widely used as conceptual models because they capture the interaction between rotation and stratification and between waves and vortices evolving on disparate time scales. The simplest shallow water equations describe the motion of a single layer of fluid with a free surface. They may be derived by averaging the three-dimensional equations of motion across the layer, under the assumption that the layer's depth is small compared with its horizontal dimensions. Many more phenomena may be described

† Email address for correspondence: dellar@maths.ox.ac.uk

by shallow water models with two or more distinct layers of different densities. These models capture some of the baroclinic effects that arise in a stratified fluid when the pressure gradient is not parallel to the density gradient. For instance, two-layer shallow water models describe the troposphere and the stratosphere (e.g. Vallis 2006), the upper mixed layer and the lower ocean (e.g. Salmon 1982*b*), or a deep ocean current flowing under relatively quiescent fluid (e.g. Nof & Olson 1993).

This paper is primarily concerned with the Coriolis force due to the Earth's rotation and its approximation in idealized models. The angular velocity vector $\boldsymbol{\Omega}$ is directed parallel to the Earth's axis, so at a typical point on the Earth's surface $\boldsymbol{\Omega}$ has components in both the locally vertical and locally horizontal directions. The exceptions are the poles, where $\boldsymbol{\Omega}$ is purely vertical, and the equator, where $\boldsymbol{\Omega}$ is purely horizontal. However, the contribution to the Coriolis force due to the locally horizontal component of $\boldsymbol{\Omega}$ is widely neglected. This approximation was named the traditional approximation by Eckart (1960*a*), on the grounds that it was widely used, but seemed to lack theoretical justification. Phillips (1966) later showed that the traditional approximation could be justified as a consequence of a shallow layer approximation, one in which vertical length scales were small compared with horizontal length scales.

However, interest has recently grown in the effects of the Coriolis terms that are neglected under the traditional approximation. This interest is driven by improvements in numerical simulations, which now reach shorter horizontal length scales for which the shallow layer approximation becomes questionable. A recent review by Gerkema *et al.* (2008) explored the sparse material that is available on this topic. The effect of including the 'non-traditional' components of the Coriolis force is sometimes quite pronounced, particularly in mesoscale flows, such as Ekman spirals and deep convection (Leibovich & Lele 1985; Marshall & Schott 1999), and in internal waves (Gerkema & Shrira 2005*a*,*b*). This is consistent with the findings of the UK Meteorological Office, which in 1992 abandoned the traditional approximation in their unified model for the atmosphere (Cullen 1993). One might expect non-traditional effects to be even more pronounced in the oceans. The oceans contain substantial wave activity at or near inertial frequencies (Munk & Phillips 1968; Fu 1981), and regions of very weak stratification where the Brunt–Väisälä or buoyancy frequency N is less than ten times the inertial frequency (Munk 1981). van Haren & Millot (2005) report observations of 'gyroscopic' waves in areas of the Mediterranean with little or no stratification ($N = 0 \pm 0.4f$) to within the uncertainty of their measurements. These gyroscopic waves cannot be explained without invoking non-traditional effects.

In this paper, we derive multilayer shallow water equations that include the complete Coriolis force, in contrast to the conventional shallow water equations that rely upon the traditional approximation in their derivation. We thus extend the derivation of single layer shallow water equations by Dellar & Salmon (2005) to encompass several superposed layers of inviscid fluid of different, constant densities flowing over topography, as illustrated in figure 1. Dellar & Salmon (2005) corrected an earlier attempt by Bazdenkov, Morozov & Pogutse (1987), whose equations failed to conserve either energy or potential vorticity in the presence of topography. Our multilayer equations provide a useful idealized setting for studying the interaction between density stratification and rotation, and the resulting sets of two-dimensional equations are practical for numerical studies of some of the phenomena listed above.

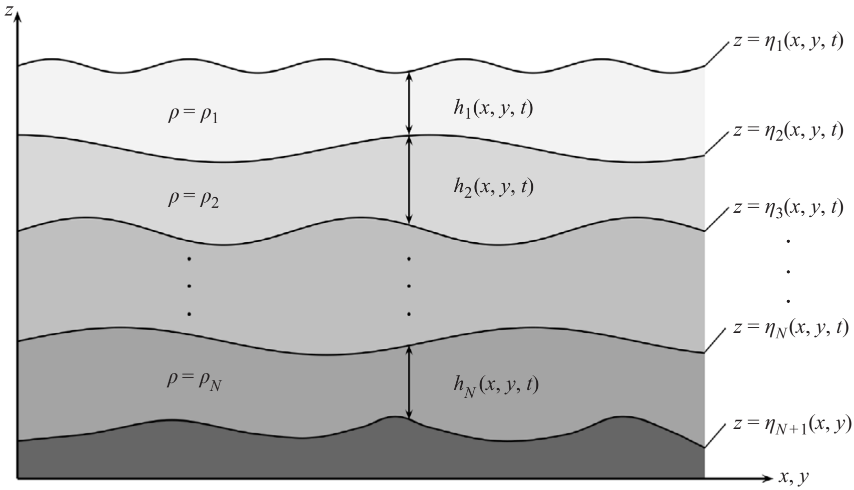

FIGURE 1. The layered model of the ocean. The upper surface of each layer is given by $z = \eta_i(x, y, t)$, and the layer thickness is given by $h_i(x, y, t)$.

The three-dimensional Euler equations for a rotating, stratified, ideal fluid possess conservation laws for energy, momentum and potential vorticity. Attention in geophysical fluid dynamics has been focused on model equations that share the same conservation laws, which are easily destroyed by making approximations directly in the equations. In addition to a derivation by averaging the three-dimensional Euler equations, we derive our multilayer shallow water equations by making approximations in a variational principle, Hamilton's principle of least action, as formulated for a three-dimensional ideal fluid. The previously mentioned conservation laws are related to symmetries in the variational principle by Noether's theorem (see § 5) and any equations derived by making approximations that preserve these symmetries will possess equivalent conservation laws. The single-layer shallow water equations may be readily derived from Hamilton's principle by integrating a three-dimensional Lagrangian across the layer (Salmon 1983, 1988, 1998). However, the extension to two or more layers is considerably more involved, because the derivation relies upon introducing Lagrangian particle labels within each layer. The transmission of pressures between layers requires some means to synchronize the positions of particles in the different layers. Our first derivation is equivalent to Salmon's (1982*b*) derivation of the two-layer traditional shallow water equations from Hamilton's principle. Salmon (1982*b*) coupled the two layers using a double integral of a delta function across both layers in the Lagrangian (see the Appendix). This approach does not readily extend to many layers, because one would need integrals across all N layers. We avoid the integrals across multiple layers by transforming each of the integrals into an integral over layer i when deriving the equations of motion for layer i. However, the calculation is still sufficiently involved that we present a second derivation that explicitly includes the work done by the pressure exerted by other layers in the Lagrangian.

The non-traditional components of the Coriolis force appear through terms involving the half-layer heights $\overline{z}_i = (1/2)\ (\eta_i + \eta_{i+1})$. This is because the non-traditional terms are linear in z when the fluid moves approximately in columns, and layer averaging a function that is linear in z is equivalent to evaluating the function at the midpoint of the layer. In particular, the potential vorticity within each layer involves the component of the planetary rotation vector $\boldsymbol{\Omega}$ that is normal to the half-layer surface (as in Dellar & Salmon 2005) rather than the vertical component as found under the traditional approximation.

The equations derived in this paper are also relevant for the development of large-scale numerical ocean models. Because of the large disparity in the horizontal and vertical length scales, many three-dimensional numerical ocean models use different discretizations in the horizontal and vertical coordinates. In particular, it is common to use an isopycnal coordinate, a constant density surface, which is also a Lagrangian coordinate, in the vertical to prevent excessive diffusion across tilted isopycnal surfaces. One may think of a layered model with many layers, as illustrated in figure 1, as arising from a Lagrangian finite-difference discretization in the vertical. The most well-known model in this class is the Miami Isopycnal Coordinate Ocean Model (MICOM) as described by Bleck *et al.* (1992) and Bleck & Chassignet (1994). The multilayer equations derived in this paper could be used to extend a layered ocean model like MICOM to include the complete Coriolis force.

2. Three-dimensional equations and coordinates

We model each layer as an inviscid, incompressible, fluid of constant density ρ_i in a frame rotating with angular velocity $\boldsymbol{\Omega}$. The fluid's motion is thus governed by the Euler equations

$$\frac{\partial \boldsymbol{u}_i}{\partial t} + (\boldsymbol{u}_i \cdot \nabla)\,\boldsymbol{u}_i + 2\,\boldsymbol{\Omega} \times \boldsymbol{u}_i + \frac{1}{\rho_i}\nabla p_i + \nabla \Phi = 0, \quad \nabla \cdot \boldsymbol{u}_i = 0, \tag{2.1}$$

in conjunction with boundary conditions at the interfaces between layers (see below). Here $\boldsymbol{u}_i$ and p_i are the velocity and pressure within the ith layer. The geopotential Φ is the combined potential for the gravitational acceleration and the centrifugal force due to rotation.

The geopotential gradient is much larger than the inertial and Coriolis terms in geophysically reasonable parameter regimes, so it must be balanced primarily by the pressure gradient. We therefore set up a coordinate system in which $\nabla\Phi = g\hat{z}$, with g being the gravitational acceleration (which by convention includes the centrifugal force). The vector $\hat{z}$ is a unit vector in the direction that is locally upward as defined by $\nabla\Phi$, and the horizontal directions are tangent to the surfaces of constant geopotential.

In theoretical studies of geophysical fluid dynamics, it is common to use Cartesian or pseudo-Cartesian coordinates (Pedlosky 1987; Salmon 1998; Vallis 2006). By pseudo-Cartesian coordinates we mean the use of curvilinear coordinates under an approximation that allows the curvilinear metric to be neglected in the equations of motion. Curvilinear coordinates are necessary because the 'horizontal' coordinates should lie within, rather than merely be tangent to, the surfaces of constant geopotential. This is the correct interpretation of the so-called beta-plane approximation to spherical geometry (Phillips 1973).

The Earth's angular velocity vector $\boldsymbol{\Omega}$ is directed parallel to the line from South pole to North pole. However, the direction of $\boldsymbol{\Omega}$ relative to local coordinates with $\hat{z}$ vertical changes with latitude, so $\boldsymbol{\Omega}$ must be spatially varying in the pseudo-Cartesian coordinates of the ocean model presented in figure 1. This approximation, retaining only the latitude dependence of the rotation vector from spherical geometry in an otherwise pseudo-Cartesian formulation, is known as the beta-plane approximation. The simpler f-plane approximation arises from taking $\boldsymbol{\Omega}$ constant, and becomes valid on length scales much smaller than the planetary radius.

We allow for arbitrary orientation of the x and y axes, generalizing the conventional axes in which the y axis points North and the x axis points East. We write $\boldsymbol{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$, and allow Ω_x, Ω_y and Ω_z to be arbitrary functions of x and y. The three-dimensional vector field $\boldsymbol{\Omega}$ must be non-divergent, $\nabla \cdot \boldsymbol{\Omega} = 0$, to ensure conservation of potential vorticity (Grimshaw 1975). To allow for spatial variation of Ω_x and Ω_y, we must therefore allow Ω_z to depend on z. We take $\Omega_z = \Omega_z(x, y, z)$ while $\Omega_x = \Omega_x(x, y)$, $\Omega_y = \Omega_y(x, y)$. This is sufficiently flexible to capture a variety of β plane approximations in which Ω_x and Ω_y, as well as Ω_z, depend on latitude. Integrating $\nabla \cdot \boldsymbol{\Omega} = 0$ with respect to z yields the following expression for Ω_z:

$$\Omega_z(x, y, z) = \Omega_{z0}(x, y) - \left(\frac{\partial \Omega_x}{\partial x} + \frac{\partial \Omega_y}{\partial y} \right) z, \tag{2.2}$$

where $\Omega_{z0} = \Omega_z|_{z=0}$.

Dellar (2009) showed that one may derive (2.1) in a pseudo-Cartesian form, together with (2.2) and expressions for Ω_x and Ω_y, by introducing suitable curvilinear coordinates into Hamilton's principle on a sphere, and then approximating for motions on length scales that are small compared with the planetary radius. In this derivation, the z-dependence of Ω_z arises as a pseudo-Cartesian approximation to the dependence of the angular momentum of a particle rotating with the planetary angular velocity $\boldsymbol{\Omega}$ on the spherical radial coordinate.

3. Derivation by layer averaging

One route to deriving our extended shallow water equations is via an extension of the standard derivation of the traditional approximation shallow water equations by averaging across layers. We obtain two-dimensional equations for the depth-averaged horizontal velocities and the layer depths by integrating the three-dimensional equations of motion over each fluid layer. Our approach follows the derivation of the non-rotating and weakly nonlinear 'great lake' equations by Camassa, Holm & Levermore (1996), as adapted by Dellar & Salmon (2005) to include the Coriolis force. Our treatment of multiple layers is similar to Liska & Wendroff's (1997) derivation of multilayer Green–Naghdi equations and to Choi & Camassa's (1996) derivation of two-layer equations for weakly nonlinear internal waves.

3.1. *Formulation and non-dimensionalization*

Within each layer we write the three-dimensional velocity vector as $(\boldsymbol{u}_i, w_i)$, where $\boldsymbol{u}_i = (u_i, v_i)$ is now a two-dimensional vector for the horizontal velocity. Separating the Euler equations (2.1) into horizontal and vertical components, we obtain

$$\frac{\partial \boldsymbol{u}_i}{\partial t} + (\boldsymbol{u}_i \cdot \nabla)\, \boldsymbol{u}_i + w_i \frac{\partial \boldsymbol{u}_i}{\partial z} + 2\Omega_z \boldsymbol{z} \times \boldsymbol{u}_i + 2\boldsymbol{\Omega} \times \hat{\boldsymbol{z}} w_i + \frac{1}{\rho_i} \nabla p_i = 0, \tag{3.1a}$$

$$\frac{\partial w_i}{\partial t} + \boldsymbol{u}_i \cdot \nabla w_i + w_i \frac{\partial w_i}{\partial z} + 2(v_i \Omega_x - u_i \Omega_y) + \frac{1}{\rho_i} \frac{\partial p_i}{\partial z} + g = 0, \tag{3.1b}$$

$$\nabla \cdot \boldsymbol{u}_i + \frac{\partial w_i}{\partial z} = 0, \tag{3.1c}$$

for $i = 1, \ldots, N$. The quantities appearing in the three-dimensional Euler equations are all functions of x, y, z and t.

We assume that each layer of fluid is bounded by an upper surface $z = \eta_i(x, y, t)$ and a lower surface $z = \eta_{i+1}(x, y, t)$. The exception is the lowest layer, the Nth layer, that flows over a fixed topography $z = \eta_{N+1}(x, y)$. For future use, we also define the layer heights $h_i = \eta_i - \eta_{i+1}$, as shown in figure 1. We assume that the upper surface of the uppermost layer is stress-free, and that the pressure is continuous across each internal surface. This leads to the following boundary conditions for the pressures:

$$p_1 = 0 \quad \text{on} \quad z = \eta_1, \quad p_i = p_{i+1} \quad \text{on} \quad z = \eta_{i+1}. \tag{3.2}$$

By considering $(\mathrm{D}/\mathrm{D}t)(z - \eta_i) = 0$ at $z = \eta_i$ in each of the two layers bounded by η_i, we obtain the kinematic boundary conditions:

$$\left.\begin{aligned} w_i &= \frac{\partial \eta_i^{(-)}}{\partial t} + \boldsymbol{u}_i \cdot \nabla \eta_i^{(-)} \quad \text{on} \quad z = \eta_i^{(-)}, \\ w_i &= \frac{\partial \eta_{i+1}^{(+)}}{\partial t} + \boldsymbol{u}_i \cdot \nabla \eta_{i+1}^{(+)} \quad \text{on} \quad z = \eta_{i+1}^{(+)}. \end{aligned}\right\} \tag{3.3}$$

The superscripts $(+)$ and $(-)$ denote that these conditions should be evaluated just above and just below the boundary, respectively, due to the discontinuity of the tangential velocity across the interfaces. Compatibility of the different expressions for $\partial_t \eta_i$ from each side of the layer is equivalent to continuity of the normal velocity across each interface.

We now apply a non-dimensionalization similar to that used by Camassa *et al.* (1996), but adapted to a rotating system. We write

$$\left.\begin{aligned} \boldsymbol{x} = L\tilde{x}, \quad z = \varepsilon L \tilde{z}, \quad \boldsymbol{u}_i = U\tilde{\boldsymbol{u}}_i, \quad w_i = \varepsilon U \tilde{w}_i, \quad p_i = 2\Omega L U \rho_i\, \tilde{p}_i, \\ t = L/U\tilde{t}, \quad \boldsymbol{\Omega} = \Omega \tilde{\boldsymbol{\Omega}}, \quad \Omega_z = \Omega \tilde{\Omega}_z, \quad \eta_i = \varepsilon L \tilde{\eta}_i, \end{aligned}\right\} \tag{3.4}$$

where U is the velocity scale, L is the length scale, $\Omega = |(\boldsymbol{\Omega}, \Omega_z)|$ is the magnitude of the Earth's angular velocity and $\varepsilon \ll 1$ is a small parameter that enforces the assumption of a shallow layer. We choose the non-dimensionalization for w_i so that the small parameter ϵ does not enter the dimensionless incompressibility condition. The dimensionless versions of equations (3.1*a*)–(3.1*c*) are thus

$$Ro\left(\frac{\partial \tilde{\boldsymbol{u}}_i}{\partial \tilde{t}} + \left(\tilde{\boldsymbol{u}}_i \cdot \tilde{\nabla}\right)\tilde{\boldsymbol{u}}_i + \tilde{w}_i \frac{\partial \tilde{\boldsymbol{u}}_i}{\partial \tilde{z}}\right) + \tilde{\Omega}_z \hat{\boldsymbol{z}} \times \tilde{\boldsymbol{u}}_i + \varepsilon\, \tilde{\boldsymbol{\Omega}} \times \hat{\boldsymbol{z}} \tilde{w}_i + \tilde{\nabla}\tilde{p}_i = 0, \tag{3.5a}$$

$$\varepsilon^2 Ro\left(\frac{\partial \tilde{w}_i}{\partial \tilde{t}} + \tilde{\boldsymbol{u}}_i \cdot \tilde{\nabla}\tilde{w}_i + \tilde{w}_i \frac{\partial \tilde{w}_i}{\partial \tilde{z}}\right) + \varepsilon(\tilde{v}_i \tilde{\Omega}_x - \tilde{u}_i \tilde{\Omega}_y) + \frac{\partial \tilde{p}_i}{\partial \tilde{z}} + Bu = 0, \tag{3.5b}$$

$$\nabla \cdot \tilde{\boldsymbol{u}}_i + \frac{\partial \tilde{w}_i}{\partial \tilde{z}} = 0, \tag{3.5c}$$

where $Ro = U/(2\Omega L)$ and $Bu = gH/(2\Omega U L)$ are the Rossby and Burger numbers, respectively. We assume Ro and Bu are both $O(1)$. Hereafter we will drop the tilde ($\tilde{}$) notation, with the understanding that all variables are dimensionless.

3.2. *Asymptotic expansion*

We wish to obtain averaged momentum equations that are accurate up to first order in the small parameter ε. We therefore pose asymptotic expansions in ε of the dependent variables $\boldsymbol{u}_i$, w_i and p_i,

$$\boldsymbol{u}_i = \boldsymbol{u}_i^{(0)} + \varepsilon \boldsymbol{u}_i^{(1)} + \cdots, \quad w_i = w_i^{(0)} + \varepsilon w_i^{(1)} + \cdots, \quad p_i = p_i^{(0)} + \varepsilon p_i^{(1)} + \cdots, \quad \text{for } i = 1, \ldots, N. \tag{3.6}$$

Substituting this expansion into (3.5*b*), we find that the leading-order pressure in each layer is just the hydrostatic pressure:

$$p_i^{(0)} = p_i^{(0)}\Big|_{z=\eta_i} + Bu(\eta_i - z). \tag{3.7}$$

The leading-order horizontal pressure gradient $\nabla p_i^{(0)}$ is thus independent of z within each layer. Additionally, non-dimensionalizing (2.2) leads to

$$\Omega_z = \Omega_{z0} - \varepsilon\,(\nabla \boldsymbol{\cdot} \boldsymbol{\Omega})\,z, \tag{3.8}$$

so Ω_z is independent of z at leading order. Equation (3.5*a*) may therefore be satisfied at leading order by a z-independent horizontal velocity $\boldsymbol{u}_i^{(0)} = \boldsymbol{u}_i^{(0)}(x, y, t)$. In other words, columnar motion is consistent with the leading-order horizontal momentum equation.

We now use the vertical momentum equation (3.5*b*) again to evaluate the first correction pressure terms:

$$p_i^{(1)} = p_i^{(1)}\Big|_{z=\eta_i} + (\eta_i - z)\left(v_i^{(0)}\Omega_x - u_i^{(0)}\Omega_y\right). \tag{3.9}$$

The combination $p_i^{(0)} + \varepsilon p_i^{(1)}$, being the result of balancing the vertical pressure gradient with the gravitational term and the vertical components of the Coriolis acceleration, is known as the 'quasi-hydrostatic' pressure (White & Bromley 1995; White *et al.* 2005). The pressure contributions from the layers above may be evaluated using the dimensionless form of the pressure boundary condition (3.2), $p_1^{(0)} + \varepsilon p_1^{(1)} = 0$ on $z = \eta_1$ and $\rho_i(p_i^{(0)} + \varepsilon p_i^{(1)}) = \rho_{i-1}(p_{i-1}^{(0)} + \varepsilon p_{i-1}^{(1)})$ on $z = \eta_i$ for $2 \leqslant i \leqslant N$. This leads to the following expression for the pressure in each layer:

$$\begin{aligned} p_i &= p_i^{(0)} + \varepsilon p_i^{(1)} + O(\varepsilon^2) \\ &= (\eta_i - z)\left(Bu + \varepsilon v_i^{(0)}\Omega_x - \varepsilon u_i^{(0)}\Omega_y\right) \\ &\quad + \sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}h_j\left(Bu + \varepsilon v_j^{(0)}\Omega_x - \varepsilon u_j^{(0)}\Omega_y\right) + O(\varepsilon^2). \end{aligned} \tag{3.10}$$

Similarly, we may determine the leading-order vertical velocity using (3.5*c*):

$$w_i^{(0)} = w_i^{(0)}\Big|_{z=\eta_{i+1}} + (\eta_{i+1} - z)\,\nabla \boldsymbol{\cdot} \boldsymbol{u}_i^{(0)}. \tag{3.11}$$

The vertical velocity in each layer acquires a contribution from those in the layers below, which may be evaluated using (3.3) in the form

$$w_i^{(0)} = w_{i+1}^{(0)} + \left(\boldsymbol{u}_i^{(0)} - \boldsymbol{u}_{i+1}^{(0)}\right) \boldsymbol{\cdot} \nabla\eta_{i+1} \quad \text{on} \quad z = \eta_{i+1}. \tag{3.12}$$

Repeated application of (3.11) and (3.12) leads to a complete expression for the leading-order vertical velocities:

$$w_i^{(0)} = \nabla \boldsymbol{\cdot} \left(\eta_{i+1}\boldsymbol{u}_i^{(0)}\right) - z\nabla \boldsymbol{\cdot} \boldsymbol{u}_i^{(0)} - \sum_{j=i+1}^{N} \nabla \boldsymbol{\cdot} \left(h_j \boldsymbol{u}_j^{(0)}\right). \tag{3.13}$$

3.3. *Averaged momentum equations*

We now derive two-dimensional equations governing the dynamics of the depth-averaged horizontal velocity in each layer. To perform averaging of (3.5*a*)–(3.5*c*), we require a result from Wu (1981) for the average of the material derivative $\mathrm{D}F/\mathrm{D}t$ across a layer of incompressible fluid bounded by the material surfaces $z = \eta_i$ and $z = \eta_{i+1}$,

$$h_i \int_{\eta_{i+1}}^{\eta_i} \left\{ \frac{\partial F}{\partial t} + \boldsymbol{u}_i \boldsymbol{\cdot} \nabla F + w_i \frac{\partial F}{\partial z} \right\} \mathrm{d}z = \frac{\partial}{\partial t}(h_i \overline{F}) + \nabla \boldsymbol{\cdot} (h_i \overline{\boldsymbol{u}_i F}). \tag{3.14}$$

Here an overbar ($\overline{}$) denotes a layer-averaged quantity. For example,

$$\overline{\boldsymbol{u}}_i = \frac{1}{h_i} \int_{\eta_{i+1}}^{\eta_i} \boldsymbol{u}_i \, \mathrm{d}z. \tag{3.15}$$

Setting $F = 1$ in (3.14) leads to an exact evolution equation for the layer height h_i:

$$\frac{\partial h_i}{\partial t} + \nabla \boldsymbol{\cdot} (h_i \overline{\boldsymbol{u}}_i) = 0. \tag{3.16}$$

Similarly, setting $F = u_i$ and $F = v_i$ allows us to integrate (3.5*a*) over each layer, as described by Camassa *et al.* (1996) and Dellar & Salmon (2005), to obtain

$$Ro \left(\frac{\partial}{\partial t} (h_i \overline{\boldsymbol{u}}_i) + \nabla \boldsymbol{\cdot} \left(h_i \overline{\boldsymbol{u}_i \ \boldsymbol{u}_i} \right) \right) + h_i \hat{\boldsymbol{z}} \times \overline{\Omega_z \boldsymbol{u}_i} \\ + \varepsilon \boldsymbol{\Omega} \times \hat{\boldsymbol{z}} \int_{\eta_{i+1}}^{\eta_i} w_i^{(0)} \, \mathrm{d}z + \int_{\eta_{i+1}}^{\eta_i} \nabla \left(p_i^{(0)} + \varepsilon p_i^{(1)} \right) \mathrm{d}z = O(\varepsilon^2). \tag{3.17}$$

To obtain evolution equations for the averaged velocities $\overline{\boldsymbol{u}}_i$, we note that $\boldsymbol{u}_i$ and Ω_z are z-independent at leading order, and so averages of their products may be factorized to sufficient accuracy (Camassa *et al.* 1996; Su & Gardner 1969) as $\overline{\boldsymbol{u}_i \boldsymbol{u}_i} = \overline{\boldsymbol{u}}_i \overline{\boldsymbol{u}}_i + O(\varepsilon^2)$, $\overline{\Omega_z \boldsymbol{u}_i} = \overline{\Omega}_z \overline{\boldsymbol{u}}_i + O(\varepsilon^2)$. We may also determine the averaged pressure gradient using (3.10),

$$\int_{\eta_{i+1}}^{\eta_i} \nabla \left(p_i^{(0)} + \varepsilon p_i^{(1)} \right) \mathrm{d}z = \tfrac{1}{2} \varepsilon h_i \left(v_i^{(0)} \Omega_x - u_i^{(0)} \Omega_y \right) \nabla (\eta_{i+1} + \eta_i) \\ + h_i \nabla \left\{ Bu \, \eta_i + \tfrac{1}{2} \varepsilon h_i \left(v_i^{(0)} \Omega_x - u_i^{(0)} \Omega_y \right) \right. \\ \left. + \sum_{j=1}^{i-1} \frac{\rho_j}{\rho_i} h_j \left(Bu + \varepsilon v_j^{(0)} \Omega_x - \varepsilon u_j^{(0)} \Omega_y \right) \right\}, \tag{3.18}$$

and the averaged vertical velocity using (3.13),

$$\int_{\eta_{i+1}}^{\eta_i} w_i^{(0)} \, \mathrm{d}z = h_i \left[\boldsymbol{u}_i^{(0)} \boldsymbol{\cdot} \nabla \eta_{i+1} - \tfrac{1}{2} h_i \nabla \boldsymbol{\cdot} \boldsymbol{u}_i^{(0)} - \sum_{j=i+1}^{N} \nabla \boldsymbol{\cdot} \left(h_j \boldsymbol{u}_j^{(0)} \right) \right]. \tag{3.19}$$

Substituting these expressions into (3.17) yields

$$
\begin{aligned}
&Ro\left(\frac{\partial}{\partial t}(h_i\overline{\boldsymbol{u}}_i)+\nabla\boldsymbol{\cdot}(h_i\overline{\boldsymbol{u}}_i\overline{\boldsymbol{u}}_i)\right)+h_i\hat{\boldsymbol{z}}\times\overline{\Omega}_z\overline{\boldsymbol{u}}_i+\tfrac{1}{2}\varepsilon h_i\left(v_i^{(0)}\Omega_x-u_i^{(0)}\Omega_y\right)\\
&\times\nabla(\eta_{i+1}+\eta_i)+h_i\nabla\left\{Bu\,\eta_i+\tfrac{1}{2}\varepsilon h_i\left(v_i^{(0)}\Omega_x-u_i^{(0)}\Omega_y\right)\right.\\
&\left.+\sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}h_j\left(Bu+\varepsilon v_j^{(0)}\Omega_x-\varepsilon u_j^{(0)}\Omega_y\right)\right\}+\varepsilon\boldsymbol{\Omega}\\
&\times\hat{\boldsymbol{z}}h_i\left[\boldsymbol{u}_i^{(0)}\boldsymbol{\cdot}\nabla\eta_{i+1}-\tfrac{1}{2}h_i\nabla\boldsymbol{\cdot}\boldsymbol{u}_i^{(0)}-\sum_{j=i+1}^{N}\nabla\boldsymbol{\cdot}\left(h_j\boldsymbol{u}_j^{(0)}\right)\right]=O(\varepsilon^2).
\end{aligned}
\tag{3.20}
$$

To complete the derivation, we note that $\boldsymbol{u}_i^{(0)}=\overline{\boldsymbol{u}}_i+O(\varepsilon)$ and that the advection terms may be simplified using (3.16). Additionally, integrating (3.8) yields an expression for $\overline{\Omega}_z$:

$$
\overline{\Omega}_z=\Omega_{z0}-\varepsilon\,\overline{z}_i\,\nabla\boldsymbol{\cdot}\boldsymbol{\Omega},
\tag{3.21}
$$

where $\overline{z}_i=(1/2)(\eta_i+\eta_{i+1})$ is the vertical position of the midsurface of the layer. The linear dependence of Ω_z on z makes the average of Ω_z across the layer equal to the value of Ω_z at the midsurface.

Neglecting terms of $O(\varepsilon^2)$ and above, and dropping the overbars on averaged velocities, we rearrange (3.20) to obtain

$$
\begin{aligned}
&Ro\left(\frac{\partial\boldsymbol{u}_i}{\partial t}+(\boldsymbol{u}_i\boldsymbol{\cdot}\nabla)\boldsymbol{u}_i\right)+\left(\Omega_{z0}-\tfrac{1}{2}\varepsilon\nabla\boldsymbol{\cdot}\left((\eta_i+\eta_{i+1})\,\boldsymbol{\Omega}\right)\right)\hat{\boldsymbol{z}}\times\boldsymbol{u}_i\\
&+\nabla\left\{Bu\,\eta_i+\tfrac{1}{2}\varepsilon h_i(v_i\Omega_x-u_i\Omega_y)+\frac{1}{\rho_i}\sum_{j=1}^{i-1}\rho_jh_j\left(Bu+\varepsilon\left(v_j\Omega_x-u_j\Omega_y\right)\right)\right\}\\
&-\varepsilon\,\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\,\nabla\boldsymbol{\cdot}\left[\tfrac{1}{2}h_i\boldsymbol{u}_i+\sum_{j=i+1}^{N}h_j\boldsymbol{u}_j\right]=0.
\end{aligned}
\tag{3.22}
$$

We thus obtain shallow water momentum equations governing the averaged horizontal fluid velocities and layer heights. We may recover the traditional multilayer shallow water equations by setting $\Omega_x=\Omega_y=0$, or equivalently by letting ε tend to zero. The terms proportional to Ω_x and Ω_y are the corrections to the traditional shallow water equations that arise from the non-traditional components of the Coriolis force.

The final term in (3.22) may be rewritten as a time derivative using the continuity equations for the layer heights:

$$
-\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\,\nabla\boldsymbol{\cdot}\left[\tfrac{1}{2}h_i\boldsymbol{u}_i+\sum_{j=i+1}^{N}h_j\boldsymbol{u}_j\right]=\frac{\partial}{\partial t}\left[\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\left(\tfrac{1}{2}h_i+\sum_{j=i+1}^{N}h_j\right)\right]=\frac{\partial}{\partial t}\left(\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\,\overline{z}_i\right),
\tag{3.23}
$$

where $\overline{z}_i(x,y,t)$ is the vertical coordinate of the midsurface of the ith layer. This term combines with the time derivative of the velocity to form the time derivative of what turns out to be the canonical momentum as shown in § 5.2. Similarly, the quantity whose gradient appears in $\nabla\{\boldsymbol{\cdot}\}$ is the pressure, given by (3.10), evaluated at the midsurface $\overline{z}_i$.

4. Derivation from a variational principle

We may also derive our extended shallow water equations (3.22) and (3.16) from the application of Hamilton's principle of least action. Hamilton's principle gives the equations of motion for a mechanical system as being those that make the action

$$\mathscr{S} = \int_{t_0}^{t_1} \mathscr{L}\,\mathrm{d}t \tag{4.1}$$

stationary over variations that vanish at the endpoints t_0 and t_1. For example, the three-dimensional Euler equations for an incompressible, inviscid fluid may be obtained from Hamilton's principle and the Lagrangian (Eckart 1960*b*)

$$\mathscr{L} = \iiint \mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c\ \frac{1}{2}\left|\frac{\partial \boldsymbol{x}}{\partial \tau}\right|^2 - p(\boldsymbol{a},\tau)\left(\frac{\partial(x,y,z)}{\partial(a,b,c)} - \frac{1}{\rho_0}\right). \tag{4.2}$$

In this formulation, the most natural extension of Lagrange's formulation of particle mechanics (as in Goldstein 1980) to hydrodynamics, fluid elements are described by their positions $\boldsymbol{x}(\boldsymbol{a},\tau)$ as functions of a set of labels $\boldsymbol{a}$ and time τ. We have returned to using $\boldsymbol{x}$ and $\boldsymbol{a}$ to denote three-dimensional vectors. A detailed description is given in the next subsection. The first term in (4.2) is identifiable as the kinetic energy of a fluid element. The second term introduces a Lagrange multiplier $p(\boldsymbol{a},\tau)$ to enforce incompressibility, expressed using the Jacobian of the label-to-particle map and a reference density ρ_0. By restricting the dependence of $\boldsymbol{x}$ on $\boldsymbol{a}$ so as to enforce columnar motion, one may derive various two-dimensional Lagrangians that lead to shallow water equations (Salmon 1982*b*, 1988; Miles & Salmon 1985).

4.1. *Particle labels*

Within each layer we let the positions of the fluid elements be $\boldsymbol{x}$, which we treat as functions of some particle labels $\boldsymbol{a}_i = (a_i, b_i, c_i)$ and time τ. In the ith layer, $\boldsymbol{x}$ denotes the position at time τ of the particle whose label is $\boldsymbol{a}_i$. To clarify, we write

$$\boldsymbol{x} = \boldsymbol{x}_i = (x_i(\boldsymbol{a}_i,\tau), y_i(\boldsymbol{a}_i,\tau), z_i(\boldsymbol{a}_i,\tau)), \tag{4.3}$$

to reflect the dependence of $\boldsymbol{x}$ on the particle labels in each layer. We use τ for time to emphasize that $\partial/\partial\tau$ means a partial derivative at fixed labels $\boldsymbol{a}_i$ rather than at fixed position $\boldsymbol{x}_i$. Thus $\partial/\partial\tau = \partial/\partial t + \boldsymbol{u}_i\cdot\nabla$ corresponds to an advective or material derivative with the velocity field defined by $\boldsymbol{u}_i(\boldsymbol{x}_i,\tau) = \partial\boldsymbol{x}_i/\partial\tau$.

We choose the particle labels $\boldsymbol{a}_i$ to be mass-weighted coordinates that satisfy

$$\mathrm{d}a_i\,\mathrm{d}b_i\,\mathrm{d}c_i = \rho_i\,\mathrm{d}x_i\,\mathrm{d}y_i\,\mathrm{d}z_i, \tag{4.4}$$

for $i = 1, \ldots, N$. This means that the density and velocity may both be expressed in terms of the label-to-particle map $\boldsymbol{x}_i(\boldsymbol{a}_i,\tau)$. Varying this map induces coordinated variations of the density and velocity, which is what distinguishes the variational principle for a fluid from the variational principle for a cloud of non-interacting particles. In particular, the density within each layer is related to the Jacobian of the map by

$$\frac{\partial(x_i, y_i, z_i)}{\partial(a_i, b_i, c_i)} = \frac{1}{\rho_i}. \tag{4.5}$$

Differentiating this relation with respect to τ leads to the incompressibility condition (3.1*c*), as in Miles & Salmon (1985). Thus, the continuity equation (kinematics) is incorporated in the label-to-particle map, while the momentum equation (dynamics) will come from Hamilton's principle.

4.2. *Formulation of the multilayer Lagrangian*

We formulate a Lagrangian for the multilayered system from the kinetic energies $\mathscr{T}_i$, potential energies $\mathscr{U}_i$ and pressure constraints $\mathscr{P}_i$ in each layer:

$$\begin{aligned}\mathscr{L} &= \sum_{i=1}^{N} \mathscr{T}_i - \mathscr{U}_i + \mathscr{P}_i, \\ &= \sum_{i=1}^{N} \iiint \mathrm{d}a_i\, \mathrm{d}b_i\, \mathrm{d}c_i \left\{ \tfrac{1}{2} \left| \frac{\partial \boldsymbol{x}_i}{\partial \tau} + \boldsymbol{R} \right|^2 - \tfrac{1}{2} |\boldsymbol{R}|^2 - g z_i + p_i(\boldsymbol{a}_i, \tau) \right. \\ &\qquad \left. \times \left(\frac{\partial(x_i, y_i, z_i)}{\partial(a_i, b_i, c_i)} - \frac{1}{\rho_i} \right) \right\}. \end{aligned} \tag{4.6}$$

The $\boldsymbol{R}$ terms arise from the Coriolis and centrifugal forces in a rotating frame, and $g z_i$ is the gravitational potential energy.

The Coriolis force is mathematically identical to the Lorentz force experienced by a charged particle in a magnetic field. We may therefore include the Coriolis force in Hamilton's principle by introducing a vector potential $\boldsymbol{R}$ such that

$$\nabla \times \boldsymbol{R} = 2\,\boldsymbol{\Omega}. \tag{4.7}$$

Here $\boldsymbol{R} = (R_x, R_y, R_z)$ and $\boldsymbol{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$ are both three-dimensional vectors. The $\boldsymbol{R}$ notation was introduced by Holm, Marsden & Ratiu (1986), by analogy with the introduction of a vector potential $\boldsymbol{A}$ for the magnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ when formulating the Lagrangian for a charged particle in a magnetic field (e.g. Goldstein 1980). However, various special cases for particular forms of $\boldsymbol{\Omega}$ appeared earlier (e.g. Salmon 1982*b*, 1983). Pursuing the analogy with magnetic fields, $\nabla \times \boldsymbol{R}$ is left unchanged by the gauge transformations $\boldsymbol{R} \to \boldsymbol{R} + \nabla\varphi$, which gives us some freedom in our choice of $\boldsymbol{R}$. Although $\boldsymbol{R}$ appears explicitly in the Lagrangian (4.6), the contribution from $\nabla\varphi$ to the Lagrangian reduces to a surface integral, which is readily shown to vanish at rigid boundaries. In addition, Dellar & Salmon (2005) showed that the integral over a free surface may be transformed into an exact time derivative, which gives no contribution to the action defined by (4.1). If $\boldsymbol{\Omega}$ is spatially uniform, $\boldsymbol{R} = \boldsymbol{\Omega} \times \boldsymbol{x}$ is a suitable vector potential. The combination $(1/2)\,|\partial \boldsymbol{x}_i/\partial\tau + \boldsymbol{R}|^2$ in (4.6) then corresponds to the kinetic energy calculated in an inertial frame. The $-(1/2)\,|\boldsymbol{R}|^2$ term in (4.6) subtracts out the effect of the centrifugal force, which we take to have been included in the gravitational acceleration, as explained in the Introduction.

More generally, our assumption that Ω_x and Ω_y are independent of z allows us to find a suitable $\boldsymbol{R}$ by imposing $R_z = 0$. We may then integrate the x and y components of (4.7) to obtain

$$\boldsymbol{R} = 2\left(F(x, y) + z\,\Omega_y,\, G(x, y) - z\,\Omega_x,\, 0\right), \tag{4.8}$$

where F and G are arbitrary functions arising from the integration of R_x and R_y with respect to z. We obtain a relation between F and G by substituting (4.8) and (2.2) for Ω_z into the z component of (4.7):

$$\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} = \Omega_{z0}(x, y). \tag{4.9}$$

This construction involving F and G is identical to that used under the traditional approximation by Salmon (1982*b*, 1983). The remaining arbitrariness in F and G is a consequence of being to make gauge transformations in $\boldsymbol{R}$ as described above.

4.3. *Dimensionless variables*

As before, we introduce dimensionless variables using (3.4), and also

$$F = \Omega L\tilde{F}, \quad G = \Omega L\tilde{G}, \quad \tau = L/U\tilde{\tau}, \quad \mathscr{L} = \frac{\tilde{\mathscr{L}}}{2\rho_1 \varepsilon L^4 \Omega U}. \tag{4.10}$$

We also introduce the dimensionless particle labels defined by

$$a_i = \rho_i^{1/3} L\tilde{a}_i, \quad b_i = \rho_i^{1/3} L\tilde{b}_i, \quad c_i = \rho_i^{1/3} \varepsilon L\tilde{c}_i, \tag{4.11}$$

so that the incompressibility condition (4.5) becomes

$$\frac{\partial(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)}{\partial(\tilde{a}_i, \tilde{b}_i, \tilde{c}_i)} = 1. \tag{4.12}$$

Here $\varepsilon \ll 1$ is introduced again to enforce the assumption that the layers of fluid are shallow.

We thus obtain the dimensionless Lagrangian

$$\begin{aligned}\tilde{\mathscr{L}} = \sum_{i=1}^{N} \frac{\rho_i}{\rho_1} \iiint \mathrm{d}\tilde{a}_i \, \mathrm{d}\tilde{b}_i \, \mathrm{d}\tilde{c}_i \Bigg\{ &\tfrac{1}{2} Ro \left(\left(\frac{\partial \tilde{x}_i}{\partial \tilde{\tau}} \right)^2 + \left(\frac{\partial \tilde{y}_i}{\partial \tilde{\tau}} \right)^2 + \varepsilon^2 \left(\frac{\partial \tilde{z}_i}{\partial \tilde{\tau}} \right)^2 \right) \\ &- Bu \, \tilde{z}_i + \varepsilon \left(\frac{\partial \tilde{x}_i}{\partial \tilde{\tau}} \tilde{\Omega}_y - \frac{\partial \tilde{y}_i}{\partial \tilde{\tau}} \tilde{\Omega}_x \right) \tilde{z}_i + \left(\frac{\partial \tilde{x}_i}{\partial \tilde{\tau}} \tilde{F} + \frac{\partial \tilde{y}_i}{\partial \tilde{\tau}} \tilde{G} \right) \\ &+ \tilde{p}_i(\tilde{\boldsymbol{a}}_i, \tau) \left(\frac{\partial(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)}{\partial(\tilde{a}_i, \tilde{b}_i, \tilde{c}_i)} - 1 \right) \Bigg\}. \end{aligned} \tag{4.13}$$

We now drop the tilde ˜ notation, with the understanding that all variables used henceforth are dimensionless.

4.4. *Restriction to columnar motion*

In § 3, we demonstrated that z-independent horizontal velocity satisfies the governing equations at leading order in ε. We will therefore follow Salmon (1983, 1988) and Miles & Salmon (1985) and restrict the fluid to columnar motion by assuming that $x_i = x_i(a_i, b_i, \tau)$ and $y_i = y_i(a_i, b_i, \tau)$ are independent of the vertical particle label c_i. Equation (4.12) then simplifies to

$$\frac{\partial z_i}{\partial c_i} = \frac{\partial(a_i, b_i)}{\partial(x_i, y_i)}. \tag{4.14}$$

Choosing $c_i = 0$ at the bottom of each layer, and $c_i = 1$ at the top, we may integrate (4.14) with respect to c_i to determine z_i:

$$z_i = h_i c_i + \eta_{i+1}. \tag{4.15}$$

This defines h_i as the reciprocal of the Jacobian of the horizontal particle positions and labels:

$$h_i = \left(\frac{\partial(a_i, b_i)}{\partial(x_i, y_i)} \right)^{-1}. \tag{4.16}$$

We write the expression this way to emphasize that h_i is more naturally treated as a function of the particle labels a_i and b_i, rather than the particle positions x_i and y_i. Differentiating $h_i(a_i, b_i, \tau)$ with respect to τ leads to the layer-averaged continuity equation (3.16), as in Miles & Salmon (1985). Substituting (4.16) into (4.15) gives

$$z_i = h_i c_i + \eta_{i+1} = h_i c_i + B + \sum_{j=i+1}^{N} h_j = h_i c_i + B + \sum_{j=i+1}^{N} \frac{\partial(a_j, b_j)}{\partial(x_j, y_j)}. \tag{4.17}$$

The vertical coordinate in each layer therefore acquires a dependence on the particle labels in all of the layers below. It is this dependence that allows each layer to respond to the motion of the layers above and below it.

Substituting these expressions into the Lagrangian (4.13), we obtain

$$\begin{aligned}\mathscr{L} = \sum_{i=1}^{N} \frac{\rho_i}{\rho_1} \iint \mathrm{d}a_i\, \mathrm{d}b_i \int_0^1 \mathrm{d}c_i \left\{ \tfrac{1}{2} Ro \left(\frac{\partial x_i}{\partial \tau}\right)^2 + \tfrac{1}{2} Ro \left(\frac{\partial y_i}{\partial \tau}\right)^2 \right. \\ \left. + \left(\frac{\partial x_i}{\partial \tau} F + \frac{\partial y_i}{\partial \tau} G\right) - \left(Bu + \varepsilon \frac{\partial y_i}{\partial \tau} \Omega_x - \varepsilon \frac{\partial x_i}{\partial \tau} \Omega_y\right)(h_i c_i + \eta_{i+1}) \right\}. \end{aligned} \tag{4.18}$$

The pressure terms involving the Lagrange multipliers $\tilde{p}_i(\tilde{\boldsymbol{a}}_i, \tau)$ have been discarded because our c_i to z_i map has been explicitly constructed to enforce incompressibility. We have also discarded terms $O(\varepsilon^2)$ and above, as in § 3, so we have also dropped the term $(\partial z_i/\partial \tau)^2$ from (4.13) to obtain (4.18). Miles & Salmon (1985) showed that retaining this term would lead to a multilayer analogue of the equations derived by Green & Naghdi (1976) using Cosserat surfaces.

We may now complete the integration over c_i in (4.18) to obtain the two-dimensional 'shallow water' Lagrangian

$$\begin{aligned}\mathscr{L} = \sum_{i=1}^{N} \frac{\rho_i}{\rho_1} \iint \mathrm{d}a_i\, \mathrm{d}b_i \left\{ \tfrac{1}{2} Ro \left(\frac{\partial x_i}{\partial \tau}\right)^2 + \tfrac{1}{2} Ro \left(\frac{\partial y_i}{\partial \tau}\right)^2 + \left(\frac{\partial x_i}{\partial \tau} F + \frac{\partial y_i}{\partial \tau} G\right) \right. \\ \left. - \left(Bu + \varepsilon \frac{\partial y_i}{\partial \tau} \Omega_x - \varepsilon \frac{\partial x_i}{\partial \tau} \Omega_y\right)\left(\tfrac{1}{2} h_i + \eta_{i+1}\right) \right\}. \end{aligned} \tag{4.19}$$

The integration over c_i leads to the appearance of $((1/2)\ h_i + \eta_{i+1})$ in the last term in the integrand. Because z_i depends linearly on c_i through (4.17), the average of any quantity that varies linearly in z_i across a layer is equal to the quantity evaluated at the midpoint of the layer.

4.5. *Derivation of momentum equations*

The most straightforward route to the shallow water equations is to require that the variations of $\mathscr{L}$ with respect to $\boldsymbol{x}_i(\boldsymbol{a}_i, \tau)$ vanish, in accordance with Hamilton's principle of least action. Having integrated over the third direction, we now return to two-dimensional vector notation and set $\boldsymbol{x}_i = (x_i, y_i)$ and $\boldsymbol{a}_i = (a_i, b_i)$. We first note that we may transform between integrals over particle labels $\mathrm{d}a_j\, \mathrm{d}b_j$ and $\mathrm{d}a_i\, \mathrm{d}b_i$ using (4.16) in the form

$$\iint \mathrm{d}a_j\, \mathrm{d}b_j\ A = \iint \mathrm{d}x_j\, \mathrm{d}y_j\ h_j A = \iint \mathrm{d}x_i\, \mathrm{d}y_i\ h_j A = \iint \mathrm{d}a_i\, \mathrm{d}b_i\ \frac{h_j A}{h_i}, \tag{4.20}$$

for any A and $j \neq i$. We see that when varying $\boldsymbol{x}_i(\boldsymbol{a}_i, \tau)$, we must transform all integrals $\mathrm{d}a_j\, \mathrm{d}b_j$ to determine their contribution to the variation. We therefore apply (4.20) to transform the Lagrangian in (4.19) into an integral over the labels in the ith layer:

$$\mathscr{L} = \iint \mathrm{d}a_i\,\mathrm{d}b_i \sum_{j=1}^{N} \frac{\rho_j h_j}{\rho_1 h_i} \left\{ \tfrac{1}{2} Ro \left(\frac{\partial x_j}{\partial \tau}\right)^2 + \tfrac{1}{2} Ro \left(\frac{\partial y_j}{\partial \tau}\right)^2 + \left(\frac{\partial x_j}{\partial \tau} F + \frac{\partial y_j}{\partial \tau} G\right) \right.$$
$$\left. - \left(Bu + \varepsilon \frac{\partial y_j}{\partial \tau} \Omega_x - \varepsilon \frac{\partial x_j}{\partial \tau} \Omega_y\right) \left(\tfrac{1}{2} h_j + B + \sum_{k=j+1}^{N} h_k\right) \right\}. \quad (4.21)$$

A more explicit approach to the transformation of integrals between layers was used by Salmon (1982*b*) and is described briefly in the Appendix. A different approach that avoids this technicality completely is described in § 4.6.

When taking variations of $\boldsymbol{x}_i$, we assume that $\partial \boldsymbol{x}_j/\partial \tau$ and h_j for all $j \neq i$ are prescribed functions of $\boldsymbol{x}$ evaluated at $\boldsymbol{x}_i$. For this we use the (non-varying) label-to-particle maps in the layers $j \neq i$, and their inverses. Similarly, B, F, G, Ω_x and Ω_y are all treated as prescribed functions of $\boldsymbol{x}$ evaluated at $\boldsymbol{x}_i$. The variation of any prescribed function $A(\boldsymbol{x})$ with respect to $\boldsymbol{x}_i$ is (Miles & Salmon 1985)

$$\delta A = \nabla A \cdot \delta \boldsymbol{x}_i. \quad (4.22)$$

To resolve the implicit dependence of h_i on $\boldsymbol{x}_i$, we rewrite (4.16) as

$$h_i \frac{\partial(x_i, y_i)}{\partial(a_i, b_i)} = 1, \quad (4.23)$$

and take variations

$$\begin{aligned} 0 = \delta \left(h_i \frac{\partial(x_i, y_i)}{\partial(a_i, b_i)} \right) &= \delta h_i \frac{\partial(x_i, y_i)}{\partial(a_i, b_i)} + h_i \frac{\partial(\delta x_i, y_i)}{\partial(a_i, b_i)} + h_i \frac{\partial(x_i, \delta y_i)}{\partial(a_i, b_i)} \\ &= \frac{\delta h_i}{h_i} + h_i \frac{\partial(x_i, y_i)}{\partial(a_i, b_i)} \left[\frac{\partial(\delta x_i, y_i)}{\partial(x_i, y_i)} + \frac{\partial(x_i, \delta y_i)}{\partial(x_i, x_i)} \right] \\ &= \frac{\delta h_i}{h_i} + \frac{\partial \delta x_i}{\partial x_i} + \frac{\partial \delta y_i}{\partial y_i}, \end{aligned} \quad (4.24)$$

to obtain (Miles & Salmon 1985)

$$\delta h_i = -h_i \nabla \cdot \delta \boldsymbol{x}_i. \quad (4.25)$$

For an arbitrary quantity Q multiplying the variation δh_i, we obtain (Miles & Salmon 1985)

$$\iint \mathrm{d}a_i\,\mathrm{d}b_i\, Q \delta h_i = -\iint \mathrm{d}a_i\,\mathrm{d}b_i\, Q h_i \nabla \cdot \delta \boldsymbol{x}_i = \iint \mathrm{d}a_i\,\mathrm{d}b_i\, \frac{1}{h_i} \nabla \left(h_i^2 Q\right) \cdot \delta \boldsymbol{x}_i. \quad (4.26)$$

The second step follows from a transformation to an integral with respect to $\mathrm{d}x\,\mathrm{d}y$, integration by parts, and then a transformation back to an integral in $\mathrm{d}a_i\,\mathrm{d}b_i$.

We now show that the majority of the terms in the integrand in (4.21) make no contribution when we vary $\boldsymbol{x}_i$. For any prescribed function $A(\boldsymbol{x})$, the variation of the functional $\mathscr{L}_A$ defined by

$$\mathscr{L}_A = \iint \mathrm{d}a_i\,\mathrm{d}b_i\, \frac{A}{h_i} \quad (4.27)$$

is given by

$$\delta\mathscr{L}_A = \iint \mathrm{d}a_i\,\mathrm{d}b_i \frac{1}{h_i}\delta A - \frac{A}{h_i^2}\delta h_i = \iint \mathrm{d}a_i\,\mathrm{d}b_i \frac{1}{h_i}\nabla A\cdot\delta\boldsymbol{x}_i - \frac{1}{h_i}\nabla\left(\frac{h_i^2 A}{h_i^2}\right)\cdot\delta\boldsymbol{x}_i = 0. \tag{4.28}$$

As we treat $\partial\boldsymbol{x}_j/\partial\tau$ and h_j as prescribed functions of $\boldsymbol{x}$ when varying $\boldsymbol{x}_i$ with $i\neq j$, many of the terms in (4.21) are of the form (4.28), and therefore make no contribution under variations of $\boldsymbol{x}_i$. Thus, when we take variations of $\mathscr{L}$ with respect to $\boldsymbol{x}_i$, we may drop all such terms, leaving

$$\begin{aligned}\delta\int \mathrm{d}\tau\mathscr{L} = \delta\int \mathrm{d}\tau \iint \mathrm{d}a_i\,\mathrm{d}b_i\frac{\rho_i}{\rho_1}\Bigg\{&-\sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}h_j\left(Bu+\varepsilon\frac{\partial y_j}{\partial\tau}\Omega_x-\varepsilon\frac{\partial x_j}{\partial\tau}\Omega_y\right)\\ &+\tfrac{1}{2}Ro\left(\frac{\partial x_i}{\partial\tau}\right)^2+\tfrac{1}{2}Ro\left(\frac{\partial y_i}{\partial\tau}\right)^2+\left(\frac{\partial x_i}{\partial\tau}F+\frac{\partial y_i}{\partial\tau}G\right)\\ &-\left(Bu+\varepsilon\frac{\partial y_i}{\partial\tau}\Omega_x-\varepsilon\frac{\partial x_i}{\partial\tau}\Omega_y\right)\left(\tfrac{1}{2}h_i+\eta_{i+1}\right)\Bigg\}.\end{aligned} \tag{4.29}$$

Thus, we are essentially taking variations of the Lagrangian for a single layer of shallow water flowing over a prescribed lower surface $\eta_{i+1}(\boldsymbol{x},t)$, as in Dellar & Salmon (2005), but with additional contributions due to the pressure inherited from each of the layers above (3.10).

Using (4.22) and (4.26) to compute the variation of (4.19) with respect to $\boldsymbol{x}_i$ gives

$$\begin{aligned}\delta\int \mathrm{d}\tau\mathscr{L} = \int \mathrm{d}\tau \iint \mathrm{d}a_i\,\mathrm{d}b_i\frac{\rho_i}{\rho_1}\Bigg\{&-Ro\,\frac{\partial^2\boldsymbol{x}_i}{\partial\tau^2}-Bu\,\nabla\eta_i\\ &-\nabla\left[\sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}h_j\left(Bu+\varepsilon\frac{\partial y_j}{\partial\tau}\Omega_x-\varepsilon\frac{\partial x_j}{\partial\tau}\Omega_y\right)\right]\\ &+\frac{\partial x_i}{\partial\tau}\nabla F_i+\frac{\partial y_i}{\partial\tau}\nabla G_i-\frac{\partial}{\partial\tau}(F_i,G_i)\\ &+\varepsilon\left(\tfrac{1}{2}h_i+\eta_{i+1}\right)\left(\frac{\partial x_i}{\partial\tau}\nabla\Omega_y-\frac{\partial y_i}{\partial\tau}\nabla\Omega_x\right)\\ &+\varepsilon\left(\frac{\partial x_i}{\partial\tau}\Omega_y-\frac{\partial y_i}{\partial\tau}\Omega_x\right)\nabla\left(\tfrac{1}{2}h_i+\eta_{i+1}\right)\\ &+\varepsilon\,\frac{\partial}{\partial\tau}\left(\left(\tfrac{1}{2}h_i+\eta_{i+1}\right)\left(-\Omega_y,\Omega_x\right)\right)\\ &+\tfrac{1}{2}\varepsilon\,\nabla\left[h_i\left(\frac{\partial x_i}{\partial\tau}\Omega_y-\frac{\partial y_i}{\partial\tau}\Omega_x\right)\right]\Bigg\}\cdot\delta\boldsymbol{x}_i,\end{aligned} \tag{4.30}$$

for $i=1,\ldots,N$. Rewriting the material time derivatives as $\partial/\partial\tau=\partial/\partial t+\boldsymbol{u}_i\cdot\nabla$, the terms involving F_i and G_i combine to give

$$u_i\nabla F_i+v_i\nabla G_i-(\boldsymbol{u}_i\cdot\nabla F_i,\boldsymbol{u}_i\cdot\nabla G_i)=(G_{ix}-F_{iy})(v_i,-u_i), \tag{4.31}$$

and $G_{ix}-F_{iy}=\Omega_{z0}$ using (4.9). Hamilton's principle, setting the integrand of (4.30) equal to zero, thus gives the same equations of motion (3.22) as before.

4.6. *Alternative formulation using a separate Lagrangian for each layer*

In the previous approach, the different layers were coupled together through the label-to-particle map. The map from the label c_i to the vertical position z_i depended upon the heights of every layer underneath. Varying the map from c_i to z_i would raise or lower every layer above, and thus change these layers' contributions to the gravitational potential energy. This is the natural way to include the pressure exerted by the layers above, but taking variations is complicated by the need to transform integrals over the upper layers into integrals with respect to a_i, b_i, c_i.

In this section, we describe a different derivation that uses a separate Lagrangian $\mathscr{L}_i$ for each layer of fluid. The label-to-particle map in each layer may be varied independently, making the derivation of the equations of motion much simpler. In particular, this approach would be much more attractive for deriving multilayer analogues of the Green & Naghdi (1976) equations that retain the vertical kinetic energy $(1/2)\dot{z}_i^2$ in each layer.

We begin with a three-dimensional Lagrangian, as before, and decompose it into the sum of contributions from each of the different layers. This leads to a Lagrangian for the multilayer system that is the sum of separate Lagrangians for each layer. The layers are coupled through an additional term representing the work done by the pressure in the layers above, analogous to the treatment of external pressure by Miles & Salmon (1985).

Returning to three-dimensional notation, we may formulate a Lagrangian for the ith layer as

$$\mathscr{L}_i = \mathscr{T}_i - \mathscr{U}_i + \mathscr{P}_i + \mathscr{W}_i, \tag{4.32}$$

where $\mathscr{T}_i$ and $\mathscr{U}_i$ are the kinetic and potential energies given in (4.6), and $\mathscr{P}_i$ is a incompressibility constraint that contains the pressure p_i as a Lagrange multiplier p_i. So far this is exactly the same as in Eckart (1960*b*) and §4.2 above. The extra contribution $\mathscr{W}_i$ is the work done on the upper surface of each layer by the layers above:

$$\mathscr{W}_i = \iiint \mathrm{d}a_i\, \mathrm{d}b_i\, \mathrm{d}c_i \left\{ -\frac{1}{\rho_i} P_i(x_i, y_i, \tau) \right\}, \tag{4.33}$$

treated analogously to the imposed external pressure on a single fluid layer by Miles & Salmon (1985). Thus, $P_i(x_i, y_i, \tau)$ is the pressure exerted on layer i by the layers above.

We may therefore write the complete three-dimensional Lagrangian as

$$\begin{aligned} \mathscr{L}_i = \iiint \mathrm{d}a_i\, \mathrm{d}b_i\, \mathrm{d}c_i \Bigg\{ & \tfrac{1}{2}\left|\frac{\partial \boldsymbol{x}_i}{\partial \tau} + \boldsymbol{R}\right|^2 - \tfrac{1}{2}|\boldsymbol{R}|^2 - g z_i \\ & + p_i(\boldsymbol{a}_i, \tau)\left(\frac{\partial(x_i, y_i, z_i)}{\partial(a_i, b_i, c_i)} - \frac{1}{\rho_i}\right) - \frac{1}{\rho_i} P_i(\boldsymbol{x}_i, \tau) \Bigg\}. \end{aligned} \tag{4.34}$$

Taking variations of (4.34) with respect to x_i, y_i z_i and p_i, and invoking Hamilton's principle of least action, we recover (3.1*a*), (3.1*b*) and (4.5), respectively. In (3.1*a*), p_i is replaced by $p_i + P_i$ and the pressure boundary condition is modified to $p_i = 0$ on $z = \eta_i$. The inclusion of P_i thereby accounts for the pressure imposed in layer i by the fluid in the layers above.

We now apply the non-dimensionalizations in (3.4), (4.10) and (4.11) to the Lagrangian in (4.34), dropping the tilde ($\tilde{}$) notation for dimensionless variables:

$$\begin{aligned}\mathscr{L}_i = \iiint \mathrm{d}a_i\,\mathrm{d}b_i\,\mathrm{d}c_i\, &\left\{\tfrac{1}{2}Ro\left(\frac{\partial x_i}{\partial \tau}\right)^2 + \tfrac{1}{2}Ro\left(\frac{\partial y_i}{\partial \tau}\right)^2\right.\\ &- Bu\, z_i + \varepsilon\left(\frac{\partial x_i}{\partial \tau}\Omega_y - \frac{\partial y_i}{\partial \tau}\Omega_x\right) z_i + \left(\frac{\partial x_i}{\partial \tau}F + \frac{\partial y_i}{\partial \tau}G\right)\\ &\left. + p_i(\boldsymbol{a}_i, \tau)\left(\frac{\partial(x_i, y_i)}{\partial(a_i, b_i)}\frac{\partial z_i}{\partial c_i} - 1\right) - P_i(x_i, y_i, \tau)\right\}.\end{aligned} \tag{4.35}$$

Here we have introduced the shallow water assumptions, restricting x_i and y_i to be independent of c_i, and neglecting terms of $O(\varepsilon^2)$. To determine $P_i(x_i, y_i, \tau)$, the unknown pressures exerted on the upper surface of each layer, we consider variations of (4.35) with respect to z_i alone:

$$\delta\int \mathrm{d}\tau\,\mathscr{L}_i = \int \mathrm{d}\tau \iiint \mathrm{d}a_i\,\mathrm{d}b_i\,\mathrm{d}c_i\left\{-Bu - \frac{\partial(x_i, y_i)}{\partial(a_i, b_i)}\frac{\partial p_i}{\partial c_i} - \varepsilon\left(\frac{\partial y_i}{\partial \tau}\Omega_x - \frac{\partial x_i}{\partial \tau}\Omega_y\right)\right\}\delta z_i. \tag{4.36}$$

Hamilton's principle states that leading-order variations of the action with respect to z_i must vanish, so it follows that the integrand in (4.36) must be uniformly equal to zero. Using (4.16) to evaluate the Jacobian multiplying $\partial p_i/\partial c_i$, this yields

$$\frac{\partial p_i}{\partial c_i} = h_i\left(Bu + \varepsilon\, v_i\Omega_x - \varepsilon\, u_i\Omega_y\right). \tag{4.37}$$

This is equivalent to what White & Bromley (1995) call quasi-hydrostatic balance in the vertical momentum equation. We may thus determine p_i by integrating with respect to c_i:

$$p_i = p_i|_{c_i=1} + (1 - c_i)h_i\left(Bu + \varepsilon\, v_i\Omega_x - \varepsilon\, u_i\Omega_y\right). \tag{4.38}$$

The Lagrangian pressure boundary condition is

$$\rho_i p_i\Big|_{c_i=1} = \rho_{i-1}p_{i-1}\Big|_{c_{i-1}=0}, \tag{4.39}$$

which corresponds to continuity of the dimensional pressure at the interface. We let $P_1 = 0$, corresponding to the stress-free boundary condition on the upper surface of the top layer, and let

$$P_i = \frac{\rho_{i-1}}{\rho_i}p_{i-1}\Big|_{c_i=1} = \sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}P_j = \sum_{j=1}^{i-1}\frac{\rho_j}{\rho_i}h_j\left(Bu + \varepsilon\, v_j\Omega_x - \varepsilon\, u_j\Omega_y\right), \tag{4.40}$$

for $i = 2, \ldots, N$. This expression for the pressure acting on the upper surface of each layer is the same as the expression calculated by layer averaging in (3.10).

We now simplify (4.35) using the assumption of columnar motion, as in §4.2. The definitions (4.15) and (4.16) mean that (4.5) is satisfied automatically, so we may drop the terms multiplied by p_i in the Lagrangian. We may then integrate with respect to

c_i by substituting in (4.15), leading to a two-dimensional 'shallow water' Lagrangian:

$$\begin{aligned}\mathscr{L}_i = \iint \mathrm{d}a_i\,\mathrm{d}b_i \int_0^1 \mathrm{d}c_i\ &\left\{\tfrac{1}{2}Ro\left(\frac{\partial x_i}{\partial \tau}\right)^2 + \tfrac{1}{2}Ro\left(\frac{\partial y_i}{\partial \tau}\right)^2 - P_i\right.\\
&\left.- \left(Bu + \varepsilon\frac{\partial y_i}{\partial \tau}\Omega_x - \varepsilon\frac{\partial x_i}{\partial \tau}\Omega_y\right)(h_i c_i + \eta_{i+1}) + \left(\frac{\partial x_i}{\partial \tau}F + \frac{\partial y_i}{\partial \tau}G\right)\right\}\\
= \iint \mathrm{d}a_i\,\mathrm{d}b_i &\left\{\tfrac{1}{2}Ro\left(\frac{\partial x_i}{\partial \tau}\right)^2 + \tfrac{1}{2}Ro\left(\frac{\partial y_i}{\partial \tau}\right)^2 - Bu\left(\tfrac{1}{2}h_i + \eta_{i+1}\right) - P_i\right.\\
&\left.+ \varepsilon\left(\frac{\partial x_i}{\partial \tau}\Omega_y - \frac{\partial y_i}{\partial \tau}\Omega_x\right)\left(\tfrac{1}{2}h_i + \eta_{i+1}\right) + \left(\frac{\partial x_i}{\partial \tau}F + \frac{\partial y_i}{\partial \tau}G\right)\right\}.\end{aligned} \tag{4.41}$$

Thus, we recover the effective Lagrangian (4.29) used to take variations with respect to $\boldsymbol{x}_i$ in §4.5.

5. Conservation properties

We now show that our non-traditional multilayer shallow water equations inherit the expected conservation laws for energy, momentum and potential vorticity from the underlying three-dimensional equations. The existence of these conservation laws is guaranteed by our variational formulation in §4, and Noether's theorem that relates symmetries in a variational principle to conservation laws (e.g. Goldstein 1980). Conservation of energy and momentum is a consequence of symmetries under translations in time and space, while material conservation of potential vorticity follows from a more subtle symmetry under relabelling of the particles (Ripa 1981; Salmon 1982*a*, 1988, 1998).

5.1. *Energy conservation*

An equation expressing conservation of energy may be obtained either by manipulating the extended shallow water equations (3.22) or by a Legendre transformation of the Lagrangian (4.19). The latter approach corresponds to finding the quantity that is conserved under time translation, as required by Noether's theorem (e.g. Goldstein 1980; Salmon 1998) because the Lagrangian does not depend explicitly upon τ. We present the energy conservation law in dimensional form for ease of interpretation:

$$\begin{aligned}&\frac{\partial}{\partial t}\left\{\sum_{i=1}^{N} \tfrac{1}{2}\rho_i h_i\,|\boldsymbol{u}_i|^2 + \tfrac{1}{2}\rho_i g h_i\,(\eta_i + \eta_{i+1})\right\}\\
&\quad + \nabla\boldsymbol{\cdot}\left\{\sum_{i=1}^{N} h_i\boldsymbol{u}_i\left(\tfrac{1}{2}\rho_i\,|\boldsymbol{u}_i|^2 + \rho_i g\eta_i + \rho_i h_i\left(v_i\Omega_x - u_i\Omega_y\right)\right.\right.\\
&\quad \left.\left. + \sum_{j=1}^{i-1}\rho_j g h_j + 2\rho_j h_j\left(v_j\Omega_x - u_j\Omega_y\right)\right)\right\} = 0.\end{aligned} \tag{5.1}$$

As usual, the energy density is unaffected by the Coriolis force, and is simply the sum of the integrals of the three-dimensional energy density $(1/2)\,\rho_i\,|\boldsymbol{u}_i|^2 + \rho_i g z$ over each layer. However, the energy flux differs from the standard shallow water form by terms proportional to Ω_x and Ω_y. These extra terms represent the work done by the quasi-hydrostatic (as opposed to purely hydrostatic) pressure.

5.2. *Canonical momenta*

The canonical momenta are best considered using the standard axes of geophysical fluid dynamics. We take the x axis pointing East, and the y axis pointing North, so that $\Omega_x = 0$, $\Omega_y = \Omega_y(y)$ and $\Omega_{z0} = \Omega_{z0}(y)$. We first turn our attention to the zonal momentum, which we expect to be conserved when the Lagrangian contains no explicit dependence on x. We therefore choose $G = 0$ and $F(y) = -\int \Omega_{z0}(y)\,\mathrm{d}y$, as in Salmon (1982*b*). We also assume a zonally symmetric topography $\eta_{N+1}(y)$ with no x dependence. The shallow water Lagrangian (4.19) then becomes

$$\mathscr{L} = \sum_{i=1}^{N} \frac{\rho_i}{\rho_1} \iint \mathrm{d}a_i\,\mathrm{d}b_i \left\{ \tfrac{1}{2} Ro\,|\dot{\boldsymbol{x}}_i|^2 + \dot{x}_i F + \left(\varepsilon \dot{x}_i \Omega_y - Bu\right)\left(\tfrac{1}{2}h_i + \eta_{i+1}\right)\right\}, \tag{5.2}$$

where $\dot{\boldsymbol{x}}_i = (\dot{x}_i, \dot{y}_i) = (\partial x/\partial\tau, \partial y/\partial\tau)$. The canonical momenta in the x direction are given by

$$p_{ix} = \frac{\delta \mathscr{L}}{\delta \dot{x}_i} = Ro\,u_i + F + \varepsilon\,\Omega_y\left(\tfrac{1}{2}h_i + \eta_{i+1}\right). \tag{5.3}$$

We do not expect any individual canonical momentum p_{ix} to be conserved. When we take variations with respect to x_i, the $\boldsymbol{x}_j$ and h_j in the other layers ($j \neq i$) are treated as prescribed functions of x_i, so there is no symmetry associated with translations in x_i alone. In other words, the form drag due to tilted interfaces between layers transfers momentum between layers.

However, there is a symmetry if we translate all of the x_i simultaneously by the same amount, letting $x_i \to x_i + \delta x$ with the same variation δx for each $i = 1, \ldots, N$. The resulting variation of the Lagrangian (5.2) is

$$\begin{aligned}\delta \int \mathrm{d}\tau\,\mathscr{L} = \int \mathrm{d}\tau \sum_{i=1}^{N} \frac{\rho_i}{\rho_1} \iint \mathrm{d}a_i\,\mathrm{d}b_i &\left\{ \left(Ro\,\dot{x}_i + F + \varepsilon\,\Omega_y\left(\tfrac{1}{2}h_i + \eta_{i+1}\right)\right)\frac{\partial(\delta x)}{\partial\tau}\right. \\ &\left. + \left(\varepsilon \dot{x}_i \Omega_y - Bu\right)\left(\tfrac{1}{2}\delta h_i + \sum_{j=i+1}^{N} \delta h_j\right)\right\}. \end{aligned} \tag{5.4}$$

Using $\delta h_i = -h_i \partial(\delta x)/\partial x$ for variations in x_i, we find that

$$\iint \mathrm{d}a_i\,\mathrm{d}b_i\,A\,\delta h_j = -\iint \mathrm{d}a_i\,\mathrm{d}b_i\,A h_j \frac{\partial(\delta x)}{\partial x} = \iint \mathrm{d}a_i\,\mathrm{d}b_i\,\frac{1}{h_i}\frac{\partial}{\partial x}\left(h_i h_j A\right)\delta x \tag{5.5}$$

for an arbitrary function A and any i and j. This result is very similar to (4.26), in that the second equality follows from a transformation to an integral over $\mathrm{d}x\,\mathrm{d}y$, integration by parts, and then transformation back to an integral over $\mathrm{d}a_i\,\mathrm{d}b_i$. It allows us to simplify (5.4) into

$$\begin{aligned}\delta \int \mathrm{d}\tau\,\mathscr{L} = \iint \mathrm{d}a_1\,\mathrm{d}b_1 \int \mathrm{d}\tau \sum_{i=1}^{N} \frac{\rho_i}{\rho_1}\frac{h_i}{h_1} &\left\{ -\frac{\partial}{\partial\tau_i}\left(Ro\,\dot{x}_i + F + \varepsilon\,\Omega_y\left(\tfrac{1}{2}h_i + \eta_{i+1}\right)\right)\right. \\ &\left. + \frac{1}{h_i}\frac{\partial}{\partial x}\left[h_i\left(\varepsilon \dot{x}_i \Omega_y - Bu\right)\left(\tfrac{1}{2}h_i + \sum_{j=i+1}^{N} h_j\right)\right]\right\}\delta x, \end{aligned} \tag{5.6}$$

where we have used (4.20) to transform each of the integrals $\mathrm{d}a_i\,\mathrm{d}b_i$ for $i = 2, \ldots, N$ into an integral $\mathrm{d}a_1\,\mathrm{d}b_1$. We write $\partial/\partial\tau_i = \partial/\partial t + \boldsymbol{u}_i \cdot \nabla$ for the material time derivative in layer i.

By Hamilton's principle of least action, the integrand in (5.6) must be equal to zero. Redimensionalizing and using (3.16), we obtain the momentum conservation equation

$$\frac{\partial}{\partial t}\left(\sum_{i=1}^{N}\rho_i h_i p_{ix}\right)+\nabla\cdot\left\{\rho_i h_i p_{ix}\boldsymbol{u}_i+\rho_i h_i\left(g-2\Omega_y u_i\right)\left(\tfrac{1}{2}h_i+\sum_{j=i+1}^{N}h_j\right)\hat{\boldsymbol{x}}\right\}=0. \tag{5.7}$$

Thus, the conserved total zonal momentum is a weighted sum of the canonical momenta over the layers.

If Ω_y, Ω_{z0} and η_{N+1} are all constants, we may find a similar conservation law for the meridional momentum by choosing $F=0$ and $G=x\Omega_{z0}$. The shallow water Lagrangian now takes the form

$$\mathscr{L}=\sum_{i=1}^{N}\frac{\rho_i}{\rho_1}\iint \mathrm{d}a_i\,\mathrm{d}b_i\left\{\tfrac{1}{2}Ro\,|\dot{\boldsymbol{x}}_i|^2+\dot{y}_i G+\left(\varepsilon\dot{x}_i\Omega_y-Bu\right)\left(\tfrac{1}{2}h_i+\eta_{i+1}\right)\right\}, \tag{5.8}$$

and the canonical y momenta are now given by

$$p_{iy}=\frac{\delta\mathscr{L}}{\delta\dot{y}_i}=Ro\,v_i+x\Omega_{z0}. \tag{5.9}$$

We thus obtain a conservation law for the y momentum by a process similar to that described above:

$$\frac{\partial}{\partial t}\left(\sum_{i=1}^{N}\rho_i h_i p_{iy}\right)+\nabla\cdot\left\{\rho_i h_i p_{iy}\boldsymbol{u}_i+\rho_i h_i\left(g-2\Omega_y u_i\right)\left(\tfrac{1}{2}h_i+\sum_{j=i+1}^{N}h_j\right)\hat{\boldsymbol{y}}\right\}=0. \tag{5.10}$$

There is no one choice for F and G that allows us to express conservation of x and y momentum simultaneously, but there are two conserved components of momentum when the rotation vector and the bottom topography are constants. On a beta-plane, for example, we would not expect a conserved meridional momentum because Ω_z depends explicitly on latitude y. We also see that the conserved zonal momentum contains terms proportional to Ω_y, whilst the conserved meridional momentum does not depend upon Ω_y. This is because in these standard axes the non-traditional components of the Coriolis force act vertically and zonally, but not meridionally; see (3.1a)–(3.1c).

5.3. *Potential vorticity*

Material conservation of potential vorticity is even more important in geophysical fluid dynamics than conservation of energy and momentum. Both energy and momentum may be transported over large distances by waves, while potential vorticity remains tied to fluid particles. Each layer of our equations possesses a potential vorticity q_i that obeys the conservation law $\partial_t q_i+\boldsymbol{u}_i\cdot\nabla q_i=0$, with

$$q_i=\frac{1}{h_i}\left\{\left[\Omega_{z0}-\tfrac{1}{2}\varepsilon\nabla\cdot\left((\eta_i+\eta_{i+1})\boldsymbol{\Omega}\right)\right]+Ro\left(\frac{\partial v_i}{\partial x}-\frac{\partial u_i}{\partial y}\right)\right\}. \tag{5.11}$$

This expression for q_i differs from the standard shallow water potential vorticity by the term $-(1/2)\,\varepsilon\nabla\cdot((\eta_i+\eta_{i+1})\boldsymbol{\Omega})$ that contains the horizontal components Ω_x and Ω_y of the rotation vector. Equivalently, if we expand the divergence into two terms, we obtain

$$\Omega_{z0}-\bar{z}_i\nabla\cdot\boldsymbol{\Omega}-\boldsymbol{\Omega}\cdot\nabla\bar{z}_i=\Omega_z-\boldsymbol{\Omega}\cdot\nabla\bar{z}_i=(\boldsymbol{\Omega},\Omega_z)\cdot\nabla(z-\bar{z}_i(x,y,t)), \tag{5.12}$$

where $\bar{z}_i = (1/2)(\eta_i + \eta_{i+1})$ is the z coordinate of the midsurface of the ith layer. The non-traditional effects therefore replace the vertical component of rotation vector, as found in the standard shallow water potential vorticity, by the component perpendicular to the layer's midsurface $z = \bar{z}_i(x, y, t)$.

The potential vorticity conservation law with the expression (5.11) for q_i may be obtained from the curl of (3.22), or we may find q_i directly from the canonical momenta. The particle relabelling symmetry (e.g. Ripa 1981; Salmon 1982*a*, 1988, 1998) implies material conservation of

$$q_i = \frac{1}{h_i}\left(\frac{\partial p_{iy}}{\partial x_i} - \frac{\partial p_{ix}}{\partial y_i}\right) \tag{5.13}$$

for any Lagrangian that depends on the particle labels a_i and b_i only through the height h_i formed from the Jacobian $\partial(x_i, y_i)/\partial(a_i, b_i)$. Moreover, the combination of p_{ix} and p_{iy} appearing in q_i is invariant under changes of gauge in $\boldsymbol{R}$, i.e. it is the same for all possible choices of F and G, even though p_{ix} and p_{iy} themselves are gauge dependent.

6. Non-canonical Hamiltonian structure

Our equations may also be expressed using the non-canonical Hamiltonian structure for multilayered shallow water equations formulated by Ripa (1993). The non-canonical Hamiltonian formalism is convenient for fluid systems expressed using Eulerian (space-fixed) variables, as described by Shepherd (1990), Morrison (1998) and Salmon (1988, 1998). It specifies the evolution of any functional $\mathscr{F}$ as being given by $\mathscr{F}_t = \{\mathscr{F}, \mathscr{H}\}$ in terms of a Hamiltonian functional $\mathscr{H}$, and a Poisson bracket $\{\cdot\,,\cdot\}$ that satisfies certain geometrical properties.

Using dimensional variables for simplicity and writing the fluid velocity as $\boldsymbol{u}_i = (u_{ix}, u_{iy})$, the evolution of the density-weighted canonical momenta

$$v_{ix} = \rho_i p_{ix} = \rho_i \left(u_{ix} + F + 2\Omega_y \left(\tfrac{1}{2}h_i + \eta_{i+1}\right)\right), \quad v_{iy} = \rho_i p_{iy} = \rho_i u_{iy}, \tag{6.1}$$

and the layer depth h_i under our non-traditional multilayer shallow water equations is given by

$$\frac{\partial}{\partial t}\begin{pmatrix} v_{ix} \\ v_{iy} \\ h_i \end{pmatrix} = -\begin{pmatrix} 0 & -\rho_i q_i & \partial_x \\ \rho_i q_i & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix}\begin{pmatrix} \delta\mathscr{H}/\delta v_{ix} \\ \delta\mathscr{H}/\delta v_{iy} \\ \delta\mathscr{H}/\delta h_i \end{pmatrix}. \tag{6.2}$$

The Hamiltonian is the energy density from § 5.1, but expressed in terms of v_{ix}, v_{iy} and h_i:

$$\mathscr{H} = \sum_{k=1}^{N} \frac{h_k}{2\rho_k}\left\{\left[v_{kx} - \rho_k\left(F + 2\Omega_y\left(\tfrac{1}{2}h_k + \eta_{k+1}\right)\right)\right]^2 + v_{ky}^2\right\} + g\rho_k h_k\left(\tfrac{1}{2}h_k + \eta_{k+1}\right). \tag{6.3}$$

Calculation of the variational derivative $\delta\mathscr{H}/\delta h_i$ is complicated by the hidden dependence of η_k on $h_k, \ldots, h_N$ through the relation

$$\eta_k = \eta_{N+1} + \sum_{j=k}^{N} h_j, \tag{6.4}$$

where $\eta_{N+1}(x, y)$ is the fixed bottom topography. The calculations are essentially the same as those computing the variation in the potential energy part of the Lagrangian in § 4. The combination $(1/2)\, h_k + \eta_{k+1}$ appearing in (6.3) is the midpoint of layer k, denoted $\tilde{h}_k$ by Ripa (1993).

All the coupling between layers is thus expressed through the Hamiltonian. The Poisson bracket that generates (6.2) may be written as a sum of standard shallow water Poisson brackets (e.g. Shepherd 1990) for each layer, as done by Ripa (1993),

$$\{\mathscr{F}, \mathscr{G}\} = \sum_{i=1}^{N} \iint \mathrm{d}x\, \mathrm{d}y\; \rho_i q_i \hat{\boldsymbol{z}} \boldsymbol{\cdot} \left(\frac{\delta F}{\delta \boldsymbol{v}_i} \times \frac{\delta G}{\delta \boldsymbol{v}_i} \right) + \frac{\delta G}{\delta h_i} \boldsymbol{\nabla} \boldsymbol{\cdot} \left(\frac{\delta F}{\delta \boldsymbol{v}_i} \right) - \frac{\delta F}{\delta h_i} \boldsymbol{\nabla} \boldsymbol{\cdot} \left(\frac{\delta G}{\delta \boldsymbol{v}_i} \right). \tag{6.5}$$

This definition holds for any functionals $\mathscr{F}$ and $\mathscr{G}$ satisfying suitable boundary conditions that allow integrations by parts in (6.5) without generating surface terms. The Poisson bracket may be shown to be bilinear, antisymmetric and to satisfy the Jacobi identity $\{\mathscr{F}, \{\mathscr{G}, \mathscr{K}\}\} + \{\mathscr{G}, \{\mathscr{K}, \mathscr{F}\}\} + \{\mathscr{K}, \{\mathscr{F}, \mathscr{G}\}\} = 0$ for all functionals $\mathscr{F}$, $\mathscr{G}$ and $\mathscr{K}$. Equation (6.2) then follows from (6.5) and the evolution equation $\mathscr{F}_t = \{\mathscr{F}, \mathscr{H}\}$ by setting $\mathscr{F}$ equal to v_{ix}, v_{iy} and h_i in turn. Conservation laws like those listed in § 5 may be derived from properties of the Poisson bracket, especially the existence of so-called Casimir functionals $\mathscr{C}$ that satisfy $\{\mathscr{C}, \mathscr{F}\} = 0$ for all functionals $\mathscr{F}$. A full description is given by Ripa (1993) and in survey articles by Shepherd (1990), Morrison (1998) and Salmon (1988, 1998).

7. Conclusion

We have derived multilayer shallow water equations that include a complete treatment of the Coriolis force, thus extending the single-layer equations of Dellar & Salmon (2005) to multiple layers. We have presented a derivation of our equations by direct averaging of the three-dimensional Euler equations across layers, and two derivations by averaging three-dimensional Lagrangians in Hamilton's variational principle. Our two variational derivations differ in their treatment of the coupling between layers. The latter derivations guarantee the existence of conservation laws for energy, momentum and potential vorticity in our equations. These laws are related to symmetries of the variational principle by Noether's theorem, and the symmetries are preserved by our averaging procedure. Our construction of a vector potential for a wide class of spatially varying $\boldsymbol{\Omega}$ extends the variational formulation of Dellar & Salmon (2005), which relied upon constant $\boldsymbol{\Omega}$, and corrects an error in their derivation by averaging the three-dimensional equations when $\partial_x \Omega_x + \partial_y \Omega_y \neq 0$.

This coupling between layers makes our derivations, especially the derivations from Hamilton's principle, much more involved than those for a single layer. Our three-dimensional variational formulation is expressed using Lagrangian particle labels. This gives a formulation very close to Hamilton's principle for particle mechanics and avoids the need to introduce extraneous variables such as Lin constraints or Clebsch potentials (see e.g. Salmon 1988). Lagrangian particle labels are also very convenient for representing the interfaces between different fluid layers, which are themselves Lagrangian surfaces. However, the reconstruction of particle positions from the labels introduces a hidden coupling between layers. The vertical position of a particle in layer i depends on the vertical position of the lower boundary of the layer, η_{i+1} in our notation, which, in turn, depends upon the labels in the layers $i+1, \ldots, N$ below.

Our first variational derivation uses the natural Lagrangian of kinetic energy minus gravitational potential energy, plus an incompressibility constraint multiplied by a pressure as a Lagrange multiplier. This is the Lagrangian that is given by Eckart (1960*b*) for a homogenous fluid. However, the derivation of the equations of motion in a layered setting requires a very intricate exchange of integration variables between the different layers. This is because the coupling between layers is exerted by particles in adjacent positions on either side of a layer, not by particles with adjacent labels. This coupling was made explicit in a two-layer formulation by Salmon (1982*b*) that contained a double integral of a delta-function to tie the particle positions in the two layers together. This formulation is equivalent to ours (see the Appendix), but does not scale up easily to three or more layers. One would need to include triple and higher integrals across all the layers in the system. As an alternative, we made the coupling between layers explicit by introducing additional terms the Lagrangian. These term represent the work done on each layer by the pressure exerted by the layers above (c.f. Miles & Salmon 1985). With these extra terms to make the previously hidden coupling explicit, we derived the same equations of motion from independent variations of the label-to-particle map within each layer.

The momentum equations we have derived are, in dimensionless form,

$$
\begin{aligned}
&Ro\left(\frac{\partial \boldsymbol{u}_i}{\partial t} + (\boldsymbol{u}_i\cdot\nabla)\boldsymbol{u}_i\right) + \left(\Omega_{z0} - \tfrac{1}{2}\varepsilon\nabla\cdot\big((\eta_i+\eta_{i+1})\,\boldsymbol{\Omega}\big)\right)\hat{\boldsymbol{z}}\times\boldsymbol{u}_i \\
&\quad + \nabla\left\{Bu\,\eta_i + \tfrac{1}{2}\varepsilon h_i(v_i\Omega_x - u_i\Omega_y) + \frac{1}{\rho_i}\sum_{j=1}^{i-1}\rho_j h_j\big(Bu + \varepsilon\left(v_j\Omega_x - u_j\Omega_y\right)\big)\right\} \\
&\quad - \varepsilon\,\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\,\nabla\cdot\left\{\tfrac{1}{2}h_i\boldsymbol{u}_i + \sum_{j=i+1}^{N} h_j\boldsymbol{u}_j\right\} = 0,
\end{aligned}
\tag{3.22}
$$

together with the usual continuity equations $\partial_t h_i + \nabla\cdot(h_i\boldsymbol{u}_i) = 0$. They contain several non-traditional corrections to the standard multilayer shallow water equations, as derived under the traditional approximation. The traditional Coriolis term $2\Omega_z\hat{\boldsymbol{z}}\times\boldsymbol{u}_i$ is modified by replacing the vertical component Ω_z with the component of the full rotation vector $\boldsymbol{\Omega}$ that is perpendicular to each layer's midsurface. Second, the pressure changes from the hydrostatic pressure to the quasi-hydrostatic pressure (White & Bromley 1995; White *et al.* 2005), due to the non-traditional Coriolis term $v_i\Omega_x - u_i\Omega_y$ in the vertical momentum equation. The last term in (3.22) has no analogue under the traditional approximation. It arises from the non-traditional Coriolis force due to the vertical velocity, as reconstructed from the divergence of the horizontal velocity under the assumption of columnar motion, and may be rewritten as the time derivative $\partial_t\,(\epsilon\,\boldsymbol{\Omega}\times\hat{\boldsymbol{z}}\,\overline{z}_i)$, where $\overline{z}_i(x, y, t)$ is the vertical coordinate of the midsurface of the ith layer. This time derivative then combines with the time derivative of the velocity to form the time derivative of the canonical momentum as shown in § 5.2.

We have shown that these equations inherit conservation laws for energy, momentum and potential vorticity from the underlying three-dimensional equations, as is guaranteed by our derivations from Hamilton's principle. The conserved components of momentum include additional non-traditional terms proportional to $\overline{z}_i$ as explained above. These terms represent the angular momentum gained or lost as fluid elements change their vertical position, and hence their distance from the rotation axis. This effect is absent in the traditional approximation, because a fluid element displaced vertically is also displaced parallel to the rotation axis. The conserved energy density is unchanged by non-traditional effects, just as it is unchanged by rotation about a vertical axis, but the energy flux gains additional terms reflecting the work done by the quasi-hydrostatic (as opposed to purely hydrostatic) pressure on the boundary of a control volume. Finally, the potential vorticity q_i that is materially conserved within each layer becomes

$$q_i = \frac{1}{h_i}\left\{\Omega_{z0} - \varepsilon\nabla\cdot(\overline{z}_i\boldsymbol{\Omega}) + Ro\left(\frac{\partial v_i}{\partial x} - \frac{\partial u_i}{\partial y}\right)\right\}. \tag{5.11}$$

The vertical component Ω_z of the rotation vector is replaced by the component perpendicular to the layer's midsurface $z=\overline{z}_i(x, y, t)$. We expect this change to be significant in the dynamics of cross-equatorial ocean currents, because the change in sign of Ω_z at the equator severely constrains the ability of fluid parcels to cross the equator (e.g. Stommel & Arons 1960; Nof & Olson 1993). In our non-traditional equations, this constraint may be at least partly alleviated by the interaction of non-traditional Coriolis effects with suitable topography.

In Part II of this paper, we focus attention on the two-layer version of our equations. We show that, like the standard two-layer shallow water equations, they are well-posed for geophysically reasonable values of the velocity difference between the two layers. We then turn to a study of linear waves and show that our two-layer equations support sub-inertial waves. These waves are permitted only by the presence of the non-traditional Coriolis terms and may play an important role in transferring energy from near-surface waves into the deep ocean, and hence in driving mixing in the deep ocean (Gerkema & Shrira 2005*a*,*b*). Our study identifies a distinguished limit in which sufficiently long near-inertial waves are substantially affected by even notionally very small non-traditional effects. These effects couple the eastward and westward propagating branches of surface and internal waves. Long eastward surface waves connect with westward internal waves and vice versa. More work will explore analytical solutions for cross-equatorial currents, like those of Nof & Olson (1993), in the two-layer version of our equations.

A.L.S. is supported by an EPSRC Doctoral Training Account award. P.J.D.'s research is supported by an EPSRC Advanced Research Fellowship grant EP/E054625/1.

Appendix. Connection with Salmon's two-layer variational formulation

The two-layer version of our derivation in §4.5 is equivalent to Salmon's (1982*b*) derivation of the two-layer shallow water equations from the Lagrangian

$$\mathscr{L} = \rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1\,L_1 + \rho_2 \iint \mathrm{d}a_2\,\mathrm{d}b_2\,L_2 + \rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1 \iint \mathrm{d}a_2\,\mathrm{d}b_2\,L_{12}, \tag{A.1}$$

which we write as $\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_{12}$. The Lagrangian densities for $i=1, 2$ are

$$L_i = \frac{1}{2}\left(\frac{\partial x_i}{\partial \tau}\right)^2 + \frac{1}{2}\left(\frac{\partial y_i}{\partial \tau}\right)^2 - \frac{1}{2} g \frac{\partial(a_i, b_i)}{\partial(x_i, y_i)} \tag{A.2}$$

in dimensional variables, and we have excluded rotation for simplicity. The two layers are coupled through $\mathscr{L}_{12}$, which is expressed as a simultaneous integral over both layers of a delta function density,

$$L_{12} = -g\delta(\boldsymbol{x}_1 - \boldsymbol{x}_2), \tag{A.3}$$

that ties together the particle positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in the two layers.

Using $\mathrm{d}a_1\,\mathrm{d}b_1 = h_1(\boldsymbol{x}_1, t)\,\mathrm{d}x_1\,\mathrm{d}y_1$ we transform $\mathscr{L}_{12}$ into

$$\mathscr{L}_{12} = -\rho_1 \iint \mathrm{d}x_1\,\mathrm{d}y_1 \iint \mathrm{d}a_2\,\mathrm{d}b_2\; g h_1(\boldsymbol{x}_1, t)\;\delta(\boldsymbol{x}_1 - \boldsymbol{x}_2), \tag{A.4}$$

and then perform the integrations over x_1 and y_1 to obtain

$$\mathscr{L}_{12} = -\rho_1 \iint \mathrm{d}a_2\,\mathrm{d}b_2\; g h_1(\boldsymbol{x}_2, t). \tag{A.5}$$

The total Lagrangian (A.1) thus becomes

$$\mathscr{L} = \rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1 \left\{\tfrac{1}{2}|\dot{\boldsymbol{x}}_1|^2 - \tfrac{1}{2} g h_1\right\} + \rho_2 \int \mathrm{d}a_2\,\mathrm{d}b_2 \left\{\tfrac{1}{2}|\dot{\boldsymbol{x}}_2|^2 - \tfrac{1}{2} g h_2 - \frac{\rho_1}{\rho_2} g h_1(\boldsymbol{x}_2, t)\right\}, \tag{A.6}$$

which is the same as (4.29) above with $N = 2$ and $i = 2$. Varying the map $(\boldsymbol{a}_2, \tau) \mapsto \boldsymbol{x}_2(\boldsymbol{a}_2, \tau)$ gives the lower layer equation of motion.

Conversely, using $\mathrm{d}a_2\,\mathrm{d}b_2 = h_2(\boldsymbol{x}_2, t)\,\mathrm{d}x_2\,\mathrm{d}y_2$, we transform $\mathscr{L}_{12}$ into

$$\mathscr{L}_{12} = -\rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1 \iint \mathrm{d}x_2\,\mathrm{d}y_2\; g h_2(\boldsymbol{x}_2, t)\;\delta(\boldsymbol{x}_1 - \boldsymbol{x}_2), \tag{A.7}$$

and then perform the integration over x_2 to obtain

$$\mathscr{L}_{12} = -\rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1\; g h_2(\boldsymbol{x}_1, t). \tag{A.8}$$

The total Lagrangian (A.1) then becomes

$$\mathscr{L} = \rho_1 \iint \mathrm{d}a_1\,\mathrm{d}b_1 \left\{\tfrac{1}{2}|\dot{\boldsymbol{x}}_1|^2 - \tfrac{1}{2} g h_1 - g h_2(\boldsymbol{x}_1, t)\right\} + \rho_2 \iint \mathrm{d}a_2\,\mathrm{d}b_2 \left\{\tfrac{1}{2}|\dot{\boldsymbol{x}}_2|^2 - \tfrac{1}{2} g h_2\right\}, \tag{A.9}$$

which is the same as (4.29) above with $N = 2$ and $i = 1$. Varying the map $(\boldsymbol{a}_1, \tau) \mapsto \boldsymbol{x}_1(\boldsymbol{a}_1, \tau)$ gives the upper layer equation of motion.

Salmon's (1982*b*) expression of $\mathscr{L}_{12}$ as an integral over both layers explicitly indicates that it contributes to the equations of motion in both layers, as found by varying $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ independently. However, extending this approach to n layers would require writing the coupling terms as integrals over all n layers. This is avoided by the approaches we presented in this paper. Our first approach is mathematically equivalent to Salmon's, but we transform directly from (A.5) to (A.8) without the intermediate multiple integral. Our second approach avoids this issue completely by expressing the coupling using explicit $\mathscr{W}_i$ terms in the Lagrangians.

REFERENCES

Bazdenkov, S. V., Morozov, N. N. & Pogutse, O. P. 1987 Dispersive effects in two-dimensional hydrodynamics. *Sov. Phys. Dokl.* **32**, 262.

Bleck, R. & Chassignet, E. 1994 Simulating the oceanic circulation with isopycnic-coordinate models. In *The Oceans: Physical–Chemical Dynamics and Human Impact* (ed. S. Majumdar, E. Miller, G. Forbe, R. Schmalz & A. Panah), pp. 17–39. Pennsylvania Academic Science.

Bleck, R., Rooth, C., Hu, D. & Smith, L. T. 1992 Salinity-driven thermocline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic. *J. Phys. Oceanogr.* **22**, 1486–1505.

Camassa, R., Holm, D. D. & Levermore, C. D. 1996 Long-time effects of bottom topography in shallow water. *Physica* D **98**, 258–286.

Choi, W. & Camassa, R. 1996 Weakly nonlinear internal waves in a two-fluid system. *J. Fluid Mech.* **313**, 83–103.

Cullen, M. J. P. 1993 The unified forecast/climate model. *Met. Magazine* **122**, 81–94.

Dellar, P. J. 2009 Variations on a beta plane: derivation of nontraditional beta plane equations from Hamilton's principle on a sphere. *J. Fluid Mech.* (submitted)

Dellar, P. J. & Salmon, R. 2005 Shallow water equations with a complete Coriolis force and topography. *Phys. Fluids* **17**, 106601–19.

Eckart, C. 1960*a* *Hydrodynamics of Oceans and Atmospheres*. Pergamon.

Eckart, C. 1960*b* Variation principles of hydrodynamics. *Phys. Fluids* **3**, 421–427.

Fu, L.-L. 1981 Observations and models of inertial waves in the deep ocean. *Rev. Geophys.* **19**, 141–170.

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

Gerkema, T. & Shrira, V. I. 2005*b* Near-inertial waves on the non-traditional β-plane. *J. Geophys. Res.* **110**, C10003.

Gerkema, T., Zimmerman, J. T. F., Maas, L. R. M. & van Haren, H. 2008 Geophysical and astrophysical fluid dynamics beyond the traditional approximation. *Rev. Geophys.* **46**, RG2004–33.

Goldstein, H. 1980 *Classical Mechanics*, 2nd edn. Addison-Wesley.

Green, A. E. & Naghdi, P. M. 1976 A derivation of equations for wave propagation in water of variable depth. *J. Fluid Mech.* **78**, 237–246.

Grimshaw, R. H. J. 1975 A note on the β-plane approximation. *Tellus* A **27**, 351–357.

van Haren, H. & Millot, C. 2005 Gyroscopic waves in the Mediterranean sea. *Geophys. Res. Lett.* **32**, L24614.

Holm, D. D., Marsden, J. E. & Ratiu, T. S. 1986 *Hamiltonian Structure and Lyapunov Stability for Ideal Continuum Dynamics*. University of Montreal Press.

Leibovich, S. & Lele, S. K. 1985 The influence of the horizontal component of Earth's angular velocity on the instability of the Ekman layer. *J. Fluid Mech.* **150**, 41–87.

Liska, R. & Wendroff, B. 1997 Analysis and computation with stratified fluid models. *J. Comput. Phys.* **137**, 212–244.

Marshall, J. & Schott, F. 1999 Open-ocean convection: observations, theory, and models. *Rev. Geophys.* **37**, 1–64.

Miles, J. & Salmon, R. 1985 Weakly dispersive nonlinear gravity waves. *J. Fluid Mech.* **157**, 519–531.

Morrison, P. J. 1998 Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.* **70**, 467–521.

Munk, W. 1981 Internal waves and small-scale processes. In *Evolution of Physical Oceanography* (ed. B. A. Warren & C. Wunsch), pp. 264–291. Massachusetts Institute of Technology Press.

Munk, W. & Phillips, N. 1968 Coherence and band structure of inertial motion in the sea. *Rev. Geophys.* **6**, 447–472.

Nof, D. & Olson, D. B. 1993 How do western abyssal currents cross the equator? *Deep Sea Res.* **40**, 235–255.

Pedlosky, J. 1987 *Geophysical Fluid Dynamics*, 2nd edn. Springer.

Phillips, N. 1973 Principles of large scale numerical weather prediction. In *Dynamic Meteorology: Lectures Delivered at the Summer School of Space Physics of the Centre National d'Études Spatiales* (ed. P. Morel), pp. 3–96. D. Reidel.

Phillips, N. A. 1966 The equations of motion for a shallow rotating atmosphere and the traditional approximation. *J. Atmos. Sci.* **23**, 626–628.

Ripa, P. 1981 Symmetries and conservation laws for internal gravity waves. In *Nonlinear Properties of Internal Waves* (ed. B. J. West), *AIP Conference Proceedings*, vol. 76, pp. 281–306. American Institute of Physics.

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

Salmon, R. 1982*a* Hamilton's principle and Ertel's theorem. In *Mathematical Methods in Hydrodynamics and Integrability of Dynamical Systems* (ed. M. Tabor & Y. M. Treve), *AIP Conference Proceedings*, vol. 88, pp. 127–135. American Institute of Physics.

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

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

Salmon, R. 1988 Hamiltonian fluid mechanics. *Annu. Rev. Fluid Mech.* **20**, 225–256.

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

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

Stommel, H. & Arons, A. B. 1960 On the abyssal circulation of the world ocean. II. An idealized model of the circulation pattern and amplitude in oceanic basins. *Deep Sea Res.* **6**, 217–218.

Su, C. H. & Gardner, C. S. 1969 Korteweg--de Vries equation and generalizations. III. Derivation of the Korteweg–de Vries equation and Burgers equation. *J. Math. Phys.* **10**, 536–539.

Vallis, G. K. 2006 *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press.

White, A. A. & Bromley, R. A. 1995 Dynamically consistent, quasi-hydrostatic equations for global models with a complete representation of the Coriolis force. *Quart. J. R. Meteorol. Soc.* **121**, 399–418.

White, A. A., Hoskins, B. J., Roulstone, I. & Staniforth, A. 2005 Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic. *Quart. J. R. Meteorol. Soc.* **131**, 2081–2107.

Wu, T. Y. 1981 Long waves in ocean and coastal waters. *J. Engng Mech. Div. (Amer. Soc. Civil Engng)* **107**, 501–522.