Pressure Forcing and Dispersion Analysis for Discontinuous Galerkin Approximations to Oceanic Fluid Flows

(*Journal of Computational Physics*, vol. 249, pp. 36–66, 2013)

Robert L. Higdon

Department of Mathematics
Oregon State University
Corvallis, Oregon 97331-4605

Abstract

This paper is part of an effort to examine the application of discontinuous Galerkin (DG) methods to the numerical modeling of the general circulation of the ocean. One step performed here is to develop an integral weak formulation of the lateral pressure forcing that is suitable for usage with a DG method and with a generalized vertical coordinate that includes level, terrain-fitted, isopycnic, and hybrid coordinates as examples. This formulation is then tested, in special cases, with analyses of dispersion relations and numerical stability and with some computational experiments. These results suggest that the advantages of DG methods may significantly outweigh their disadvantages, in the settings tested here. This paper also outlines some other issues that need to be addressed in future work.

Keywords: ocean modeling, multi-layer ocean models, shallow water equations, well-balanced forcing, discontinuous Galerkin method

1. Introduction

The purpose of this paper is to derive and examine some properties of discontinuous Galerkin (DG) methods, as applied to the numerical modeling of ocean circulation.

Operational ocean models have traditionally used finite difference and finite volume methods on structured rectangular grids, although the idea of unstructured Voronoi grids is presently under development (Ringler et al. [26]). In practice, structured rectangular grids have usually been used with staggered arrangements of grid points known as the B-grid and the C-grid. Such grids have an advantage of simplicity, but they can allow troublesome grid noise and

Email address: higdon@math.oregonstate.edu (Robert L. Higdon)

can give inaccurate propagation of inertia-gravity waves and/or Rossby waves, depending on the relation between the grid size and a length scale known the Rossby radius. More extensive discussions of these grids are included in Sections 6.2 and 6.3 and, for example, in Griffies [14] and Higdon [18].

For the class of DG methods, some often-quoted advantages include applicability to unstructured grids and the ability to attain high-order accuracy while maintaining high locality (Cockburn and Shu [7]). Some disadvantages, related to efficiency, include restrictive conditions on the maximum allowable time step for numerical stability (Kubatko, et al. [21]) and the need to compute multiple degrees of freedom for each dependent variable.

The present work is the first part of an effort to examine the applicability of DG methods to ocean circulation modeling and to assess whether they have a net advantage over the methods that are used currently. Due to the overall complexity of this matter, the scope of the present paper is limited to the following goals.

(1) Formulate the pressure forcing in the partial differential equations that describe the conservation of momentum. Here, it is assumed that the vertical coordinate is a generalized coordinate that includes level, terrain-fitted, isopycnic, and hybrid coordinates as special cases. When a generalized vertical coordinate is used, the pressure forcing is the sum of two terms, and in some circumstances the terms can have similar magnitude but opposite signs, so that their sum can be dominated by error. This representation can also be awkward for implementing in a weak form that is required for a Galerkin numerical method. Here, we go back to physical principles and derive, and then analyze, an integral weak form of the pressure forcing that avoids these difficulties.

(2) Analyze the accuracy of this approach by developing dispersion relations for the resulting DG spatial discretizations, and in particular compare the accuracy of such discretizations with finite difference approximations on the B-grid and C-grid. This methodology is then extended to give stability analyses of some time-stepping methods. In order to limit the complexity of the present paper, these analyses assume a reduced-dimension setting in which all flow variables are independent of one horizontal spatial variable. In this setting, the Coriolis parameter is nonzero and both components of velocity can be nonzero; in effect, we consider flow in an infinite straight channel in a rotating reference frame. The complexity is limited further by assuming linearized flow in a constant-density fluid. One conclusion is that the DG spatial discretizations can be much more accurate than the B- and C-grids. In particular, the DG formulation is not vulnerable to the problem of grid size versus Rossby radius for inertia-gravity waves that was mentioned above.

(3) Test the preceding ideas in some numerical computations. In one test problem described here, the higher spatial accuracy of the DG method can more than compensate for the restrictive bound on the time step and the need to compute multiple degrees of freedom. Another computation illustrates the well-balanced nature of the pressure forcing formulated here, in the constant-density case.

The DG method has been used extensively to solve the shallow water equa-

tions, which describe a single-layer (homogeneous) hydrostatic fluid (e.g., Giraldo and Warburton [12], Kubatko et al. [22], Nair et al. [25]). In addition, Kärnä et al. [20] have recently developed a DG model of three-dimensional coastal flows that uses a terrain-fitted vertical coordinate with a moving vertical mesh. Nair et al. [24] have developed a dynamical core for three-dimensional atmospheric circulation which uses a DG method on a cubed sphere for the horizontal discretization and a Lagrangian coordinate for the vertical discretization.

The purpose of the present paper is to develop a framework for a general vertical coordinate for ocean modeling and to perform the mathematical and computational analyses that are described above. An outline of this paper is the following.

In Section 2 we describe the equations for conservation of mass and momentum in terms of a generalized vertical coordinate, with two horizontal dimensions. In Section 3 we derive the weak forms of these equations for the reduced-dimension setting described above, with no assumption of linearity or constant density. This discussion emphasizes the formulation of the pressure term. The Appendix at the end of the paper gives the corresponding results for the general case of two horizontal orthogonal curvilinear coordinates on a rotating spheroid, with a generalized vertical coordinate.

Section 4 summarizes the remaining issues that will be discussed in the present paper, and it also gives an outline of other issues that are the subject of continuing work.

Section 5 describes the special case of a hydrostatic fluid of constant density, i.e., the shallow water equations. This discussion includes a detailed discussion of the pressure term, including Lax-Friedrichs interpolation to obtain pressure forcing at cell (element) edges, implementation with variable and discontinuous bottom topography, and a proof that the pressure forcing is well-balanced in this case.

Section 6 gives the analysis of numerical dispersion relations, and Section 7 gives analyses of some time-stepping methods. Some numerical computations are described in Section 8. Section 9 gives a summary.

2. Governing equations

The paper by Higdon [18] contains a derivation of the partial differential equations for conservation of mass, momentum, and tracers in a fluid that is in motion relative to a rotating spheroid. In that derivation the horizontal coordinates are arbitrary orthogonal curvilinear coordinates, and the vertical coordinate is a generalized coordinate, in a sense discussed in Section 2.1 below. Here we re-state the equations for conservation of mass and momentum. Curvilinear coordinates are used in the Appendix, but elsewhere in this paper the horizontal coordinates are taken to be rectangular Cartesian coordinates for the sake of notational simplicity.

In [18] it is assumed that the depth of the fluid is much smaller than the horizontal extent of the motions being studied. This shallow-water assumption

implies that the fluid is very nearly in hydrostatic balance, i.e., vertical accelerations are small, and this condition will be assumed throughout the present paper.

2.1. Vertical coordinate

The partial differential equations that describe three-dimensional oceanic flows include a vertical coordinate, and in the numerical modeling of ocean circulation several such coordinates are in use. These include the following.

(i) The elevation z. This choice is the most traditional.

(ii) A terrain-fitted coordinate σ. This quantity has constant values at the top and bottom of the fluid, with a continuous transition between the top and bottom, and it is well-suited for representing bottom topography.

(iii) An isopycnic coordinate, which is a quantity related to density. In the interior of the ocean, away from boundary layers, such a quantity is nearly constant along fluid trajectories. In this case, surfaces of constant vertical coordinate are nearly material surfaces, so a vertical discretization divides the fluid into water masses having distinct physical properties. This feature could be an advantage for long-time integrations, such as in climate modeling.

(iv) A hybrid coordinate, which may use z near the upper boundary, σ in near-shore regions, and an isopycnic coordinate in the ocean's interior (Bleck [6]).

Further discussions of these coordinates are given, for example, in [14] and [19]. In the following, the vertical coordinate will be denoted by s and will be regarded as a generalized vertical coordinate that could include any of the above possibilities. It will be assumed that s is non-decreasing function of z, at each horizontal location and time.

2.2. Conservation of mass

For a dependent variable in an equation for conservation of mass, it is commonplace to use the fluid density ρ. However, in the isopycnic case the density is essentially an independent variable, so ρ would not be a suitable dependent variable in that case.

For an alternative, let $p(x, y, s, t)$ denote the pressure in the fluid at horizontal position (x, y), vertical position s, and time t. For a mass variable, use the nonnegative quantity

$$-\frac{\partial p}{\partial s} = -p_s = |p_s|$$

(Bleck [6]). In the case where $s = z$, the hydrostatic condition implies $-p_s = -p_z = \rho g$, where g is the magnitude of the acceleration due to gravity. More generally, consider two coordinate surfaces defined by $s = s_0$ and $s = s_1$, where s_0 and s_1 are constants with $s_1 < s_0$. The size of $-p_s$ then indicates the amount of mass between those surfaces; in particular, $\int_{s_1}^{s_0} -p_s(x, y, s, t)\, ds = p(x, y, s_1, t) - p(x, y, s_0, t) = \Delta p(x, y, t)$. In the case of a hydrostatic fluid, $\Delta p(x, y, t)$ is the weight per unit horizontal area between the coordinate surfaces, or g times the mass per unit horizontal area.

The conservation of mass is described by the equation

$$\frac{\partial}{\partial t}\big(p_s\big) + \frac{\partial}{\partial x}\big(up_s\big) + \frac{\partial}{\partial y}\big(vp_s\big) + \frac{\partial}{\partial s}\big(\dot{s}p_s\big) \;=\; 0. \tag{1}$$

Here, $\dot{s} = \frac{Ds}{Dt}$ denotes the material derivative of s, i.e., the time derivative of s following fluid parcels, and $u(x, y, s, t)$ and $v(x, y, s, t)$ are the x- and y-components of fluid velocity, respectively. In the case where $s = z$, we have $\dot{s} = \dot{z} = \frac{Dz}{Dt} = w$ and $-p_s = -p_z = \rho g$, so equation (1) becomes $\rho_t + (\rho u)_x + (\rho v)_y + (\rho w)_z = 0$. In the case of an ideal isopycnic coordinate where $\dot{s} = 0$, there is no exchange of fluid between coordinate layers, and equation (1) reduces to a two-dimensional equation for conservation of mass within a coordinate layer.

2.3. Conservation of momentum

For the x- and y-components of momentum density, use the quantities $u(-p_s)$ and $v(-p_s)$, where u, v, and $-p_s$ all depend on (x, y, s, t). The x-component of the momentum equation is

$$\frac{\partial}{\partial t}\big[u(-p_s)\big] + F_u - f\big(v(-p_s)\big) \;= -(gz_s)\frac{\partial P}{\partial x} + g\frac{\partial \tau_u}{\partial s}. \tag{2}$$

Here, f is the Coriolis parameter, $z_s = \partial z/\partial s$ is the rate of change of elevation with respect to the generalized vertical coordinate s, $P(x, y, z, t)$ is the pressure (discussed below),

$$F_u = \frac{\partial}{\partial x}\Big[u\big(u(-p_s)\big)\Big] + \frac{\partial}{\partial y}\Big[v\big(u(-p_s)\big)\Big] + \frac{\partial}{\partial s}\Big[\dot{s}\big(u(-p_s)\big)\Big] \tag{3}$$

denotes the momentum advection terms, and τ_u is a shear stress of the form

$$\tau_u = \tau_u^{wb} + \rho A_D \frac{1}{z_s}\frac{\partial u}{\partial s},$$

where A_D is a vertical viscosity coefficient and τ_u^{wb} is the sum of the wind stress at the top of the fluid and frictional stress along the bottom. Similarly, the y-component of the momentum equation has the form

$$\frac{\partial}{\partial t}\big[v(-p_s)\big] + F_v + f\big(u(-p_s)\big) \;= -(gz_s)\frac{\partial P}{\partial y} + g\frac{\partial \tau_v}{\partial s}. \tag{4}$$

In the momentum equations (2) and (4), horizontal viscosity has been neglected. In a discontinuous Galerkin method, such a term can be represented with a "local" DG method, in which the gradient of velocity is an auxiliary variable (e.g., Dawson et al. [9]). This matter will not be included in the present paper.

2.4. An issue with the pressure term

In the above equations, the pressure P is expressed in terms of (x, y, z, t), so $\partial P/\partial x$ and $\partial P/\partial y$ represent derivatives with respect to x and y, respectively, for a fixed elevation z. However, z is not necessarily the vertical coordinate, and the pressure forcing needs to be expressed in terms of the coordinate that is actually used. The quantities $p(x, y, s, t)$ and $P(x, y, z, t)$ are related by $p(x, y, s, t) = P\big(x, y, z(x, y, s, t), t\big)$, where $z(x, y, s, t)$ is the elevation associated with vertical coordinate s. Then

$$\begin{aligned} & \frac{\partial p}{\partial x}(x, y, s, t) \\ = \; & \frac{\partial P}{\partial x}\big(x, y, z(x, y, s, t), t\big) + \frac{\partial P}{\partial z}\big(x, y, z(x, y, s, t), t\big)\frac{\partial z}{\partial x}(x, y, s, t) \\ = \; & \frac{\partial P}{\partial x} \; - \; \rho g \frac{\partial z}{\partial x}, \end{aligned}$$

so

$$\frac{\partial P}{\partial x} \; = \; \frac{\partial p}{\partial x} \; + \; \rho g \frac{\partial z}{\partial x}. \tag{5}$$

The quantity $\partial p/\partial x$ is a derivative for fixed s and thus represents a derivative along an s-coordinate surface that could slant. However, the pressure forcing requires a pressure gradient with respect to directions that are truly horizontal, and the term $\rho g \partial z/\partial x$ provides the necessary correction to $\partial p/\partial x$. Analogous remarks apply to $\partial P/\partial y$ and $\partial p/\partial y$.

Equation (5) includes a sum of terms that could have similar magnitudes but opposite signs; for example, consider a sloping coordinate surface when the free surface at the top of the fluid is level. When the derivatives in (5) are approximated numerically, the resulting approximation to (5) could then be dominated by errors (Adcroft et al. [1], Griffies [14]). Similar remarks apply if the horizontal pressure forcing is expressed in terms of the Montgomery potential, which has been used in isopycnic and hybrid-coordinate ocean modeling (Bleck [6]). The development of pressure forcing in Section 3.2 avoids this difficulty and, in addition, uses a representation that is natural for a weak form that can be used with a discontinuous Galerkin method.

3. Weak forms for a reduced-dimension case

Next we develop weak forms of the equations for conservation of momentum and mass, with an emphasis on the pressure forcing terms in the momentum equations. This is given for a simplified setting in the present section and for the general case in the Appendix.

For the present case, assume that the fluid is confined to an infinite straight channel aligned with the y-direction and that all of the velocity and mass variables are independent of y. This setting allows the possibility of nonzero flow in the y-direction, so the problem is not entirely one-dimensional in the horizontal directions. In particular, we assume that the Coriolis parameter f is

nonzero and constant. In this configuration, the Coriolis effect is present, but with the simplifying assumption that all quantities depend only on (x, s, t) instead of (x, y, s, t). The horizontal dependence in this system is thus quasi-one-dimensional. For notational simplicity, also assume $\dot{s} = Ds/Dt = 0$ so that the vertical advection terms are not carried through the calculations; this choice does not affect the formulas for the pressure forcing.

Let $u(x, s, t)$ and $v(x, s, t)$ denote the x- and y-components of fluid velocity, and let $p(x, s, t) = P\big(x, z(x, s, t), t\big)$ denote the pressure. In this case the x-component of the momentum equation is

$$\begin{aligned} &\frac{\partial}{\partial t}\big[u(-p_s)\big] + \frac{\partial}{\partial x}\Big[u\big(u(-p_s)\big)\Big] - f\big(v(-p_s)\big) \\ &= -gz_s \frac{\partial P}{\partial x}\big(x, z(x, s, t), t\big) + g\frac{\partial \tau_u}{\partial s}. \end{aligned} \tag{6}$$

The right side of equation (6) includes the function $\partial P/\partial x$ evaluated at $(x, z(x, s, t), t)$. This is *not* the derivative of a composite function, for which the chain rule would apply; instead, the notation $\partial P/\partial x$ refers to a derivative for fixed z, and this function is evaluated at the location $(x, z(x, s, t), t)$. In the development given below, we will not write this term in the form (5) described earlier, but will instead proceed directly to an integral formulation.

The y-component of the momentum equation is an analogue of equation (6), except that the pressure term is not present due to the assumption that all quantities are independent of y. In the present case, the mass equation (1) reduces to $\partial\big(p_s\big)/\partial t + \partial\big(up_s\big)/\partial x = 0$.

For a discretization of the fluid domain, partition vertically with coordinate surfaces defined by $s = s_0, s_1, \ldots, s_R$, with $s_0 > s_1 > \ldots > s_R$. Denote the horizontal interval by $a \le x \le b$, and partition the interval $[a, b]$ into grid cells of the form $D_j = [x_{j-1/2}, x_{j+1/2}]$ for $1 \le j \le J$.

In order to develop weak forms of the governing equations that would be suitable for usage in a Galerkin method, consider the volume of fluid associated with grid cell D_j in the horizontal and layer r in the vertical; the latter is the region between the coordinate surfaces $s = s_{r-1}$ and $s = s_r$. In terms of the vertical coordinate s, the volume of fluid is

$$\tilde{V}_{j,r} \;=\; D_j \times [s_r, s_{r-1}] \;=\; \{(x, s) : x \in D_j,\; s_r < s < s_{r-1}\};$$

in terms of the vertical coordinate z, this volume is

$$V_{j,r}(t) \;=\; \{(x, z) : x \in D_j,\; z_r(x, t) < z < z_{r-1}(x, t)\},$$

where $z_r(x, t) = z(x, s_r, t)$ and $z_{r-1}(x, t) = z(x, s_{r-1}, t)$ denote the elevations of the lower and upper boundaries, respectively, of layer r.

3.1. Weak form of the u-momentum equation

For a weak form of the u-momentum equation (6) on grid cell D_j in layer r, multiply (6) by an arbitrary smooth test function ψ and integrate on the region

$\tilde{V}_{j,r}$. For reasons associated with the pressure term, as described below, ψ is assumed to depend only on x and not on the vertical coordinate s. This process yields

$$\int_{D_j} \left\{ \frac{\partial}{\partial t} \left(\int_{s_r}^{s_{r-1}} u(-p_s)\, ds \right) + \frac{\partial}{\partial x} \left(\int_{s_r}^{s_{r-1}} uu(-p_s)\, ds \right) - f \int_{s_r}^{s_{r-1}} v(-p_s)\, ds \right\} \psi(x)\, dx$$
$$= \Pi_u(j,r,\psi) + g \int_{D_j} \Big\{ \tau_u(x,s_{r-1},t) - \tau_u(x,s_r,t) \Big\} \psi(x)\, dx. \qquad (7)$$

Here, $\Pi_u(j,r,\psi)$ is the pressure term that is discussed below.

Let

$$\Delta p_r(x,t) = \int_{s_r}^{s_{r-1}} (-p_s)\, ds = p(x,s_r,t) - p(x,s_{r-1},t)$$

denote the vertical pressure increment across layer r, i.e., g times the mass per unit horizontal area in that layer. Also let

$$u_r(x,t) = \frac{1}{\Delta p_r} \int_{s_r}^{s_{r-1}} u(-p_s)\, ds$$

denote the mass-weighted vertical average of u in layer r. A mass-weighted vertical average $v_r(x,t)$ of v is defined similarly. A calculation shows $\int_{s_r}^{s_{r-1}} uu(-p_s)\, ds = u_r(u_r \Delta p_r) + \mathcal{O}(\Delta s)^3$; this term represents a lateral flux of the momentum density $u_r \Delta p_r$ by the velocity u_r. If the error $\mathcal{O}(\Delta s)^3$ is deleted, and if an integration by parts is performed on the flux term, then equation (7) can be written as

$$\int_{D_j} \left\{ \frac{\partial}{\partial t} (u_r \Delta p_r) - f v_r \Delta p_r \right\} \psi(x)\, dx$$
$$+ \Big[u_r(u_r \Delta p_r) \psi(x) \Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} - \int_{D_j} u_r(u_r \Delta p_r) \psi'(x)\, dx$$
$$= \Pi_u(j,r,\psi) + g \int_{D_j} \Big\{ (\tau_u)_{r-1}(x,t) - (\tau_u)_r(x,t) \Big\} \psi(x)\, dx. \qquad (8)$$

Here, $(\tau_u)_{r-1}(x,t) = \tau_u(x,s_{r-1},t)$ and $(\tau_u)_r(x,t) = \tau_u(x,s_r,t)$ denote the shear stresses at the top and bottom of layer r, respectively. Equation (8) is the weak form of the u-momentum equation, for the reduced-dimension case considered here.

3.2. The pressure term $\Pi_u(j,r,\psi)$

The pressure term $\Pi_u(j,r,\psi)$ in equation (8) is obtained by multiplying the pressure term on the right side of equation (6) by the test function ψ and then integrating over the region $\tilde{V}_{j,r}$ in (x,s) space. Some calculations yield

$$\Pi_u(j,r,\psi) = -\int_{D_j} \left[\int_{s_r}^{s_{r-1}} \frac{\partial P}{\partial x}\Big(x, z(x,s,t), t\Big)\, g z_s\, ds \right] \psi(x)\, dx$$

$$
\begin{aligned}
&= -g \int_{D_j} \left[\int_{z_r(x,t)}^{z_{r-1}(x,t)} \frac{\partial P}{\partial x}(x,z,t)\ dz \right] \psi(x)\ dx \\
&= -g \int_{V_{j,r}(t)} \frac{\partial P}{\partial x}(x,z,t)\ \psi(x)\ dz\, dx \\
&= -g \int_{V_{j,r}(t)} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right) \cdot \Big(P,\ 0 \Big)\ \psi(x)\ dz\, dx \\
&= -g \int_{\partial V_{j,r}(t)} \Big(P,\ 0 \Big) \cdot \mathbf{n}\ \psi(x)\ dS \\
&\qquad\qquad + g \int_{V_{j,r}(t)} \Big(P,\ 0 \Big) \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right) \psi(x)\ dz\, dx. \qquad (9)
\end{aligned}
$$

In the first line, the function $\partial P/\partial x$ is evaluated at $(x, z(x,s,t), t)$, and the next line is obtained by a change of variable in the integral over s. In the last line, $\partial V_{j,r}(t)$ denotes the boundary of the region $V_{j,r}(t)$ in (x,z) space, and $\mathbf{n}$ is the outward unit normal vector along $\partial V_{j,r}(t)$.

If the test function ψ were to depend on both x and s, then the factor $\psi(x,s)$ would be placed inside the inner integral in the first line of (9). Changing to an integration with respect to z, which appears in the second line, would require that s be determined as a function of z over that interval. This would introduce additional complexity to the algorithm, especially in the case of an isopycnic or hybrid vertical coordinate. On the other hand, if the vertical coordinate is z or the terrain-fitted coordinate σ, which has a relatively simple relation to z, then this step might be more feasible and could lead to higher-order discretizations in the vertical dimension. However, this possibility will not be pursued in the present paper.

In the last line of equation (9), an integral of P with respect to z is contained in the boundary terms for the left and right boundaries and also in the integral on the interior of $V_{j,r}(t)$. For the sake of subsequent formulas, denote the vertical integral of the horizontal pressure forcing over layer r by

$$
\begin{aligned}
H_r(x,t) &= g \int_{z_r(x,t)}^{z_{r-1}(x,t)} P(x,z,t)\ dz \qquad (10) \\
&= \int_{p_{r-1}(x,t)}^{p_r(x,t)} \alpha\, p\ dp, \qquad (11)
\end{aligned}
$$

where $\alpha = 1/\rho$ is the specific volume (volume per unit mass). The form in (10) is motivated directly by the structure of the last line in (9). In equation (11), $p_{r-1}(x,t)$ and $p_r(x,t)$ denote the pressures at the top and bottom of layer r, i.e., $p_r(x,t) = P(x, z_r(x,t), t) = p(x, s_r, t)$. The derivation of (11) from (10) uses the hydrostatic condition $dP/dz = -\rho g = -g/\alpha$.

On the left and right boundaries of the region $\partial V_{j,r}(t)$, the unit normal vectors are $\mathbf{n} = (-1, 0)$ and $\mathbf{n} = (1, 0)$, respectively. On the graph of a function ϕ, the upward unit normal vector at horizontal position x is $(-\phi'(x), 1)/((\phi')^2 +$

$1)^{1/2}$, and the element of arclength is $((\phi')^2 + 1)^{1/2} dx$. The representation of $\Pi_u(j, r, \psi)$ in equation (9) can then be written as

$$\begin{aligned} \Pi_u(j,r,\psi) &= -\Big[H_r(x,t)\ \psi(x)\Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} + \int_{D_j} H_r(x,t)\ \psi'(x)\ dx \\ &+ g\int_{D_j} \left\{ p_{r-1}(x,t)\frac{\partial z_{r-1}}{\partial x} - p_r(x,t)\frac{\partial z_r}{\partial x} \right\} \psi(x)\ dx. \end{aligned} \tag{12}$$

In Section 2.4 it was noted that the expression (5) for pressure forcing includes a sum of terms that could have opposite signs, and when the derivatives in those terms are approximated numerically the sum could be dominated by errors. In the representation (12) for $\Pi_u(j, r, \psi)$, the boundary terms at $x_{j\pm 1/2}$ also involve a subtraction of terms that could be nearly equal. However, the representation (11) of $H_r(x, t)$ can be evaluated exactly (up to roundoff error) for a fluid of constant density and for a stack of layers of constant densities. More generally, the evaluation of (11) could require a highly-accurate quadrature involving an equation of state that relates pressure and density; however, such a computation would be completely local in the horizontal dimension, unlike higher-order difference approximations to derivatives.

3.3. Pointwise form and alternate derivation

An alternative to the preceding derivation is to integrate with respect to s over layer r to obtain an equation that is pointwise in x, and then multiply by a test function and integrate by parts. This approach requires fundamentally that the test function ψ depend only on x and not on s, whereas, in principle, the preceding approach could be extended to use $\psi(x, s)$ and obtain higher-order approximations in the vertical direction.

If the continuous u-momentum equation (6) is integrated over the interval $s_r < s < s_{r-1}$, the result is

$$\begin{aligned} &\frac{\partial}{\partial t}\left(u_r \Delta p_r\right) + \frac{\partial}{\partial x}\Big[u_r\left(u_r \Delta p_r\right)\Big] - f v_r \Delta p_r \\ &= -g\int_{s_r}^{s_{r-1}} \frac{\partial P}{\partial x}\Big(x, z(x,s,t), t\Big)\ z_s\ ds \\ &+ g\Big[(\tau_u)_{r-1}(x,t) - (\tau_u)_r(x,t)\Big], \end{aligned} \tag{13}$$

with an error term $\mathcal{O}(\Delta s)^3$ deleted. Now express the integral on the right side as an integral over z and compare with the representation of $H_r(x, t)$ in (10) to calculate

$$\begin{aligned} \frac{\partial H_r}{\partial x} &= g\int_{z_r(x,t)}^{z_{r-1}(x,t)} \frac{\partial P}{\partial x}\ dz \\ &+ gP(x, z_{r-1}(x,t), t)\frac{\partial z_{r-1}}{\partial x} - gP(x, z_r(x,t), t)\frac{\partial z_r}{\partial x}. \end{aligned} \tag{14}$$

Equation (13) can then be written as

$$\begin{aligned}\frac{\partial}{\partial t}\left(u_r\Delta p_r\right) &+ \frac{\partial}{\partial x}\Big[u_r\left(u_r\Delta p_r\right)\Big] - f v_r \Delta p_r \\ = -\frac{\partial H_r}{\partial x} &+ g\left[p_{r-1}(x,t)\frac{\partial z_{r-1}}{\partial x} - p_r(x,t)\frac{\partial z_r}{\partial x}\right] \\ &+ g\Big[(\tau_u)_{r-1}(x,t) - (\tau_u)_r(x,t)\Big]. \end{aligned} \tag{15}$$

This is a pointwise form of the u-momentum equation, and from this form the weak form in (8) and (12) can be derived. The preceding derivation resembles some calculations used by Adcroft et al. [1] during a finite-volume development that addresses some problems related to the effects of compressibility on the horizontal pressure forcing.

3.4. The mass and v-momentum equations

For the reduced-dimension case considered in the present section, the mass equation (1) is $\partial\big(p_s\big)/\partial t + \partial\big(up_s\big)/\partial x = 0$. The vertically-integrated pointwise form of this equation is

$$\frac{\partial}{\partial t}\left(\Delta p_r\right) + \frac{\partial}{\partial x}\left(u_r\Delta p_r\right) = 0, \tag{16}$$

and the weak form is

$$\begin{aligned}\int_{D_j}\frac{\partial}{\partial t}\left(\Delta p_r\right)\psi(x)\;dx &+ \Big[u_r(\Delta p_r)\psi(x)\Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} \\ &- \int_{D_j} u_r(\Delta p_r)\psi'(x)\;dx = 0. \end{aligned} \tag{17}$$

The weak form of the v-momentum equation is an analogue of the u-momentum equation (8), except that there is no pressure term, and it can be written as

$$\begin{aligned}&\int_{D_j}\left\{\frac{\partial}{\partial t}\left(v_r\Delta p_r\right) + f u_r\Delta p_r\right\}\;\psi(x)\;dx \\ &+ \Big[u_r(v_r\Delta p_r)\psi(x)\Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} - \int_{D_j} u_r(v_r\Delta p_r)\psi'(x)\;dx \\ &= g\int_{D_j}\Big\{(\tau_v)_{r-1}(x,t) - (\tau_v)_r(x,t)\Big\}\psi(x)\;dx. \end{aligned} \tag{18}$$

4. Plan of analysis

Numerous issues need to be addressed in order to incorporate the preceding ideas about pressure forcing into a working DG algorithm for three-dimensional ocean modeling, and, more generally, to assess DG methods for possible usage in this area.

4.1. Present paper

The remainder of the present paper begins an analysis of these issues. In order to limit complexity and length, the remaining discussion (except for the Appendix) is restricted to the special case of a hydrostatic fluid of constant density in the reduced-dimension setting described above, in which the horizontal dependence is quasi-one-dimensional. This discussion includes the following.

(1) Representation of the pressure forcing at the edges of grid cells, where the dependent variables may be discontinuous.

(2) Implementation of the pressure forcing with variable and possibly discontinuous bottom topography.

(3) Verification that this formulation of the pressure forcing is well-balanced.

(4) For the linearized case, analysis of numerical dispersion relations and comparison to some classical staggered finite-difference grids. Due to the assumption of linearity and the restriction on the spatial dependence, this analysis is restricted to inertia-gravity waves.

(5) For the linearized case, analysis of stability of some time-stepping methods.

(6) Numerical experiments that illustrate the preceding points.

4.2. Continuing work

Some other issues are beyond the scope of one paper and are the subject of continuing work. These include the following.

(i) Extend the analysis, implementation, and testing from the constant-density case mentioned above to a stratified and hydrostatic fluid having variable density.

(ii) Evaluate the choice and implementation of different vertical coordinates.

(iii) Extend the analysis, implementation, and testing to two horizontal dimensions, and use numerical experiments that include Rossby waves and more realistic configurations. The following discussion of the reduced-dimension case includes algorithms that are applied at endpoints of one-dimensional grid cells (i.e., intervals). For the case of two horizontal dimensions, it may suffice to apply these ideas pointwise at each point on a cell edge and then integrate over the edge.

5. The case of constant density: the shallow water equations

In Section 2 it was assumed that the depth of the fluid is much smaller than the horizontal length scales of the motions being studied, and thus the fluid is hydrostatic. For the sake of limiting the complexity of the analysis in the remainder of this paper, now also assume that the density of the fluid is constant. Together, these assumptions lead to the shallow water equations (Gill [11]).

The shallow water equations are of interest in their own right, such as, for example, in modeling storm surges (Dawson et al. [9]) and the propagation of tsunamis (LeVeque et al. [23]). These equations are also of interest in the more

general case of three-dimensional modeling. In the case of an isopycnic coordinate, a vertically-discrete three-dimensional model can be regarded as a stack of two-dimensional shallow water models, with various means for communicating between layers. In addition, for three-dimensional ocean circulation modeling it is common practice to split the fast and slow dynamics into separate subproblems, with the fast motions being modeled by a vertically-integrated two-dimensional system that closely resembles the shallow water equations (Higdon [18]).

5.1. A representation of the shallow water equations

We first develop a representation of the shallow water equations which is an analogue of the formulations developed in Section 3. It that setting, it was assumed that the horizontal dependence is quasi-one-dimensional, in the sense of flow in an infinite straight channel in a rotating reference frame. For the present case of constant density, a vertical coordinate is not needed as an independent variable, so in this case the overall system is quasi-one-dimensional.

To apply the derivations from Section 3 to the present case of constant density, regard the top and bottom of the fluid domain as coordinate surfaces, which would be the case for a σ-coordinate representation. Vertical integration between those two surfaces then gives equations that describe the entire fluid.

Instead of using the notation Δp_r to refer to the mass variable, let $p_b(x,t)$ denote the difference between the pressure at the bottom of the fluid and the atmospheric pressure p_0 at the top of the fluid. The quantity $p_b(x,t)$ is the weight of the water column per unit horizontal area, i.e., g times the mass per unit horizontal area, and $p_0 + p_b(x,t)$ is the total bottom pressure. Also let $u(x,t)$ and $v(x,t)$ denote the x- and y-components of fluid velocity, respectively. According to the analysis in Section 3.1, these would be mass-weighted vertical averages over the water column. However, it can be shown (e.g., [19]) that the horizontal components of fluid velocity are actually independent of vertical position if this is the case at some initial time, for the present case of a hydrostatic fluid of constant density.

In this notation, the pointwise form (15) of the u-component of the momentum equation can be expressed as

$$\begin{aligned} \frac{\partial}{\partial t}\left(p_b u\right) &+ \frac{\partial}{\partial x}\left[u\left(p_b u\right)\right] - f p_b v \\ = -\frac{\partial H}{\partial x} &+ g\left[p_0 \frac{\partial z_{top}}{\partial x} - \left(p_0 + p_b(x,t)\right)\frac{\partial z_{bot}}{\partial x}\right] \\ &+ g\Big[(\tau_u)^{wind}(x,t) - (\tau_u)^{bot}(x,t)\Big]. \end{aligned} \tag{19}$$

Here, $p_b u$ equals g times the u-component of momentum per unit horizontal area, $z_{bot}(x)$ is the elevation of the bottom topography, $z_{top}(x,t) = z_{bot}(x) + p_b(x,t)/(\rho g)$ is the elevation of the free surface at the top of the fluid, $(\tau_u)^{wind}(x,t)$ denotes wind stress at the top of the fluid, $(\tau_u)^{bot}(x,t)$ denotes frictional stress

along the bottom of the fluid, and

$$\begin{aligned} H(x,t) &= g\int_{z_{bot}(x)}^{z_{top}(x,t)} P(x,z,t)\ dz \\ &= \int_{p_0}^{p_0+p_b(x,t)} \alpha\, p\ dp \\ &= \frac{\alpha}{2}\Big[\big(p_0+p_b(x,t)\big)^2 - p_0^2\Big] \end{aligned} \tag{20}$$

is the vertically-integrated horizontal pressure forcing. The corresponding weak form (8) of the u-momentum equation is

$$\begin{aligned} &\int_{D_j}\left\{\frac{\partial}{\partial t}\left(p_b u\right) - f p_b v\right\}\ \psi(x)\ dx \\ &+\ \Big[u(p_b u)\psi(x)\Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} \ - \int_{D_j} u(p_b u)\psi'(x)\ dx \\ &=\ \Pi_u(j,\psi)\ +\ g\int_{D_j}\Big\{(\tau_u)^{wind}(x,t) - (\tau_u)^{bot}(x,t)\Big\}\psi(x)\ dx, \end{aligned} \tag{21}$$

where

$$\begin{aligned} \Pi_u(j,\psi) &= -\Big[H(x,t)\ \psi(x)\Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} \ +\ \int_{D_j} H(x,t)\ \psi'(x)\ dx \\ &+\ g\int_{D_j}\left\{p_0\frac{\partial z_{top}}{\partial x}\ -\ (p_0+p_b(x,t))\frac{\partial z_{bot}}{\partial x}\right\}\ \psi(x)\ dx. \end{aligned} \tag{22}$$

If the atmospheric pressure p_0 is constant, then $\partial P/\partial x = \partial/\partial x(P-p_0)$. The preceding derivations can then be performed with $P-p_0$ replacing P, and the final results are the same, with $p-p_0$ replacing p. The effect on the formulas in (19), (20), and (22) is to set $p_0 = 0$. From now on, it will be assumed that the atmospheric pressure p_0 is constant; without further loss of generality, p_0 can then be deleted from the representation of the pressure forcing. Note that this step does not state that the atmospheric pressure is actually zero; instead, it says that this pressure does not affect the preceding representations of the lateral pressure forcing within the fluid, and one can use $p_0 = 0$ in those formulas.

The v-component of the momentum equation is analogous to the u-component, except that the pressure term is absent in the case considered here. The pointwise form of this component is

$$\frac{\partial}{\partial t}\left(p_b v\right) + \frac{\partial}{\partial x}\big[u\left(p_b v\right)\big] + f p_b u\ =\ g\Big[(\tau_v)^{wind}(x,t) - (\tau_v)^{bot}(x,t)\Big], \tag{23}$$

and the weak form is

$$\int_{D_j}\left\{\frac{\partial}{\partial t}\left(p_b v\right) + f p_b u\right\}\ \psi(x)\ dx$$

$$+ \left[u(p_b v)\psi(x)\right]_{x=x_{j-1/2}}^{x=x_{j+1/2}} - \int_{D_j} u(p_b v)\psi'(x)\, dx$$
$$= g \int_{D_j} \Big\{(\tau_v)^{wind}(x,t) - (\tau_v)^{bot}(x,t)\Big\}\psi(x)\, dx. \tag{24}$$

The equation for conservation of mass is

$$\frac{\partial p_b}{\partial t} + \frac{\partial}{\partial x}(p_b u) = 0, \tag{25}$$

and the corresponding weak form is

$$\int_{D_j} \frac{\partial p_b}{\partial t}\psi(x)\, dx + \left[(p_b u)\psi(x)\right]_{x=x_{j-1/2}}^{x=x_{j+1/2}} - \int_{D_j} (p_b u)\psi'(x)\, dx = 0. \tag{26}$$

5.2. Comparison with previous representations of the shallow water equations and pressure forcing

The preceding representation of the shallow water equations is not the standard one that is found in many references. This is ultimately a result of (i) using pressure in Sections 2 and 3 to define a mass variable, which results from the usage of an arbitrary vertical coordinate in that discussion, and (ii) the decision in Section 3 to represent the pressure forcing with a horizontal gradient of pressure within the fluid and then multiply by a test function and integrate by parts. In the present subsection we derive the standard representation, for the quasi-one-dimensional case, partly for later usage and partly in order to compare to the more standard representation of the pressure forcing for the shallow water equations.

First, let $h(x,t) = z_{top}(x,t) - z_{bot}(x) = p_b(x,t)/(\rho g) = \alpha p_b(x,t)/g$ denote the linear thickness of the fluid layer. Then h is the volume of the water column per unit horizontal area, and the relation $h = p_b/(\rho g)$ implies that the mass equation (25) is equivalent to

$$h_t + (hu)_x = 0. \tag{27}$$

In the pointwise form (19) of the u-momentum equation, the term $-\partial H/\partial x$ can be expressed as a derivative of the integral representation in the first line of (20). The derivative of that expression produces terms that involve the derivatives of the limits of integration, and these terms cancel the terms in (19) that involve z_{top} and z_{bot}. Equation (19) then implies

$$\frac{\partial}{\partial t}(p_b u) + \frac{\partial}{\partial x}\left[u\,(p_b u)\right] - f p_b v = -g \int_{z_{bot}(x)}^{z_{top}(x,t)} \frac{\partial P}{\partial x}(x,z,t)\, dz, \tag{28}$$

where in this case the wind stress and bottom stress terms are deleted for notational simplicity. The hydrostatic assumption can be expressed as $\partial P/\partial z = -\rho g$, so $P(x,z,t) = \rho g(z_{top}(x,t) - z)$, and thus the right side of equation (28) equals

$$\begin{aligned} -g\int_{z_{bot}(x)}^{z_{top}(x,t)} \rho g \frac{\partial z_{top}}{\partial x}(x,t)\; dz &= -g\rho g\Big(z_{top}(x,t) - z_{bot}(x)\Big)\frac{\partial z_{top}}{\partial x} \\ &= -gp_b\frac{\partial z_{top}}{\partial x}. \end{aligned}$$

Substitute this result into equation (28) and use the relation $p_b(x,t) = \rho g h(x,t)$ to obtain

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}\big[u\,(hu)\big] - fhv = -gh\frac{\partial z_{top}}{\partial x}. \tag{29}$$

As suggested by equation (27), the thickness h of the fluid layer can serve as the mass variable in the system. The extra variable z_{top} can be eliminated by using the relation $z_{top}(x,t) = z_{bot}(x) + h(x,t)$ to obtain

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}\left[u\,(hu) + \frac{1}{2}gh^2\right] - fhv = -gh\frac{\partial z_{bot}}{\partial x}. \tag{30}$$

The right side of equation (30) is a source term that arises from variations in the elevation of the bottom topography. A numerical discretization of this term should be chosen so that it does not produce spurious forcing due to errors in the discretization. For this purpose, a criterion that is widely used is that a numerical method should be *well-balanced*. That is, in a region where the free surface is level, the computed pressure forcing should be zero, regardless of the nature of the bottom topography. This problem has been the subject of a great deal of recent research; see, for example, Le Veque et al. [23] and Xing et al. [28] and the papers referenced therein.

The alternative developed in the present paper is to proceed directly from the continuous momentum equation (6), in which the pressure forcing is expressed with a gradient within the fluid in a direction that is truly horizontal, regardless of the nature of the generalized vertical coordinate s. The first step is to use an integral over a fluid volume defined by a grid cell in the horizontal and a coordinate layer in the vertical. (In the special case of the constant-density shallow water equations, the "coordinate layer" is the entire water column.) The resulting representation of pressure forcing is natural for usage in a discontinuous Galerkin numerical method, and it also leads naturally to well-balanced forcing, provided that the pressures at cell edges are defined appropriately. The latter points are addressed in Sections 5.3 and 5.4.

5.3. Numerical implementation of pressure in a discontinuous Galerkin method; Lax-Friedrichs interpolation

A discontinuous Galerkin method is based on the weak forms of the governing equations and is developed as follows. For each grid cell, choose a basis for

the space of polynomials of a specified degree. Express each of the dependent variables as a linear combination of these basis functions, with coefficients (degrees of freedom) that depend on t. In the weak form, let the test function ψ be each of the basis functions, and thereby produce a system of ordinary differential equations for the degrees of freedom for each of the dependent variables. This process is discussed in greater detail in Sections 6.1 and 8.1.

In the present subsection we discuss the implementation of the pressure term $\Pi_u(j, \psi)$ in (22). In each grid cell and for each time t, the mass variable $p_b(x, t)$ is represented with a polynomial in x, with different polynomials being used in different cells. The vertically-integrated horizontal pressure forcing $H(x, t)$ in (20) is then a polynomial in x in each cell, for each t. As part of the configuration of the problem, assume that the bottom topography $z_{bot}(x)$ is represented with a polynomial in each cell; the free-surface elevation $z_{top}(x, t) = z_{bot}(x) + p_b(x, t)/(\rho g)$ is then also a polynomial. In the integrals on the cell D_j that appear in (22), the integrands are polynomials, so the integrals can be computed essentially exactly, assuming a sufficient number of quadrature points.

There remains the matter of defining suitable values of $H(x, t)$ when x is a cell edge, for the sake of the boundary terms in (22). At a cell edge $x_{j-1/2}$, p_b may be discontinuous. Thus the left and right limits of $H(x, t)$ as $x \to x_{j-1/2}$ may be unequal, so there is some ambiguity about a value of $H(x_{j-1/2}, t)$. To express the problem in physical terms, if the right limit $H(x^+_{j-1/2}, t)$ is used in the pressure forcing in cell D_j and the left limit $H(x^-_{j-1/2}, t)$ is used in cell D_{j-1}, then the force exerted by cell D_j on cell D_{j-1} would not equal the force exerted by cell D_{j-1} on cell D_j. That is, the numerical method would not satisfy Newton's third law of motion.

The preceding is the same basic difficulty that is encountered when one defines mass and momentum fluxes at cell edges for finite volume and discontinuous Galerkin methods. In fact, in the formula for $\Pi_u(j, \psi)$ in (22), the quantity $H(x, t)$ plays the structural role of a flux, and H could be regarded as a flux of momentum due to pressure forcing.

In order to define an edge value of $H(x_{j-1/2}, t)$ that is suitable for usage in (22), we will use a Lax-Friedrichs interpolation of the left and right limits of $H(x, t)$. This choice is motivated by the solution of a Riemann problem for the linearized shallow water equations, as described below, in Section 5.6.

Before this interpolation is described, we address a technical point that arises if one allows the elevation z_{bot} of the bottom topography to be discontinuous across cell edges. At cell edge $x_{j-1/2}$, let

$$z_{j-1/2} = \max\bigl\{ z_{bot}(x^-_{j-1/2}), z_{bot}(x^+_{j-1/2}) \bigr\} \tag{31}$$

denote the greater of the elevations on either side of that edge. At locations above $z_{j-1/2}$, the fluid masses in cells D_{j-1} and D_j are in contact with each other, and the vertically-integrated horizontal pressure forcing over that range of elevations should be interpolated between the two cells. However, if z_{bot} is discontinuous at $x_{j-1/2}$, then the deeper of the two cells is also in contact with a

solid vertical wall that represents the transition in bottom topography between the two elevations at cell edge $x_{j-1/2}$. That portion of the pressure forcing will be represented separately.

The usage of the elevation $z_{j-1/2}$ in the following discussions resembles some procedures used, for example, by Higdon [17] and Xing et al. [28], to compute mass and momentum fluxes at cell edges.

Denote the left and right limits of the pressures at edge $x_{j-1/2}$ at elevation $z_{j-1/2}$ by

$$\begin{aligned} (p_b)^-_{j-1/2}(t) &= p_b(x^-_{j-1/2}) - \rho g\Big(z_{j-1/2} - z_{bot}(x^-_{j-1/2})\Big) \\ (p_b)^+_{j-1/2}(t) &= p_b(x^+_{j-1/2}) - \rho g\Big(z_{j-1/2} - z_{bot}(x^+_{j-1/2})\Big). \end{aligned} \tag{32}$$

(Strictly speaking, the actual pressures in the fluid are equal to these values plus the atmospheric pressure p_0. However, as noted after equation (22), one can use $p_0 = 0$ in the formulas for the pressure forcing, and the representations in (32) will be used solely for that purpose.) In (32), the terms involving $\rho g = g/\alpha$ are hydrostatic adjustments to the bottom values $p_b(x^\pm_{j-1/2})$. If z_{bot} is discontinuous at $x_{j-1/2}$, then one of those adjustments is zero; if z_{bot} is continuous at $x_{j-1/2}$, then both of those adjustments are zero.

Now define corresponding values of the vertically-integrated horizontal pressure forcing, over the range where the two fluid masses are in contact, by

$$H^\pm_{j-1/2}(t) = \int_0^{(p_b)^\pm_{j-1/2}(t)} \alpha\, p\; dp. \tag{33}$$

If the elevation of the free surface at the top of the fluid is continuous at $x_{j-1/2}$, then $H^-_{j-1/2}(t) = H^+_{j-1/2}(t)$; otherwise, these two quantities are unequal. The case of inequality can arise even if the bottom topography is continuous, so one therefore needs to interpolate between the two states represented in (33). The Lax-Friedrichs interpolation, as motivated in Section 5.6, is

$$\begin{aligned} H^{LF}_{j-1/2}(t) &= \frac{1}{2}\left[H^-_{j-1/2}(t) \;+\; H^+_{j-1/2}(t)\right] \\ &+ \frac{c}{2}\left[(p_b u)^-_{j-1/2}(t) \;-\; (p_b u)^+_{j-1/2}(t)\right]. \end{aligned} \tag{34}$$

Here, the terms

$$(p_b u)^\pm_{j-1/2}(t) \;=\; (p_b)^\pm_{j-1/2}(t)\; u(x^\pm_{j-1/2}, t) \tag{35}$$

represent the one-sided limits of the momentum density (times g) at the cell edge, for the portion of the fluid that lies above the edge elevation $z_{j-1/2}$; the interpolation in (34) concerns only that portion of the fluid. The quantity c in (34) is a representative value of the speed of gravity waves in the nondispersive limit, which we will take here to be $\sqrt{g h_{j-1/2}}$ (see Section 5.5), where $h_{j-1/2}$ is the depth of the bottom topography as implied by (31), i.e., $h_{j-1/2}$ is the equilibrium elevation of the free surface minus $z_{j-1/2}$.

The quantity $H^{LF}_{j-1/2}(t)$ in (34) represents a common value for the vertically-integrated horizontal pressure forcing exerted by the fluid in cell D_{j-1} on the fluid in cell D_j, and vice-versa, over the vertical range on which these fluid masses are in direct contact. However, if the bottom topography is discontinuous at $x_{j-1/2}$, then the deeper of the two cells also exerts a force on a vertical wall at $x_{j-1/2}$, and thus the wall exerts a force on the fluid in that cell. This effect should also be included in the pressure term.

To that end, let

$$\left(H^{topog}_{j-1/2}\right)^{\pm}(t) = \int_{(p_b)^{\pm}_{j-1/2}(t)}^{p_b(x^{\pm}_{j-1/2},t)} \alpha\, p\; dp. \tag{36}$$

The lower limit $(p_b)^{\pm}_{j-1/2}(t)$ is defined in (32) and represents a pressure at the edge elevation $z_{j-1/2}$. The upper limit $p_b(x^{\pm}_{j-1/2}, t)$ is a one-sided limit of the pressure at the bottom of the fluid within a cell. If the bottom topography is discontinuous at $x_{j-1/2}$, then one of the terms $\left(H^{topog}_{j-1/2}\right)^{-}(t)$ and $\left(H^{topog}_{j-1/2}\right)^{+}(t)$ is positive and the other is zero; if the bottom topography is continuous at $x_{j-1/2}$ then both terms are zero.

For the values of $H(x,t)$ that appear in the boundary terms in the pressure term $\Pi_u(j,\psi)$ in (22), one can then use

$$H^{LF}_{j-1/2}(t) \;+\; \left(H^{topog}_{j-1/2}\right)^{+}(t) \tag{37}$$

at edge $x_{j-1/2}$ and

$$H^{LF}_{j+1/2}(t) \;+\; \left(H^{topog}_{j+1/2}\right)^{-}(t) \tag{38}$$

at edge $x_{j+1/2}$. In each of these formulas, the first term represents the force exerted on the fluid in cell D_j by the fluid in a neighboring cell, and the second represents the force exerted by a vertical wall along the bottom topography (if any).

5.4. Well-balanced forcing

A useful check on a numerical representation of pressure forcing is to verify whether the forcing is "well-balanced" in the following sense. Suppose that, in a neighborhood of cell D_j, the elevation of the free surface is constant and the fluid velocity is constant. This circumstance could be found in the global rest state for which the entire system is at rest, or it could be found in a local steady state with varying activity elsewhere. The numerical representation of the pressure forcing for cell D_j should then be zero, at least up to roundoff error and errors in numerical quadrature. (Also see Section 5.2.)

Theorem 1. *The pressure forcing term $\Pi_u(j,\psi)$ in (22) is well-balanced, in the sense described above.*

PROOF. Assume that, in a neighborhood of cell D_j, the elevation of the free surface is constant. In that case, the limits of integration $(p_b)^-_{j-1/2}(t)$ and $(p_b)^+_{j-1/2}(t)$ in (33) are the same, and thus $H^-_{j-1/2}(t) = H^+_{j-1/2}(t)$. If, in addition, u is constant in a neighborhood of D_j, then the Lax-Friedrichs interpolation in (34) reduces to

$$H^{LF}_{j-1/2}(t) \;=\; H^-_{j-1/2}(t) \;=\; H^+_{j-1/2}(t). \tag{39}$$

(In physical terms, the numerical representation $H^{LF}_{j-1/2}(t)$ of the pressure force exerted between the fluid masses in cells D_{j-1} and D_j, over the vertical range where the masses are in direct contact, is equal to the representations that are obtained within each cell.)

In the representation (22) of $\Pi_u(j,\psi)$, the value of $H(x,t)$ that is used in the boundary term at $x_{j-1/2}$ is given by the sum in (37). In the present circumstances, this sum equals

$$\lim_{x \to x^+_{j-1/2}} H(x,t) = \lim_{x \to x^+_{j-1/2}} \int_0^{p_b(x,t)} \alpha\, p \; dp, \tag{40}$$

i.e., the limit of $H(x,t)$ as $x \to x_{j-1/2}$ from within cell D_j. Similarly, the value of $H(x,t)$ that is used in the boundary term at $x_{j+1/2}$ is given by the sum in (38), and this equals the limit of $H(x,t)$ as $x \to x_{j+1/2}$ from within cell D_j.

A comparison with Section 3.2 reveals that the representation of $\Pi_u(j,\psi)$ in (22) is obtained by integrating $-g\psi(x)\partial P/\partial x$ over the entire water column on cell D_j. Denote this region of fluid by V_j. In a numerical implementation, this representation is modified by using interpolated values of $H(x,t)$ at the cell edges $x_{j\pm1/2}$. However, in the present case, those interpolated values are the one-sided limits of the interior values of $H(x,t)$. Thus

$$\Pi_u(j,\psi) = -g \int_{V_j} \frac{\partial P}{\partial x}(x,z,t)\; \psi(x)\; dz\, dx. \tag{41}$$

But $\partial P/\partial x = 0$ in the present case, since the free surface is level, so $\Pi_u(j,\psi) = 0$.

The pressure forcing in (22) is thus well-balanced. This completes the proof.

Section 8.3 describes a numerical experiment that illustrates this well-balancing in the presence of discontinuous and sloping bottom topography.

5.5. Linearization

Here we derive linearized forms of the equations for momentum and mass given in (19)–(26). This is done partly for the analysis of dispersion relations and stability in Sections 6 and 7 and partly for the motivation for Lax-Friedrichs interpolation of H given in Section 5.6.

Assume that the bottom of the fluid domain is level and that the flow is a small perturbation of the rest state for which the free surface is level and the velocity is zero. Also assume that the wind stress and bottom stress are zero.

In the formulas for the pressure forcing, neglect the atmospheric pressure (i.e., let $p_0 = 0$ in those formulas), as justified after equation (22). Let $\tilde{p}_b$ denote the constant value of the weight of the water column per unit horizontal area in the rest state; if the atmospheric pressure were zero, then $\tilde{p}_b$ would be the equilibrium value of bottom pressure. Let $p_b(x,t)$ denote the perturbation in bottom pressure, so that $\tilde{p}_b + p_b(x,t)$ denotes the weight of the water column per unit horizontal area in a general non-equilibrium state. That is, in the present discussion the quantity $\tilde{p}_b + p_b$ will take the place of the symbol "p_b" in the preceding discussions.

Also assume that the perturbation p_b and the velocity components u and v are small, in the sense that products of "small" quantities can be neglected in the governing equations. With this approximation, the pointwise form (19) of the u-component of the momentum equation simplifies to

$$\tilde{p}_b \frac{\partial u}{\partial t} - f\tilde{p}_b v = -\frac{\partial H}{\partial x}, \tag{42}$$

and the vertically-integrated horizontal pressure forcing (20) becomes

$$\begin{aligned} H(x,t) &= \int_0^{\tilde{p}_b + p_b(x,t)} \alpha\, p\, dp \quad = \quad \frac{\alpha}{2}\left[(\tilde{p}_b)^2 + 2\tilde{p}_b p_b + p_b^2\right] \\ &\approx \frac{1}{2}\alpha(\tilde{p}_b)^2 + (\alpha\tilde{p}_b)p_b. \end{aligned}$$

Let $c^2 = \alpha\tilde{p}_b = g\tilde{h}$, where $\tilde{h}$ is the thickness of the fluid layer at the rest state, and $c > 0$. As seen in Section 5.6, c is the speed of propagation of gravity waves in the case $f = 0$. The linearization of H is then

$$H(x,t) = \frac{1}{2}c^2\tilde{p}_b + c^2 p_b, \tag{43}$$

and the linearized u-momentum equation (42) can be written as

$$\frac{\partial}{\partial t}\left(\frac{u}{c}\right) - f\left(\frac{v}{c}\right) = -c\frac{\partial}{\partial x}\left(\frac{p_b}{\tilde{p}_b}\right).$$

Now let $U(x,t) = u(x,t)/c$ and $V(x,t) = v(x,t)/c$ denote non-dimensional components of velocity, and let $\eta(x,t) = p_b(x,t)/\tilde{p}_b$ denote the perturbation in bottom pressure relative to $\tilde{p}_b$. Due to the hydrostatic condition and the assumption of constant density, $\eta(x,t)$ is also the perturbation in the elevation of the free surface relative to the mean depth. The linearized u-momentum equation can then be expressed as

$$U_t - fV = -c\eta_x. \tag{44}$$

Similarly, the v-component of the momentum equation is

$$V_t + fU = 0, \tag{45}$$

in the quasi-one-dimensional configuration presently considered here. The linearization of the mass equation (25) is

$$\frac{\partial p_b}{\partial t} + \tilde{p}_b \frac{\partial u}{\partial x} = 0,$$

or

$$\eta_t + cU_x = 0. \tag{46}$$

The weak form of the linearized u-momentum equation (44) is

$$\begin{aligned} & \int_{D_j} \left\{ \frac{\partial U}{\partial t} - fV \right\} \psi(x)\ dx \\ = \ & - \Big[c\eta(x,t)\psi(x) \Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} + \int_{D_j} c\eta(x,t)\psi'(x)\ dx, \end{aligned} \tag{47}$$

the weak form of the linearized v-momentum equation (45) is

$$\int_{D_j} \left\{ \frac{\partial V}{\partial t} + fU \right\} \psi(x)\ dx = 0, \tag{48}$$

and the weak form of the linearized mass equation (46) is

$$\int_{D_j} \frac{\partial \eta}{\partial t}\ \psi(x)\ dx + \Big[cU(x,t)\psi(x) \Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} - \int_{D_j} cU(x,t)\psi'(x)\ dx = 0. \tag{49}$$

5.6. Riemann problem and motivation for Lax-Friedrichs interpolation

Section 5.3 describes the problem of determining values of H, and fluxes in general, at cell edges where the solution may be discontinuous. A standard approach to this problem is to solve a Riemann problem with piecewise constant initial data defined by the left and right limits of the solution at that edge; the solution of this problem then produces values of the dependent variables at the edge that could be used in a numerical method. This can be regarded as a kind of "intelligent interpolation" that uses the dynamics of the partial differential equation to interpolate between left and right states, as opposed to simple averaging.

Here, consider the case of the one-dimensional linearized shallow water equations with $f = 0$. In this case, equations (44) and (46) have the forms $U_t + c\eta_x = 0$ and $\eta_t + cU_x = 0$, respectively. Addition and subtraction yield the equations

$$(U+\eta)_t + c(U+\eta)_x = 0 \qquad \text{and} \qquad (U-\eta)_t - c(U-\eta)_x = 0.$$

The quantities $w_1 = (U+\eta)/2$ and $w_2 = (U-\eta)/2$ thus propagate with characteristic velocities c and $-c$, respectively, so c is the speed of the gravity waves that are propagated in this system. Now consider the Riemann problem for U and η with initial data $U(x,t_0) = U_L$ and $\eta(x,t_0) = \eta_L$ for $x < x_{j-1/2}$ and $U(x,t_0) = U_R$ and $\eta(x,t_0) = \eta_R$ for $x > x_{j-1/2}$. Convert these initial states to

initial data for w_1 and w_2, solve for w_1 and w_2, and use the relations $U = w_1 + w_2$ and $\eta = w_1 - w_2$ to obtain

$$\begin{aligned} U(x_{j-1/2},t) &= \frac{1}{2}\Big(U_L + U_R\Big) + \frac{1}{2}\Big(\eta_L - \eta_R\Big) \\ \eta(x_{j-1/2},t) &= \frac{1}{2}\Big(\eta_L + \eta_R\Big) + \frac{1}{2}\Big(U_L - U_R\Big). \end{aligned} \tag{50}$$

In equations (44) and (46) the quantities $c\eta$ and cU, respectively, play the role of fluxes. When the above expressions for η and U are multiplied by c, the results are Lax-Friedrichs fluxes; that is, in each case use the average of the flux values on either side of the edge, plus $c/2$ times a diffusive difference of the dependent variable being computed.

The linearization of the vertically-integrated horizontal pressure forcing H is given in equation (43), and its Lax-Friedrichs interpolation at cell edge $x_{j-1/2}$ is

$$\begin{aligned} H^{LF}(x_{j-1/2},t) &= \frac{1}{2}c^2\tilde{p}_b + c^2\tilde{p}_b\eta(x_{j-1/2},t) \\ &= \frac{1}{2}c^2\tilde{p}_b + \frac{1}{2}c^2\tilde{p}_b\Big(\eta(x^-_{j-1/2},t) + \eta(x^+_{j-1/2},t)\Big) \\ &\qquad + \frac{1}{2}c^2\tilde{p}_b\Big(U(x^-_{j-1/2},t) - U(x^+_{j-1/2},t)\Big) \\ &= \frac{1}{2}\Big[H(x^-_{j-1/2},t) + H(x^+_{j-1/2},t)\Big] \\ &\qquad + \frac{c}{2}\Big[\tilde{p}_b u(x^-_{j-1/2},t) - \tilde{p}_b u(x^+_{j-1/2},t)\Big]. \end{aligned} \tag{51}$$

This derivation uses the relations $p_b = \tilde{p}_b\eta$ and $u = cU$. The expression (51) provides motivation for the interpolation formula (34) stated earlier.

6. Analysis of numerical dispersion relations

In this section we develop a discontinuous Galerkin representation of the system (47)–(49) of linearized weak forms of the shallow water equations and then give an analysis of numerical dispersion relations for this system.

General analyses of dispersion relations for DG methods have been given previously (Ainsworth [2], Bernard et al. [5], Guo et al. [15]). One purpose of the present analysis is to provide a check on the accuracy resulting from the representation of pressure forcing developed earlier, including the usage of Lax-Friedrichs interpolation. This analysis also includes a comparison with finite difference approximations on the classical B- and C-grids that are often used in ocean modeling. The formulation developed here will then be used in Section 7 to analyze the stability of time-stepping methods.

6.1. Discontinuous Galerkin formulation

In Section 3 it was assumed that the spatial interval $[a,b]$ is partitioned into grid cells of the form $D_j = [x_{j-1/2}, x_{j+1/2}]$ for $1 \le j \le J$. For each

j, let x_j denote the center of cell D_j, and for notational simplicity assume that the grid cells have equal length Δx. Then the endpoints of cell D_j are $x_{j\pm1/2} = x_j \pm \Delta x/2$.

6.1.1. Basis functions

For any positive integer N, let $\mathcal{P}^N(D_j)$ denote the space of all polynomials of degree N or less on cell D_j. In order to develop a basis of $\mathcal{P}^N(D_j)$, let P_m be the Legendre polynomial of degree $m \geq 0$ on the reference interval $[-1, 1]$, and assume that P_m is normalized so that $P_m(1) = 1$. Then $P_m(-1) = (-1)^m$ for all $m \geq 0$, since P_m is an even function if m is even and is an odd function if m is odd. The Legendre polynomials satisfy the orthogonality condition $\int_{-1}^{1} P_m P_n = 2/(2m+1)$ if $m = n$ and $\int_{-1}^{1} P_m P_n = 0$ otherwise.

Now let

$$\psi_m^{(j)}(x) = P_m\left(\frac{2}{\Delta x}(x - x_j)\right) \tag{52}$$

for $x_{j-1/2} < x < x_{j+1/2}$ and $0 \leq m \leq N$. The set $\left\{\psi_0^{(j)}, \psi_1^{(j)}, \ldots, \psi_N^{(j)}\right\}$ is a (modal) basis of the space $\mathcal{P}^N(D_j)$. Also, $\psi_m^{(j)}(x_{j+1/2}) = 1$, $\psi_m^{(j)}(x_{j-1/2}) = (-1)^m$, and

$$\int_{D_j} \psi_m^{(j)} \psi_n^{(j)} \;=\; \begin{cases} \frac{\Delta x}{2m+1} & \text{if } m = n \\ 0 & \text{if } m \neq n. \end{cases} \tag{53}$$

The following property will also be used later.

Remark 1. *Let*

$$B_{mn} = \int_{D_j} \frac{d\psi_m^{(j)}}{dx} \psi_n^{(j)}(x)\, dx. \tag{54}$$

Then $B_{mn} = 0$ if $m \leq n$ and $B_{mn} = 1 - (-1)^{m+n}$ if $m > n$.

Proof. A change of variable shows $B_{mn} = \int_{-1}^{1} P_m'(\xi) P_n(\xi)\, d\xi$. If $m \leq n$, then P_m' has degree strictly less than n. But P_n is orthogonal on $[-1, 1]$ to all polynomials of degree less than n, so $B_{mn} = 0$ in this case. On the other hand, if $m > n$ then

$$B_{mn} = \Big[P_m(\xi) P_n(\xi)\Big]_{\xi=-1}^{\xi=1} - \int_{-1}^{1} P_m P_n'.$$

The integral is zero, since P_n' has degree less than m, so $B_{mn} = 1 - (-1)^m(-1)^n$. This completes the proof.

In the system (47)–(49) the dependent variables are the nondimensional velocity components U and V and the relative perturbation η in bottom pressure (or free-surface elevation). Represent the polynomial approximations to U, V, and η in terms of the basis functions by

$$U(x,t) \;=\; \sum_{n=0}^{N} U_n^{(j)}(t)\; \psi_n^{(j)}(x)$$

$$
\begin{aligned}
V(x,t) &= \sum_{n=0}^{N} V_n^{(j)}(t)\ \psi_n^{(j)}(x) \qquad (55)\\
\eta(x,t) &= \sum_{n=0}^{N} \eta_n^{(j)}(t)\ \psi_n^{(j)}(x)
\end{aligned}
$$

for all $x \in D_j$, for all t. (With a slight abuse of notation, the approximations to U,V,η are also denoted by U,V,η.) The time-dependent coefficients $U_n^{(j)}(t)$, $V_n^{(j)}(t)$, and $\eta_n^{(j)}(t)$ will be referred to as "degrees of freedom" for U, V, and η, respectively.

6.1.2. Ordinary differential equations for the degrees of freedom

Insert the representations (55) into the weak form (47) of the linearized u-momentum equation, and let the test function ψ in (47) be the basis function $\psi_m^{(j)}$ for $0 \le m \le N$. The result is

$$
\begin{aligned}
&\sum_{n=0}^{N} \left(\int_{D_j} \psi_m^{(j)} \psi_n^{(j)} \right) \frac{dU_n^{(j)}}{dt} \;-\; f \sum_{n=0}^{N} \left(\int_{D_j} \psi_m^{(j)} \psi_n^{(j)} \right) V_n^{(j)}(t)\\
&= \; -\, c \Big[\eta^* \psi_m^{(j)}(x) \Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} \;+\; c \sum_{n=0}^{N} \left(\int_{D_j} \frac{d\psi_m^{(j)}}{dx} \psi_n^{(j)} \right) \eta_n^{(j)}(t). \qquad (56)
\end{aligned}
$$

Here, η^* refers to values of η at cell edges, as specified below. The orthogonality relation (53) and the definition of B_{mn} in (54) imply

$$
\begin{aligned}
&\frac{\Delta x}{2m+1} \left(\frac{dU_m^{(j)}}{dt} - f V_m^{(j)}(t) \right)\\
&= \; -\, c \Big[\eta^* \psi_m^{(j)}(x) \Big]_{x=x_{j-1/2}}^{x=x_{j+1/2}} \;+\; c \sum_{n=0}^{N} B_{mn} \eta_n^{(j)}(t). \qquad (57)
\end{aligned}
$$

Let Δt be a time increment, such as what might be used in a time-stepping method, and let $\nu = c\Delta t/\Delta x$ be the corresponding Courant number. Then

$$
\begin{aligned}
(\Delta t)\ \frac{dU_m^{(j)}}{dt} &= (f\Delta t)\ V_m^{(j)}(t) \;-\; \nu(2m+1) \Big[\eta_{j+1/2}^*(t) - (-1)^m \eta_{j-1/2}^*(t) \Big]\\
&\quad + \; \nu(2m+1) \sum_{n=0}^{N} B_{mn} \eta_n^{(j)}(t) \qquad (58)
\end{aligned}
$$

for $0 \le m \le N$. The coefficients $f\Delta t$ and $\nu(2m+1)$ are dimensionless. The form of the boundary terms in (58) follows from the relations $\psi_m^{(j)}(x_{j+1/2}) = 1$ and $\psi_m^{(j)}(x_{j-1/2}) = (-1)^m$.

In the linearized weak form (47) the quantity $c\eta(x,t)$ plays the same role as does $H(x,t)$ in the general weak form (21)–(22). (Also compare to the linearization of $H(x,t)$ given in (43).) In keeping with the discussion in Sections 5.3 and

5.6, Lax-Friedrichs interpolation will be used here to determine the values of η^* at the cell edges. A comparison with (50) yields

$$\begin{aligned} \eta^*_{j-1/2}(t) &= \frac{1}{2}\left(\eta(x^-_{j-1/2},t)+\eta(x^+_{j-1/2},t)\right) \\ &+ \frac{1}{2}\left(U(x^-_{j-1/2},t)-U(x^+_{j-1/2},t)\right). \end{aligned} \tag{59}$$

In this representation, the left limit uses data from cell D_{j-1} and the right limit uses data from cell D_j. The representations of η and U in (55) yield

$$\begin{aligned} \eta^*_{j-1/2}(t) &= \frac{1}{2}\sum_{n=0}^{N}\left(\eta_n^{(j-1)}(t)\ \psi_n^{(j-1)}(x_{j-1/2})+\eta_n^{(j)}(t)\ \psi_n^{(j)}(x_{j-1/2})\right) \\ &+ \frac{1}{2}\sum_{n=0}^{N}\left(U_n^{(j-1)}(t)\ \psi_n^{(j-1)}(x_{j-1/2})-U_n^{(j)}(t)\ \psi_n^{(j)}(x_{j-1/2})\right) \\ &= \frac{1}{2}\sum_{n=0}^{N}\left(\eta_n^{(j-1)}(t)+(-1)^n\eta_n^{(j)}(t)\right) \\ &+ \frac{1}{2}\sum_{n=0}^{N}\left(U_n^{(j-1)}(t)-(-1)^nU_n^{(j)}(t)\right). \end{aligned} \tag{60}$$

The DG implementation of the linearized u-momentum equation (47) is then given by the ordinary differential equations in (58), with η^* defined in (60).

The weak form (49) of the linearized mass equation has the same structure as the weak form (47) of the linearized u-momentum equation, with the roles of η and U reversed and the term fV deleted. If the Lax-Friedrichs flux is used to approximate the mass flux at cell edges, then the DG implementation of the mass equation is obtained by modifying (58) to yield

$$\begin{aligned} (\Delta t)\ \frac{d\eta_m^{(j)}}{dt} &= -\nu(2m+1)\Big[U^*_{j+1/2}(t)-(-1)^mU^*_{j-1/2}(t)\Big] \\ &+ \nu(2m+1)\sum_{n=0}^{N}B_{mn}U_n^{(j)}(t) \end{aligned} \tag{61}$$

for $0 \le m \le N$, where the edge value $U^*_{j-1/2}(t)$ is obtained by interchanging the roles of U and η in (60).

The v-momentum equation (48) is analogous to the u-momentum equation (47), except that a pressure term is not present in the case considered here. Its DG implementation is

$$(\Delta t)\ \frac{dV_m^{(j)}}{dt} = -\left(f\Delta t\right)\ U_m^{(j)}(t) \tag{62}$$

for $0 \le m \le N$.

6.1.3. Fourier representation

The DG representation of the momentum and mass equations consists of the system (58), (61), (62), with the unknowns consisting of the degrees of freedom $U_m^{(j)}$, $V_m^{(j)}$, $\eta_m^{(j)}$ for $0 \le m \le N$ and $1 \le j \le J$. In order to make an analysis of this system tractable, assume that for fixed t the degrees of freedom have Fourier representations with respect to the spatial index j, or equivalently, with respect to position x_j. The actual DG approximations to U, V, and η are piecewise polynomials in x, for each t, and these are not assumed to have Fourier representations. Instead, the degrees of freedom are assumed to have such representations, and modal solutions for these degrees of freedom will be developed and analyzed.

Instead of inserting general Fourier superpositions into the system (58), (61), (62), we will simplify notation somewhat by inserting simple oscillatory solutions

$$\begin{aligned} U_m^{(j)}(t) &= \hat{U}_m(k,t)e^{ikx_j} \\ V_m^{(j)}(t) &= \hat{V}_m(k,t)e^{ikx_j} \\ \eta_m^{(j)}(t) &= \hat{\eta}_m(k,t)e^{ikx_j} \end{aligned} \tag{63}$$

into that system. The Fourier transform of the u-momentum equation (58) is then

$$\begin{aligned} (\Delta t)\,\frac{\partial \hat{U}_m}{\partial t}(k,t) &= (f\Delta t)\hat{V}_m(k,t) \\ &- \nu(2m+1)\Big[e^{ik\Delta x} - (-1)^m\Big]\hat{\eta}^*(k,t) \\ &+ \nu(2m+1)\sum_{n=0}^{N} B_{mn}\hat{\eta}_n(k,t), \end{aligned} \tag{64}$$

for $0 \le m \le N$, where, from (60),

$$\begin{aligned} \hat{\eta}^*(k,t) &= \frac{1}{2}\sum_{n=0}^{N}\Big(e^{-ik\Delta x} + (-1)^n\Big)\hat{\eta}_n(k,t) \\ &+ \frac{1}{2}\sum_{n=0}^{N}\Big(e^{-ik\Delta x} - (-1)^n\Big)\hat{U}_n(k,t). \end{aligned} \tag{65}$$

Now insert (65) into (64) to produce

$$\begin{aligned} (\Delta t)\,\frac{\partial \hat{U}_m}{\partial t}(k,t) &= (f\Delta t)\hat{V}_m(k,t) \\ + \nu(2m+1)\sum_{n=0}^{N}&\Big[B_{mn} + \frac{1}{2}\Big((-1)^m - e^{ik\Delta x}\Big)\Big(e^{-ik\Delta x} + (-1)^n\Big)\Big]\hat{\eta}_n(k,t) \\ + \nu(2m+1)\sum_{n=0}^{N}&\frac{1}{2}\Big((-1)^m - e^{ik\Delta x}\Big)\Big(e^{-ik\Delta x} - (-1)^n\Big)\hat{U}_n(k,t) \end{aligned} \tag{66}$$

for $0 \le m \le N$.

The Fourier transform of the mass equation (61) is obtained by interchanging the roles of $\hat{U}$ and $\hat{\eta}$ in (66) and deleting the term involving f. The Fourier transform of the v-momentum equation (62) is

$$(\Delta t)\ \frac{\partial \hat{V}_m}{\partial t}(k,t) = -(f\Delta t)\hat{U}_m(k,t). \tag{67}$$

The Fourier transform of the entire system can be written in matrix-vector form as follows. Define a column vector $q(k,t)$ by

$$q(k,t) = \left(\hat{U}_0, \ldots, \hat{U}_N, \hat{V}_0, \ldots, \hat{V}_N, \hat{\eta}_0, \ldots, \hat{\eta}_N\right)^T, \tag{68}$$

where all quantities on the right side depend on (k,t). Then

$$\frac{\partial q}{\partial t} \ = \ \left(\frac{1}{\Delta t}\right) Aq, \tag{69}$$

where

$$A(k\Delta x, f\Delta t) = \begin{pmatrix} E & (f\Delta t)I & F \\ -(f\Delta t)I & 0 & 0 \\ F & 0 & E \end{pmatrix}. \tag{70}$$

Here, I is the $(N+1)\times(N+1)$ identity matrix, 0 denotes the $(N+1)\times(N+1)$ matrix of zeros, and E and F are nonsymmetric $(N+1)\times(N+1)$ matrices whose entries are implied by (66). The entries in the matrix A are dimensionless.

Solutions of the system (69) can be constructed with eigenvalues and eigenvectors of the matrix A in (70). Let λ be an eigenvalue of A with corresponding eigenvector z; since A is dimensionless, λ is also dimensionless. Denote λ by $\lambda = (\mathrm{Re}\,\lambda) - i\omega\Delta t$, where ω is real and has units of 1/time. Let t_n be a given time, such as a time level that is encountered with time-stepping method. A corresponding solution of the system (69) is then

$$q(k,t) = \exp\left[\frac{\lambda}{\Delta t}(t-t_n)\right] z = \exp\left[(\mathrm{Re}\,\lambda)\left(\frac{t-t_n}{\Delta t}\right) - i\omega(t-t_n)\right] z, \tag{71}$$

and the solution at time $t_n + \Delta t$ is

$$q(k, t_n + \Delta t) = e^{\lambda} z = e^{\mathrm{Re}\,\lambda} e^{-i\omega\Delta t} z. \tag{72}$$

A comparison with the form (63) of oscillatory (in x) solutions shows that ω is a frequency corresponding to spatial wavenumber k. Plots of the (dimensionless) imaginary part $\omega\Delta t$ of λ versus the dimensionless wavenumber $k\Delta x$ can be used to illustrate phase velocity and group velocity of wave motions. Plots of the quantity $e^{\mathrm{Re}\,\lambda}$ versus $k\Delta x$ illustrate growth and/or decay over a time increment Δt and thus can be used to assess stability and numerical dissipation.

6.2. Exact solution and staggered finite-difference grids

For purposes of comparison, the present subsection states the dispersion relations for oscillatory wave solutions of the exact system (44)–(46) and of second-order finite difference approximations on the B-grid and the C-grid. The labeling and classification of these and some other finite difference grids are due to Arakawa and Lamb [3], and the wave propagation properties of these grids have long been established ([3], Dukowicz [10]). The dispersion relations for the B- and C-grids for the present context are included here in order to make the discussion more self-contained.

In analogy to the oscillatory solutions (63) of the DG approximation, insert solutions

$$\begin{aligned} U(x,t) &= \hat{U}(k,t)e^{ikx} \\ V(x,t) &= \hat{V}(k,t)e^{ikx} \\ \eta(x,t) &= \hat{\eta}(k,t)e^{ikx} \end{aligned} \tag{73}$$

into (44)–(46) to produce

$$\frac{\partial}{\partial t}\begin{pmatrix} \hat{U} \\ \hat{V} \\ \hat{\eta} \end{pmatrix} = \begin{pmatrix} 0 & f & -ikc \\ -f & 0 & 0 \\ -ikc & 0 & 0 \end{pmatrix}\begin{pmatrix} \hat{U} \\ \hat{V} \\ \hat{\eta} \end{pmatrix}. \tag{74}$$

A calculation shows that the 3×3 coefficient matrix has eigenvalues $\lambda = -i\omega$, where

$$\omega = 0 \qquad \text{or} \qquad \omega^2 = c^2k^2 + f^2. \tag{75}$$

A comparison with (73) shows that the exact system (44)–(46) has oscillatory wave solutions $e^{ikx-i\omega t}z$, where z is a 3×1 vector. For such waves, the phase velocity (velocity of single, pure sinusoidal waves) is ω/k, and the group velocity (velocity of wave packets) is $d\omega/dk$. The nonzero values of ω in (75) correspond to inertia-gravity waves, and the case $\omega = 0$ is a stationary mode that would transform into a Rossby mode if the Coriolis parameter f were variable. The equations in (75) constitute the "dispersion relation" for the system (44)–(46).

For purposes of comparison to numerical methods, let Δx denote a grid spacing (or cell length), and write the dispersion relation in (75) for nonzero ω as

$$\left(\frac{\omega}{f}\right)^2 = R^2(k\Delta x)^2 + 1, \tag{76}$$

where ω/f is a nondimensional frequency and

$$R = \frac{c}{(f\Delta x)} = \frac{c/f}{\Delta x}. \tag{77}$$

The quantity c/f is the Rossby radius associated with speed c, so R is the ratio of the Rossby radius to the grid size.

In order to define the B- and C-grids on a general two-dimensional region, assume that the spatial domain is partitioned into rectangular grid cells with

mass variables defined at the centers of those cells. In the case of the B-grid, the components of horizontal velocity are defined at the corners of the mass cells. With the C-grid, the normal components of velocity are defined at the centers of the cell edges, i.e., U is defined at the centers of the edges corresponding to minimal and maximal x, and V is defined at the centers of the edges corresponding to minimal and maximal y.

For the quasi-one-dimensional setting considered here, assume that the mass variable η is defined at points x_j having integer indices. For the B-grid, U and V are then defined at points $x_{j\pm1/2}$ having half-integer indices. In the case of the C-grid, U is defined at points of the form $x_{j\pm1/2}$, and V is defined at the same points as η. For each grid, the spatial derivatives in the system (44)–(46) are approximated with centered second-order finite differences that are natural for the grid in question. The Coriolis terms require values of V at U-points and values of U at V-points, and on the C-grid these values are obtained with simple averages. On a B-grid with two horizontal dimensions, a pressure term $-c\eta_x$ (see (44)) is implemented with an average in y of a difference in x; in the quasi-one-dimensional case considered here, such averaging has no effect since all quantities are independent of y.

Some calculations, analogous to those given above, show that the dispersion relation for the B-grid is

$$\omega = 0 \qquad \text{or} \qquad \left(\frac{\omega}{f}\right)^2 = R^2 \left(2\sin\frac{k\Delta x}{2}\right)^2 + 1, \tag{78}$$

and the dispersion relation for the C-grid is

$$\omega = 0 \qquad \text{or} \qquad \left(\frac{\omega}{f}\right)^2 = R^2 \left(2\sin\frac{k\Delta x}{2}\right)^2 + \left(\cos\frac{k\Delta x}{2}\right)^2. \tag{79}$$

for $|k\Delta x| \le \pi$. The cosine term in (79) arises from the spatial averaging that is required to implement the Coriolis term on the C-grid. In (78) and (79) it is assumed that there is no discretization with respect to t, i.e., the effects of numerical time-stepping methods are not included here.

6.3. Remarks about Rossby radius and the B- and C-grids

The dispersion relations (76), (78), and (79) express the nondimensional frequency ω/f in terms of the nondimensional wavenumber $k\Delta x$, and in those relations the only free parameter is the ratio R of the Rossby radius c/f to the grid size Δx.

For a rotating spheroid, the Coriolis parameter is $f = 2\Omega\sin\theta$, where Ω is the angular rate of rotation and θ is the latitude; for the mid-latitudes of the earth, $f \approx 10^{-4}\,\mathrm{sec}^{-1}$. For the constant-density shallow water equations presently considered here, the gravity wave speed (if f were zero) is $c = \sqrt{gh}$, where h is the thickness of the fluid layer. For a generic mid-ocean depth of 4000 meters, $c \approx 200$ m/sec, and the Rossby radius is then $c/f \approx 2000$ kilometers. This is far greater than any possible grid size, so $R \gg 1$ in this case.

However, the variable-density ocean also admits a multitude of internal modes. For the case of linearized flow in a region with a flat bottom, separation of variables reveals an infinite sequence of internal modes, with the vertical dependence given by an eigenfunction and with the time and horizontal dependences modeled by the linearized shallow water equations. This is derived, for example, by Higdon [18]. The value of c for the fastest internal mode is typically on the order of one to two meters per second, and the other speeds are smaller, with the sequence of speeds tending to zero. In a vertically-discrete three-dimensional model, the number of internal modes is equal to the number of vertical coordinate surfaces, e.g., layer interfaces for an isopycnic model.

If $c = 2$ m/sec and $f = 10^{-4}\,\mathrm{sec}^{-1}$, then the Rossby radius is 20 kilometers; the Rossby radius for slower modes is then smaller. The grid sizes for ocean circulation models can be on the order of kilometers or tens of kilometers or perhaps more, depending on the application, so it would be worthwhile to use a numerical method that performs well for all of the regimes $R > 1$, $R \approx 1$, and $R < 1$.

Analyses of the shallow water equations in two horizontal dimensions have indicated that when $R > 1$, the C-grid is generally better than the B-grid for propagating inertia-gravity waves; when $R < 1$, the situation is reversed (Arakawa and Lamb [3]). Inertia-gravity waves participate, for example, in adjustment processes in which a change in forcing leads to a shift in the mean state of the system. The B-grid has been judged to be somewhat better for propagating vorticity-driven Rossby waves (Dukowicz [10]), although this can depend on the location in wavenumber space. Rossby waves are involved in the development of large-scale current systems (Gill [11]).

This variation of performance of the B- and C-grids has been one motivation for the present exploration of discontinuous Galerkin methods as an alternate method of spatial discretization for ocean modeling.

6.4. Comparison of DG approximations with the B- and C-grids

We now use the ideas developed in Sections 6.1.3 and 6.2 to compare the accuracy of DG approximations to the accuracy of the B- and C-grids. The present analysis is restricted to the quasi-one-dimensional setting, due to the relative complexity of DG methods. Here, the shallow water equations are studied partly for their own sake and partly as a proxy for representing the dynamics of internal modes.

This comparison is based on plots of the nondimensional frequency ω/f versus the nondimensional wavenumber $k\Delta x$. For the DG system (69), the imaginary parts $-\omega\Delta t$ of the eigenvalues of the matrix A in (70) will be divided by the nondimensional Coriolis parameter $f\Delta t$. Equivalently, A is divided by $f\Delta t$, and a comparison with (66) and (70) shows that the resulting matrix depends only on $k\Delta x$ and the parameter $\nu/f\Delta t = (c\Delta t/\Delta x)/(f\Delta t) = R$. This is the same parameter seen in the dispersion relation (76) for the exact system and the dispersion relations (78)–(79) for the B- and C-grids.

Equation (72) indicates that any growth or decay in the solution over a time increment Δt is given by the factor $e^{\mathrm{Re}\,\lambda}$. If the normalized imaginary part

$(\omega\Delta t)/f\Delta t)$ is plotted, then it is natural to use the same normalization for the real part and plot $e^{\mathrm{Re}\,\lambda/(f\Delta t)}$; the resulting plots then indicate the amount of growth or decay over a time interval of length $\Delta t = 1/f$.

Figure 1 shows plots of the dispersion relations (78) and (79) for the B-grid and C-grid, respectively, for the case $R = 2$. The horizontal coordinate is restricted to the range $0 \le k\Delta x \le \pi$, as the case $k\Delta x = \pi$ corresponds to waves of length $2\Delta x$, which are the shortest waves that can be seen on a grid with spacing Δx. The vertical range is restricted to $\omega/f \ge 0$, as the case $\omega/f < 0$ is simply a reflection. In each frame, the dashed curve shows the positive root of the exact dispersion relation (76) for inertia-gravity waves. The solid curve and solid line illustrate the inertia-gravity mode and stationary mode, respectively, in the difference approximations. The near-agreement of the dashed and solid curves for $k\Delta x$ near 0 illustrates the consistency of the finite difference approximations. However, for larger $k\Delta x$ the finite difference approximations show substantial error in the phase velocity ω/k and group velocity $d\omega/dk$.

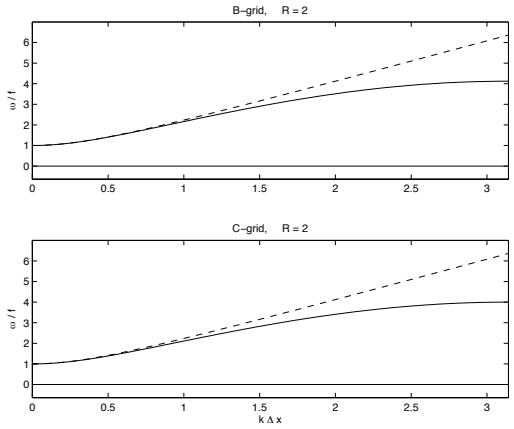

Figure 1: Dispersion relations for centered second-order finite difference approximations on the B-grid and C-grid, for the case $R = 2$. Here, R is the ratio of the Rossby radius c/f to the grid spacing Δx. In each frame, the dashed curve shows the inertia-gravity mode in the exact solution, and the solid curve and solid line show the dispersion relation for the finite difference approximation. It is assumed here that there is no discretization with respect to time.

Figure 2 shows plots of the dispersion relations for the B- and C-grids for the case $R = 1/2$. In this case, the group velocity for the inertia-gravity mode on the C-grid is $d\omega/dk = 0$, and no energy can be propagated. This conclusion assumes that there is no discretization with respect to time. In a plot for the C-grid in the case $R = 1/4$ (not shown here), ω/f is actually a decreasing function of $k\Delta x$ for $0 \le k\Delta x \le \pi$, so the group velocity has the wrong sign in that case, and thus energy propagates in the wrong direction.

Next consider the dispersion relations for the DG approximation. Figures

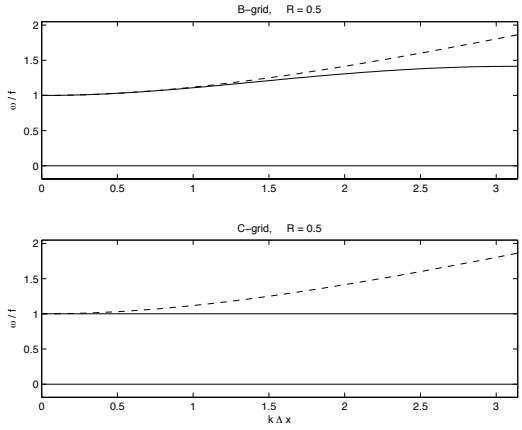

Figure 2: Dispersion relations for the B-grid and C-grid, for the case $R = 1/2$. In the case of the C-grid, the group velocity for the inertia-gravity mode is $d\omega/dk = 0$. In the case $R = 1/4$ (not shown here), the group velocity for the C-grid has the wrong sign, so energy is propagated in the wrong direction in that case.

3 and 4 illustrate the case of piecewise quadratic approximations (i.e., $N = 2$) with ratios $R = 2$ and $R = 1/2$, respectively. A comparison with (55) shows that if $N = 2$ then there are three degrees of freedom for each of the unknowns U, V, and η. The system (69), which represents the Fourier transform of the DG method in this case, then consists of nine equations in nine unknowns. For each value of $k\Delta x$, the matrix A in (70) has nine eigenvalues, which correspond to nine modal solutions of the system (69). In comparison, the exact solution and the B- and C-grid approximations each have two inertia-gravity modes and one stationary mode, so the DG method admits six computational modes in the present case.

In order to produce the plots in Figures 3 and 4, the eigenvalues of A were computed numerically, for each value of $k\Delta x$ in a finely-spaced mesh in the interval $[0, \pi]$. In each of the Figures, the top frame contains plots of ω/f, and the bottom frame contains plots of the damping/growth factor $e^{\mathrm{Re}\,\lambda/(f\Delta t)}$, as described earlier in this subsection. These quantities were sorted prior to plotting, in order to produce continuous curves. For the cases plotted here, inspections of some numerical output indicate that there are two inertia-gravity modes (with positive and negative ω), three stationary modes with $\omega = 0$, and four other modes with $\omega \neq 0$.

In the top frame in each Figure, the solid sloping curve shows the DG representation of the inertia-gravity mode with $\omega > 0$, and the solid horizontal line shows a stationary mode with $\omega = 0$. The inertia-gravity mode in the exact solution is shown with a dashed curve that coincides almost exactly with the DG representation of the inertia-gravity mode over the entire range $0 \leq k\Delta x \leq \pi$. This illustrates a high level of accuracy in the DG method that is radically

different from what is seen with the B- and C-grids in Figures 1 and 2. In particular, the close agreement for larger values of $k\Delta x$ suggests high accuracy at low resolution, and this is illustrated by some numerical experiments described in Section 8.

In the bottom frame in each of Figures 3 and 4, the solid curve represents the damping factor for the inertia-gravity modes. (For given $k\Delta x$, this factor is the same for the two inertia-gravity modes with $\omega > 0$ and $\omega < 0$.) These curves indicate very little spurious numerical dissipation for these modes.

Of the three modes with $\omega = 0$, two have damping factor 1 and one has damping factor less than 1; these are indicated by the horizontal dash-dot lines in the lower frames. The four remaining (computational) modes have values of ω/f that come in plus/minus pairs, and the positive values are illustrated with dash-dot curves in the upper frames that in some portions exceed the vertical ranges of those plots. However, these modes are damped very rapidly in time, as indicated by the remaining dash-dot curves in the lower frames of Figures 3 and 4. In Figure 3, corresponding to $R = 2$, one of the damping factors is so close to zero that it is not visible in the plot.

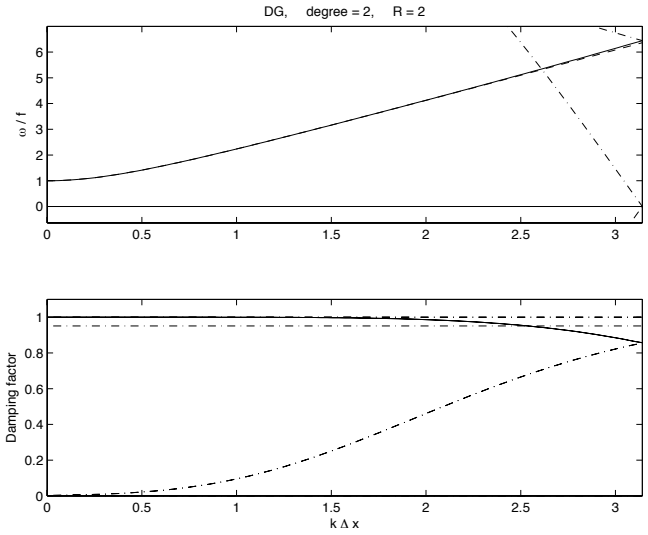

Figure 3: The upper frame shows the dispersion relation for the discontinuous Galerkin approximation to the linearized shallow water equations in the quasi-one-dimensional case, with the Lax-Friedrichs representation of the mass flux and pressure forcing at cell edges, piecewise quadratic approximations, and $R = 2$. It is assumed here that there is no discretization with respect to time. The solid curves represent the physical modes, and the dash-dot curves represent computational modes. The exact dispersion relation for the inertia-gravity wave mode is represented by a dashed curve in the upper frame, which is almost identical to the plot of the inertia-gravity mode in the DG approximation. The lower frame shows damping factors for the various modes, with the solid curve corresponding to the physical inertia-gravity wave mode.

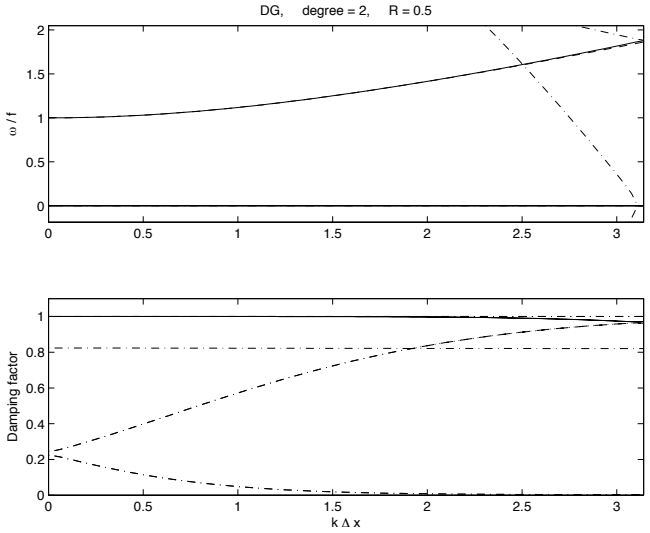

Figure 4: This Figure shows the same situation as Figure 3, except that $R = 0.5$ in this case.

For piecewise linear approximations ($N = 1$), the dispersion relations for inertia-gravity waves are not quite as accurate as for $N = 2$, and the numerical dissipation is noticeably larger. For piecewise cubic approximations ($N = 3$), the dispersion relations for inertia-gravity waves are slightly more accurate than for $N = 2$, and the numerical dissipation for those modes is less. Further experiments with other values of R indicate that the DG approximation is accurate for a wide range of values of that parameter. In this respect, the DG method presents a distinct advantage over second-order finite differences on the B- and C-grids.

7. Analysis of time-stepping methods

The dispersion analysis in the preceding section assumes that there is no discretization with respect to time. The present section extends that analysis to a time-stepping method developed by Higdon [16], [17]. We also compare the stability of that method to the stability of some standard Runge-Kutta methods, as applied to the discontinuous Galerkin representation of the linearized shallow water equations in the quasi-one-dimensional setting. The starting point is the formulation (68)–(70), which is obtained after a Fourier transform in x.

7.1. A two-level time-stepping method

The method developed in [16] and [17] is a two-time-level, second-order method for ocean circulation modeling with a barotropic-baroclinic time splitting. With such a splitting, the fast external motions are modeled with a two-dimensional vertically-integrated subsystem that resembles the shallow water

equations, and the slow motions are modeled with a subsystem that is fully three-dimensional. If the barotropic equations are discarded and the baroclinic equations are reduced to the case of a single homogeneous layer, then the method in [16] and [17] becomes a time-stepping method for the shallow water equations.

Here, we apply that method to the Fourier-transformed system in (69). In that system, Lax-Friedrichs interpolation is used to compute the mass fluxes and the pressure forcing at the edges of grid cells. Write (69) in the form

$$\frac{\partial q}{\partial t} \;=\; \frac{\partial}{\partial t}\begin{pmatrix} \hat{U} \\ \hat{V} \\ \hat{\eta} \end{pmatrix} \;=\; \left(\frac{1}{\Delta t}\right) Aq,$$

where

$$\hat{U}(k,t) = \left(\hat{U}_0, \ldots, \hat{U}_N\right)^T, \quad \hat{V}(k,t) = \left(\hat{V}_0, \ldots, \hat{V}_N\right)^T, \quad \hat{\eta}(k,t) = \left(\hat{\eta}_0, \ldots, \hat{\eta}_N\right)^T.$$

Consider the evolution of the system from time t_n to time $t_{n+1} = t_n + \Delta t$, and use superscripts on dependent variables to denote their dependence on the time level. In the present situation, the method in [16] and [17] consists of the following steps.

(i) Predict $\hat{U}$ and $\hat{\eta}$ with a forward Euler step, to produce $\hat{U}^{pred}$ and $\hat{\eta}^{pred}$. (In the quasi-one-dimensional case, the role of $\hat{V}$ is limited to the Coriolis terms, and $\hat{V}$ is updated in step (iii).)

(ii) Correct $\hat{\eta}$ to produce $\hat{\eta}^{n+1}$. The Lax-Friedrichs mass flux uses both $\hat{U}$ and $\hat{\eta}$; during this step, use unweighted averages of $\hat{U}^n$ and $\hat{U}^{pred}$ and of $\hat{\eta}^n$ and $\hat{\eta}^{pred}$.

(iii) Correct $\hat{U}$ and update $\hat{V}$, to produce $\hat{U}^{n+1}$ and $\hat{V}^{n+1}$. Implement the Coriolis term implicitly with the trapezoidal rule, and in the Lax-Friedrichs interpolation in the pressure forcing use unweighted averages of $\hat{U}^n$ and $\hat{U}^{pred}$ and of $\hat{\eta}^n$ and $\hat{\eta}^{n+1}$.

The prediction step (i) can be expressed in matrix-vector form as

$$\begin{pmatrix} \hat{U}^{pred} \\ \hat{V}^{pred} \\ \hat{\eta}^{pred} \end{pmatrix} \;=\; \begin{pmatrix} \hat{U}^n \\ \hat{V}^n \\ \hat{\eta}^n \end{pmatrix} + \Delta t \left(\frac{1}{\Delta t}\right) \begin{pmatrix} E & (f\Delta t)I & F \\ 0 & 0 & 0 \\ F & 0 & E \end{pmatrix} \begin{pmatrix} \hat{U}^n \\ \hat{V}^n \\ \hat{\eta}^n \end{pmatrix},$$

or

$$q^{pred} \;=\; \begin{pmatrix} I+E & (f\Delta t)I & F \\ 0 & I & 0 \\ F & 0 & I+E \end{pmatrix} q^n \;\equiv\; Q_{pred}\, q^n. \tag{80}$$

The matrices E and F depend on the dimensionless wavenumber $k\Delta x$, for fixed values of the Courant number $\nu = c\Delta t/\Delta x$. The quantity $\hat{V}^{pred}$ plays no substantive role and is used here only for notational convenience. The combination of steps (ii) and (iii) can be written in matrix-vector form as

$$\begin{pmatrix} I & -\frac{1}{2}(f\Delta t)I & -\frac{1}{2}F \\ \frac{1}{2}(f\Delta t)I & I & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} \hat{U}^{n+1} \\ \hat{V}^{n+1} \\ \hat{\eta}^{n+1} \end{pmatrix}$$

$$
\begin{aligned}
&= \begin{pmatrix} I + \frac{1}{2}E & \frac{1}{2}(f\Delta t)I & \frac{1}{2}F \\ -\frac{1}{2}(f\Delta t)I & I & 0 \\ \frac{1}{2}F & 0 & I + \frac{1}{2}E \end{pmatrix} \begin{pmatrix} \hat{U}^n \\ \hat{V}^n \\ \hat{\eta}^n \end{pmatrix} \\
&+ \begin{pmatrix} \frac{1}{2}E & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{2}F & 0 & \frac{1}{2}E \end{pmatrix} \begin{pmatrix} \hat{U}^{pred} \\ \hat{V}^{pred} \\ \hat{\eta}^{pred} \end{pmatrix},
\end{aligned}
$$

or

$$
G_1 q^{n+1} \;=\; G_0 q^n + G_{pred}\, q^{pred}. \tag{81}
$$

Equations (80) and (81) can be combined to yield

$$
q^{n+1} \;=\; G_1^{-1}\left(G_0 + G_{pred} Q_{pred}\right) q^n \;\equiv\; G q^n. \tag{82}
$$

For fixed values of ν and $f\Delta t$, the matrix G depends on $k\Delta x$. For given values of those parameters, let λ be an eigenvalue of G with eigenvector z. A corresponding solution to (82) is $q^n = \lambda^n z$, where the superscript on q is a time index and the superscript on λ is an exponent. Write λ as $\lambda = |\lambda| \exp(-i\omega\Delta t)$, where ω is real and has units of 1/time; the dimensionless quantity $-\omega\Delta t$ is then an argument of the complex number λ. A corresponding solution at time t_n to the spatially-dependent problem (i.e., before Fourier transform in x) is

$$
q^n e^{ikx} = |\lambda|^n e^{-i\omega n\Delta t} e^{ikx} = |\lambda|^n e^{ikx - i\omega t_n}. \tag{83}
$$

(Compare to (63).)

Plots of all of the values of $|\lambda|$ versus $k\Delta x$, for fixed ν and $f\Delta t$, can then be used to assess the stability of the method. If $|\lambda| \le 1$ for all modes and all $k\Delta x$, then the method is stable for that ν and $f\Delta t$. If $|\lambda| < 1$ for some eigenvalue, then the method is dissipative for that mode. Furthermore, plots of $\omega\Delta t$ versus $k\Delta x$ give dispersion relations for the various modes and can be compared to the relation (75), $(\omega\Delta t)^2 = \nu^2(k\Delta x)^2 + (f\Delta t)^2$, for the exact inertia-gravity mode.

In a manner analogous that used in Section 6.4, the eigenvalues of G are computed here for each value of $k\Delta x$ on a finely-spaced mesh in the interval $0 \le k\Delta x \le \pi$. Some experiments for the cases $N = 1$ (piecewise linear), $N = 2$ (piecewise quadratic), and $N = 3$ (piecewise cubic) suggest that the stability of the method depends on the value of the Courant number ν but is independent of the value of $f\Delta t$. For each case, let ν_{max} denote the maximum possible value of ν for which the method is stable. The experiments show that if $N = 1$ then $\nu_{max} \approx 0.33$; if $N = 2$ then $\nu_{max} \approx 0.16$; if $N = 3$ then $\nu_{max} \approx 0.09$. These values are also listed in Table 1, which is given later.

The values of ν_{max} for this method are considerably less than the values that are typically encountered when finite differences are used to discretize in space. However, in the next subsection these values are seen to be competitive with those encountered with some standard Runge-Kutta methods, and some numerical experiments described in Section 8 suggest that the higher spatial accuracy with the DG method can more than offset the disadvantage of the greater restriction on the time step.

Figures 5 and 6 show plots of $\omega\Delta t$ versus $k\Delta x$ and $|\lambda|$ versus $k\Delta x$ for the case $N = 2$; $\nu = 0.16$; and $f\Delta t = 0.08$ and $f\Delta t = 0.32$, respectively. In these two plots the ratio $R = (c/f)/\Delta x = \nu/(f\Delta t)$ has values $R = 2$ and $R = 1/2$, respectively, which are the same values used in Figures 1–4.

In the top frames in Figures 5 and 6, the solid curve and solid line show the inertia-gravity and stationary modes as represented in the discrete system, and the dashed curve shows the inertia-gravity mode in the exact solution to the continuous problem. As in the time-continuous case discussed in Section 6.4, the representation of the inertia-gravity mode in the fully-discrete system is nearly exact. The solid curves in the lower frames illustrate the damping factor $|\lambda|$ in the inertia-gravity mode, and for this mode the dissipation is small.

Among the computational modes, there are typically one or more stationary modes that are nearly undamped, and there are other modes for which the dissipation is much stronger. The upper frames do not show dispersion relations for the computational modes in order to reduce clutter, as there are cases where a real eigenvalue migrates across the origin as $k\Delta x$ varies, with the consequence that its argument changes suddenly between 0 and $\pm\pi$. Such modes decay very rapidly in time.

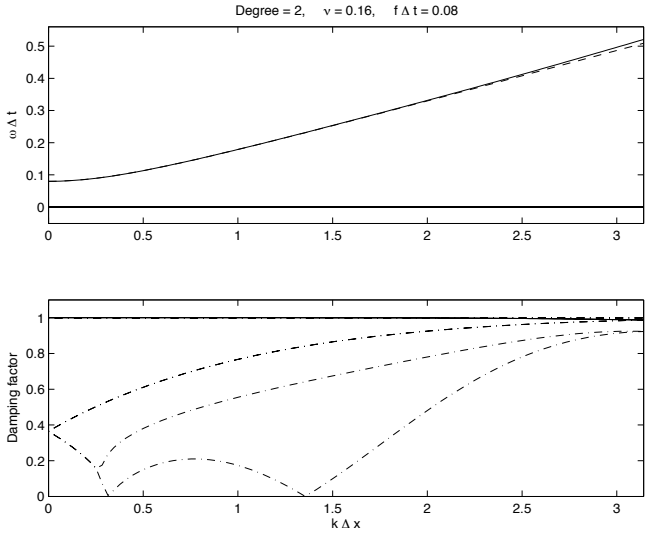

Figure 5: Plots of dispersion relations and damping factors for the fully-discrete problem consisting of a discontinuous Galerkin spatial discretization and the two-level time-stepping method of Higdon [16], [17]. The case shown here uses piecewise quadratic approximations, with $\nu = 0.16$ and $f\Delta t = 0.08$ and thus $R = 2$. In the upper frame, the solid curve shows the inertia-gravity mode as represented in the discrete problem, and the dashed curve shows the inertia-gravity mode in the exact solution. These nearly coincide. In the lower frame, the solid curve shows $|\lambda|$ for the inertia-gravity mode in the discrete system, and the dash-dot curves correspond to other modes.

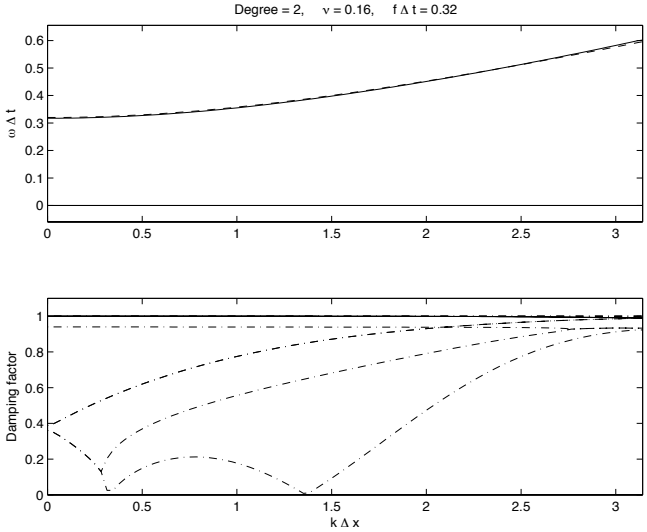

Figure 6: Same configuration as in Figure 5, except that $f\Delta t = 0.32$ and thus $R = 1/2$.

7.2. Runge-Kutta methods

Next consider Runge-Kutta time-stepping methods, as applied to the formulation (68)–(70) of the Fourier transform in x of the discontinuous Galerkin spatial discretization.

The system (69) has the form $q'(t) = (1/\Delta t)Aq$, where the matrix A is independent of t. In that case, a p-stage Runge-Kutta method of order p reduces to a Taylor expansion of the matrix exponential (Ascher and Petzold [4]), i.e., the method is

$$q^{n+1} = \left(\sum_{k=0}^{p} \frac{1}{k!} A^k \right) q^n \;\equiv\; G_{RK} q^n \tag{84}$$

in this case. Runge-Kutta methods with p stages and order p exist only for $p \leq 4$.

A recently-developed alternative to the classical Runge-Kutta methods is the class of strong-stability-preserving Runge-Kutta methods (Gottlieb et al. [13]). Within this class, improvements on the maximum allowable time step can be obtained with methods for which the number of stages exceeds the order (Spiteri and Ruuth [27], Kubatko et al. [21]). For example, the optimal four-stage method of order three for the equation $dq/dt = F(q)$ is

$$\begin{aligned} Q^{(1)} &= q^n + \frac{\Delta t}{2} F\big(q^n\big) \\ Q^{(2)} &= Q^{(1)} + \frac{\Delta t}{2} F\big(Q^{(1)}\big) \end{aligned}$$

$$
\begin{aligned}
Q^{(3)} &= \frac{2}{3}q^n + \frac{1}{3}\left[Q^{(2)} + \frac{\Delta t}{2}F\big(Q^{(2)}\big)\right] \\
q^{n+1} &= Q^{(3)} + \frac{\Delta t}{2}F\big(Q^{(3)}\big). \qquad (85)
\end{aligned}
$$

For the special case $q'(t) = (1/\Delta t)Aq$, this method reduces to

$$
q^{n+1} = \left(I + \frac{1}{2}A\right)\left[\frac{2}{3}I + \frac{1}{3}\left(I + \frac{1}{2}A\right)^3\right]q^n \equiv G_{43}\,q^n. \qquad (86)
$$

Similar to the technique used in Section 7.1, the stability of the above Runge-Kutta methods can be explored by computing eigenvalues of the matrices G_{RK} and G_{43} for all $k\Delta x$ on a finely-spaced mesh in the interval $[0, \pi]$, for various values of the parameter ν and $f\Delta t$. The results of some experiments are listed in Table 1.

In the case of the two-level method discussed in Section 7.1, the stability appears to be independent of the value of $f\Delta t$. However, with the Runge-Kutta methods considered here, the value of that parameter has an influence on stability. The values of ν_{max} given in Table 1 for the Runge-Kutta methods are approximations to the maximum value of ν for which the method is stable when $f\Delta t$ is restricted to the range $0 \le f\Delta t \le 1$.

	$N = 1$	$N = 2$	$N = 3$
Two-level	0.33	0.16	0.09
RK(3,3)	0.40	0.20	0.13
RK(4,4)	0.46	0.23	0.14
SSPRK(4,3)	0.58	0.30	0.18

Table 1: Stability of time-stepping methods. Here, "Two-level" refers to the two-level method described in Section 7.1, $RK(p, p)$ refers to a p-stage Runge-Kutta method of order p, and $SSPRK(4, 3)$ refers to the method in (86). The columns of numbers give approximate values of the maximum Courant number ν_{max} for which the method is stable, for piecewise linear, piecewise quadratic, and piecewise cubic approximations to the solution. The values of ν_{max} for the Runge-Kutta methods are the maximum values of ν when $f\Delta t$ is restricted to the range $0 \le f\Delta t \le 1$. The method $RK(2, 2)$ is not listed in the table, as it is unconditionally unstable for nonzero values of $f\Delta t$. The two-level method has an operation count similar to a two-stage Runge-Kutta method, so the lesser values of ν_{max} relative to the other methods is not a disadvantage.

The values of ν_{max} for the two-level method are smaller than those for the three-stage and four-stage Runge-Kutta methods. However, the two-level method is a predictor-corrector method with an operation count similar to that of a two-stage Runge-Kutta method. If a four-stage Runge-Kutta method is used with twice the value of Δt as is used with the two-level method, then over a given length of model time the two methods would have comparable costs.

Figures 5 and 6 suggest that the two-level method represents the inertia-gravity mode nearly exactly, if a piecewise quadratic DG method is used for

the spatial discretization. The accuracy of the two-level method may suffice, at least for the situation covered by the present analysis.

8. Numerical experiments

Compared to finite difference methods, DG methods present some potential disadvantages related to efficiency. In particular, a DG method requires the computation of multiple degrees of freedom per dependent variable, in each grid cell and at each time step, whereas finite difference methods require only one degree of freedom. In addition, DG methods encounter more restrictive bounds on the allowable time step for stable computations. On the other hand, the multiple degrees of freedom in a DG method allow for higher accuracy, and this raises the question of whether the higher accuracy can compensate for the disadvantages in efficiency that were just mentioned.

The present section describes a numerical experiment which suggests that this is the case, in a simple setting in which a DG method is compared to second-order finite differencing on the B- and C-grids. In the situation described here, DG methods and finite difference methods are equally applicable, and in this test the DG method is at least as good if not better. More generally, DG methods are well-suited for usage on unstructured meshes, and they can obtain high-order accuracy while maintaining high locality. These features represent fundamental advantages over finite difference methods, but a discussion of these points is beyond the scope of the present paper.

The present section also describes a second numerical experiment, with variable bottom topography, which illustrates the well-balanced nature of the pressure forcing that is developed in this paper.

8.1. Configuration of the computations

The DG code used here is an implementation of the weak forms (21), (24), (26) of the nonlinear equations for conservation of momentum and mass in a constant-density fluid in the quasi-one-dimensional configuration that has been discussed here. That is, we consider the shallow water equations in an infinite straight channel in a rotating reference frame.

The formulation of the DG method for this case is similar to the formulation used in Section 6.1 for the linearized equations (47)–(49), in the sense that the polynomial basis functions are Legendre polynomials under a change of independent variable. However, in Section 6.1 the integrals appearing in the weak forms (47)–(49) are evaluated exactly by using orthogonality properties, in order to carry out the analyses of dispersion and stability in Sections 6 and 7, whereas the code used for the present computations employs Gauss-Legendre quadrature to compute all of the integrals. Numerical quadrature is appropriate when an explicit time-stepping method is used, since at any stage in the computation the time tendencies are determined by information that has already been computed. In the simulations described here, five quadrature points are used for all integrals.

For a time-stepping method, this code uses the two-level method of Higdon ([16], [17]), as outlined in Section 7.1. For the nonlinear case implemented here, the prediction of momentum involves both $p_b u$ and $p_b v$. The momentum equations include momentum advection terms that are not present in the linear case discussed in Sections 6 and 7; in equations (21) and (24), these are the terms involving $u(p_b u)$ and $u(p_b v)$, respectively. During the computations, the momentum advection is evaluated at time t_n during both the prediction and correction steps, and the boundary terms for the momentum advection are evaluated with the Lax-Friedrichs flux.

8.2. Test #1: Wave propagation at low resolution

The analyses of dispersion relations in Sections 6.4 and 7.1 indicate that the DG spatial discretization produces very little error in group velocity and phase velocity for the physical inertia-gravity modes, and in addition the numerical dissipation for such modes is very low. These remarks apply even when the dimensionless wavenumber $k\Delta x$ is large, i.e., at low resolution, and they apply both to the time-continuous case and to the case where the system is discretized in time with the two-level time stepping method described in Section 7.1. The purpose of the present set of computations is to test the ability of the DG method to propagate an inertia-gravity wave at low resolution and to compare the results with those produced with centered second-order finite differences on the B-grid and the C-grid.

For this test, the spatial interval has the form $-x_{max} \le x \le x_{max}$, where $x_{max} = 10,000$ km. The initial state consists of a localized pulse centered at $x = 0$ at time $t = 0$, and the resulting waves propagate both to the left and to the right for $t > 0$. The graphs displayed below show only the interval $0 \le x \le x_{max}$, and for purposes of these observations the test consists of initiating a signal at $x = 0$ and then observing its propagation to the right.

The fluid domain has constant depth, and the gravity wave speed in the nondispersive limit is chosen to be $c = 1$ m/sec. This speed is within the typical range for internal waves (Section 6.3). In general, $c = \sqrt{gh}$ for the shallow water equations, where g is the acceleration due to gravity and h is the thickness of the fluid layer. The present computations use the arbitrary value $h = 100$ meters, and the corresponding value of g is then 0.01 m/sec^2. The effect of this procedure is to produce a weak restoring force so that the inertia-gravity waves move slowly. The small value of g is a "reduced gravity" that is sometimes used in studies of internal waves (Cushman-Roisin [8], Gill [11]). Equivalently, one could use the physical value $g \approx 9.8$ m/sec^2 and use a small "equivalent depth" h that produces the desired value of c.

In these tests the Coriolis parameter is the constant value $f = 10^{-4}$ sec^{-1}, which is a representative value for the mid-latitudes. The Rossby radius for the present wave motion is then $c/f = 10$ km.

The spatial interval is divided into grid cells having equal width Δx, with several different values of Δx being used here. The largest of these is $\Delta x_0 = 40$ km. For that value, the ratio R of Rossby radius to cell width is $R = 0.25$.

According to the discussions in Sections 6.3 and 6.4, the C-grid is not expected to work well for this value of R.

In order to define the initial conditions, let $L = 4\Delta x_0$ and $M = 16\Delta x_0$. Also let $k_0 = 2\pi/L$ denote the wavenumber corresponding to wavelength L. The initial condition for the x-component u of velocity is

$$u(x,0) = u_0(x) = A_0 e^{-(x/M)^2} \sin k_0 x, \tag{87}$$

where $A_0 = 0.01$ m/sec. At time $t = 0$, the perturbation in the elevation of the free surface is set to zero, as is the y-component v of velocity. The amplitude A_0 is chosen to have a small value so that the dynamics of the solution are nearly linear, which enables a comparison with the results of simple implementations of the linearized shallow water equations on the B- and C-grids.

Numerical experiments show that the solution breaks cleanly into left-going and right-going wave packets from $x = 0$. (Other experiments show that if the initial pulse consists of a perturbation in the free-surface elevation instead of u, then much of the energy resides in the stationary mode $\omega = 0$, and not much energy is propagated.) The Fourier transform of the function u_0 in (87) is a Gaussian centered at wavenumber k_0. According to the dispersion relation (75) for exact solutions of the linearized shallow water equations, the group velocity corresponding to wavenumber k_0 in the right-going part of the exact solution is

$$\frac{d\omega}{dk}(k_0) = \frac{k_0 c^2}{\sqrt{c^2 k_0^2 + f^2}}. \tag{88}$$

Since the solution is localized in wavenumber space, the solution is a wave packet that travels with the group velocity (88). This velocity will be used below to determine the location of the packet at a specified time and thereby assess the accuracy of numerical solutions.

Because of the choice of the wavenumber k_0, the factor $\sin k_0 x$ in (87) has wavelength $L = 4\Delta x_0$, or four grid cells when the grid spacing is Δx_0. For a finite difference method, this is a low resolution; here, we test how the DG discretization performs at this resolution.

However, a fair comparison of the two types of methods should acknowledge that a DG method uses multiple degrees of freedom for each dependent variable in each grid cell, whereas a finite difference method uses only one degree of freedom for each dependent variable in each cell. A DG method with cell size Δx_0 should therefore be compared with finite difference methods that use smaller values of Δx. A comparison of methods should also account for different restrictions on the allowable time increment Δt.

In these experiments, solutions were computed to time $T = 20,000,000$ seconds, or about 231 days. According to the value of the group velocity (88), the wave packet should travel a distance of approximately 7310 kilometers over that time interval.

Figure 7 shows the results obtained with the DG spatial discretization with piecewise quadratic polynomials. For the computation shown in the upper frame, the grid spacing is $\Delta x_0 = 40$ km, and the time increment is $\Delta t =$

6400 sec. Since $c = 1$ m/sec, the corresponding Courant number is $\nu = c\Delta t/\Delta x = 0.16$, which is the approximate upper bound listed in Table 1 for the two-level time-stepping method with $N = 2$. The plot indicates that the location of the wave packet is essentially equal to the location found in the exact solution.

For a check on this solution, an additional computation was run with Δx and Δt cut in half. Again, the location of the wave packet is essentially exact. The slight increase in amplitude can be attributed to the fact that the dimensionless wavenumbers $k\Delta x$ for this packet are cut in half, and this reduces the slight numerical dissipation that is found with the DG method.

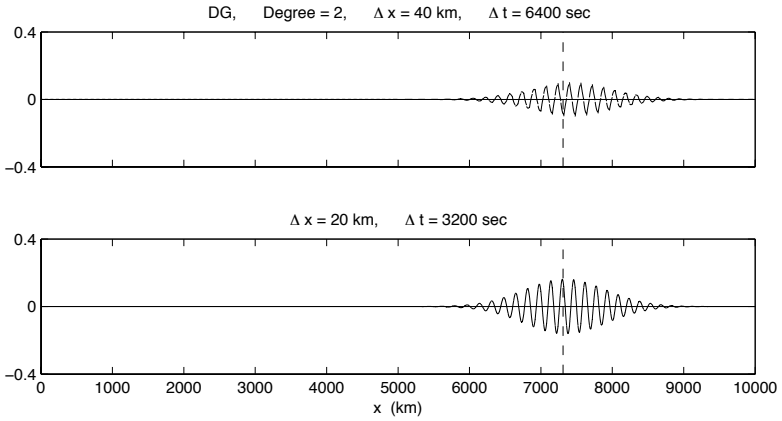

Figure 7: Propagation of a wave packet with a DG spatial discretization that uses piecewise quadratic approximations. In each frame, the plotted curve shows the elevation of the free surface, in meters, and the vertical dashed line indicates the location of the center of the wave packet in the exact solution, as determined by the group velocity (88). For the computation shown in the upper plot, there are four grid cells per wavelength, and for the lower plot there are eight grid cells per wavelength.

Figure 8 shows the results obtained with centered second-order finite differences on a C-grid. The time-stepping method used here is the same two-level method that was used for the DG computations. In this case, the Courant number must satisfy $\nu < 1$ in order for the method to be stable (Higdon [16]); this condition is much less restrictive than the ones encountered with the DG spatial discretizations. The computations shown here use Courant number $\nu = 0.8$.

For the solution shown in the top frame of Figure 8, the grid size is $\Delta x_0 = 40$ km, which is the same size used in the DG solution shown in the upper frame of Figure 7. In this case the ratio R of Rossby radius to grid size is 0.25, and the solution computed with the C-grid is highly inaccurate. For the middle frame of Figure 8, the grid size and time step are cut in half, and for the bottom frame these quantities are reduced by a factor of 5. For the last case, the time step is $\Delta t = 6400$ sec, which is the same value used for the DG solution in the upper frame of Figure 7. In this case, the location of the wave packet in the C-grid

solution is still not nearly as accurate as in that DG solution.

In addition, the C-grid solution in the bottom frame requires the computation of more dependent variables than does the DG solution. Compared to the C-grid computation with grid spacing $\Delta x_0 = 40$ km, the C-grid solution with $\Delta x = 8$ km requires the computation of 5 times as many quantities per time step. On the other hand, the DG solution with quadratic approximations in the upper frame of Figure 7 requires the computation of only 3 times as many quantities per time step.

The DG solution in the top frame of Figure 7 contains a small amplitude error, due to a small amount of numerical dissipation that is inherent in that method. Such an amplitude error appears not to be present in the C-grid solution in the bottom frame of Figure 8. However, the numerical algorithm used here for the linearized equations on the C-grid is completely nondissipative, whereas any algorithm used in an ocean model would contain some dissipation due to the usage of numerical advection schemes. A comparison of wave amplitudes in the present test would therefore be misleading.

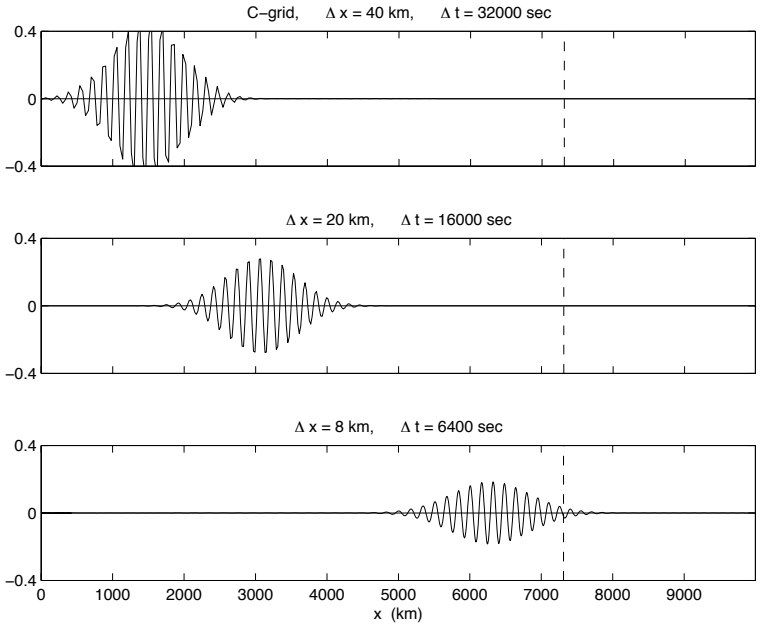

Figure 8: Propagation of a wave packet with second-order finite differences on a C-grid, for a sequence of decreasing values of Δx and Δt. The plotting format is the same as in Figure 7. Compared to the DG solution shown in the upper frame in Figure 7, the C-grid solution shown in the top frame of the present figure uses the same grid size but a larger value of Δt. The C-grid solution shown in the bottom frame uses the same Δt but a smaller grid size; it requires the computation of 5/3 times as many dependent variables and produces a less accurate location of the wave packet. A comparison of wave amplitudes is of limited value here, as the simple code used for this C-grid computation does not include a numerical advection scheme and is entirely nondissipative.

Figure 9 shows the results of some computations in which the B-grid is used in place of the C-grid. The top two frames in that Figure use the same values of Δx and Δt as in the top two frames of Figure 8. Again, the location of the wave packet is not as accurate as the location produced by the DG method. For this grid, an accurate location of the wave packet is obtained by reducing the original Δx and Δt by a factor of 4 instead of 5, and the results are shown in the bottom frame. Compared to the DG solution in the top frame of Figure 7, this computation produces 4/3 as many unknowns per time step and uses 4/5 as many steps.

The results shown in Figure 9 suggest that the B-grid may not be at a disadvantage relative to the DG method, in this particular test. However, this quasi-one-dimensional setting does not fully display the deficiencies of the B-grid in propagating inertia-gravity waves. It was mentioned in Section 6.3 that the B-grid is generally less accurate than the C-grid when the ratio R of Rossby radius to grid size is greater than 1. However, some contour plots of regions of error in Figure 4 of Dukowicz [10] indicate that the B-grid and C-grid give similar results for wave propagation parallel to the coordinate axes. Instead, the real deficiencies of the B-grid are displayed with waves that propagate in oblique directions, and such waves are not present in the quasi-one-dimensional configuration considered here. A related remark is that, as noted in Section 6.2, the spatial averaging that is normally required on the B-grid is not actually used in the quasi-one-dimensional case.

8.3. Test #2: Well-balanced pressure forcing

For the nonlinear shallow water equations in an infinite straight channel in a rotating reference frame, a weak form of the u-momentum equation is given in equation (21). The pressure forcing for this equation is given by the term $\Pi_u(j, \psi)$, which is specified in (22). Section 5.3 describes a method for implementing this pressure term in the presence of sloping and discontinuous bottom topography, and Section 5.4 gives a proof that this forcing is well-balanced. Here we describe a numerical experiment that illustrates this result.

For this computation, the spatial interval has the form $0 \le x \le x_{max}$, where $x_{max} = 500$ km. This interval is partitioned into 50 grid cells of equal width 10 km. The bottom topography of the channel is assumed to have a trapezoidal cross-section, as illustrated in the upper frame of Figure 10. On each end of the interval, the sloping portion of the topography occupies one-fourth the total width of the interval. On the sloping portions, the elevation z_{bot} of the bottom topography is discontinuous across cell edges, and within each of those cells the slope is half the slope that would be used if z_{bot} were to be linear and continuous. The small number of grid cells was chosen here so that the discontinuous nature of the bottom topography would be readily visible in the plot.

The depth of the channel in the middle is 1000 meters, and the acceleration due to gravity was set to a physical value of $g = 9.81$ m/sec^2. For a layer thickness $h = 1000$ m, the speed of gravity waves in the nondispersive limit is then $c = \sqrt{gh} \approx 99$ m/sec. The time step was chosen to be 16 sec, so the Courant number is $\nu = c\Delta t/\Delta x \approx 0.158$. This is slightly less than the

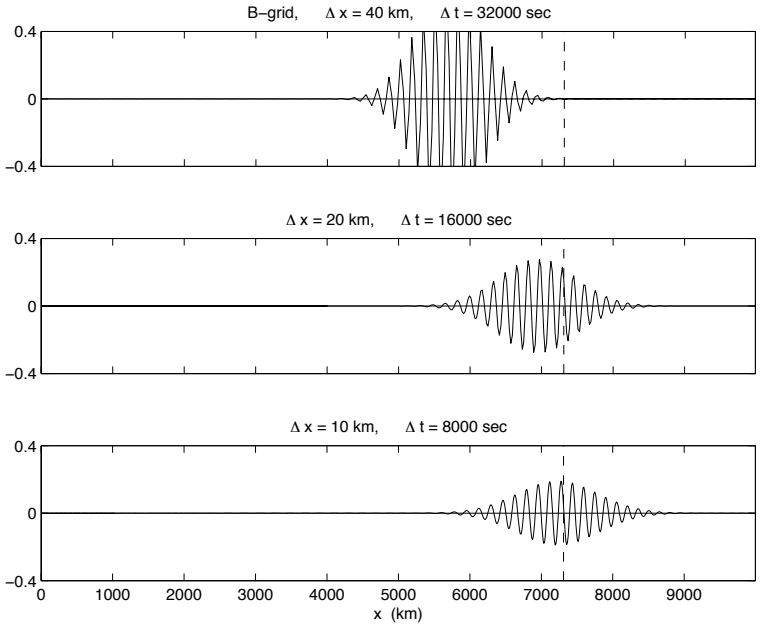

Figure 9: Propagation of a wave packet with second-order finite differences on a B-grid. For the computation shown in the bottom frame, the number of dependent variables computed is similar to the number computed for the DG solution in the top frame of Figure 7. Again, a comparison of wave amplitudes is of limited value. As noted in the text, the quasi-one-dimensional case considered here does not display the full deficiencies of the B-grid for inertia-gravity waves.

upper bound stated in Table 1 for the case $N = 2$, i.e., for piecewise quadratic approximations to the solution. Such approximations were used in the present computation. The Coriolis parameter was defined to be the constant value $f = 10^{-4}$ sec^{-1}.

For one simple test of well-balancing, the system was initialized to a rest state consisting of a level free surface and zero velocity, and the algorithm was executed for 10^6 time steps with zero wind forcing. In the computed solution, the system remained at rest, as expected.

Another test was obtained by initializing the system to a geostrophically-balanced state, which was then maintained by the algorithm over a long time. To define such a state, assume that all time derivatives are zero, the cross-channel velocity u is zero, and the wind and bottom stresses are zero. In this case, the pointwise form (29) of the u-momentum equation reduces to

$$-fv = -g\frac{\partial z_{top}}{\partial x}, \tag{89}$$

and the v-momentum equation (23) and the mass equation (25) both reduce to $0 = 0$. The free-surface elevation z_{top} was initialized to be continuous and

piecewise linear, with mean zero, a slope of 1 m / 50 km in the middle 100 km of the interval, and constant elevation ± 1 meter elsewhere. Equation (89) was then used to initialize the along-channel velocity v. The resulting v is piecewise constant, with $v = (1/50,000)g/f = 1.962$ m/sec on the subinterval where z_{top} varies, and $v = 0$ elsewhere.

The algorithm was run for 10^6 time steps. The resulting elevation of the free surface is shown in the lower frame of Figure 10. This elevation is essentially equal to the initial state, so the geostrophic balance is maintained in this simulation. In this system, the only physical processes present are the Coriolis effect due to the rotating reference frame and the horizontal pressure forcing due to the variation in the elevation of the free surface. As predicted by the analysis in Section 5.4, the numerical results do not indicate any spurious forcing due to the variable and discontinuous bottom topography.

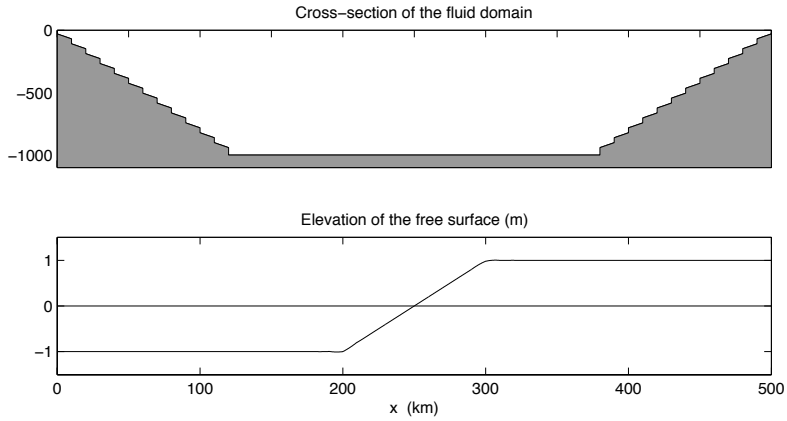

Figure 10: A test of well-balanced pressure forcing. The system is initialized to a geostrophically-balanced state consisting of a piecewise linear free-surface elevation and a corresponding component of velocity along the direction of the channel (i.e., into the page). The lower frame shows the free-surface elevation after 10^6 time steps. The elevation shown is essentially equal to the elevation at the initial time, so the geostrophic balance is maintained. In this system, the only physical processes present are the Coriolis effect due to the rotating reference frame and the horizontal pressure force due to variations in the elevation of the free surface. The sloping and discontinuous bottom topography shown in the upper frame does not contribute any spurious forcing to the system.

9. Summary

This paper is part of a longer-term study of the application of discontinuous Galerkin methods to the numerical modeling of ocean circulation. One step taken here is to develop a weak integral formulation of the lateral pressure forcing in a three-dimensional hydrostatic fluid that is described by an arbitrary,

generalized vertical coordinate. Such a formulation is suitable for a discontinuous Galerkin numerical method, and it avoids some known problems with the pressure forcing in a generalized coordinate.

We then begin an analysis of this approach by considering a hydrostatic fluid of constant density in a quasi-one-dimensional setting consisting of a flow in an infinite straight channel in a rotating reference frame. One issue is the practical implementation of the pressure forcing derived here, as one needs values of mass variables at cell edges, but in the case of a DG method those quantities are discontinuous at cell edges. The method developed here is applicable to a fluid domain in which the bottom topography is variable within grid cells and discontinuous across cell edges. The pressure forcing is shown to be well-balanced, in this setting. The well-balancing is also illustrated with a numerical experiment.

In addition, an analysis of numerical dispersion relations in the linearized case indicates that this DG discretization is much more accurate than second-order finite difference approximations on the B- and C-grids, which are widely used in ocean modeling. A related analysis demonstrates the stability of some time-stepping schemes, subject to restrictions on the Courant number. In a simple numerical experiment in which the DG method and the B- and C-grids are equally applicable, the additional accuracy of the DG approach can more than offset its disadvantages, which result from a reduced time step and the need to compute multiple degrees of freedom.

Section 4 outlines some other issues that are the subject of continuing work.

10. Acknowledgment

I thank Clint Dawson for numerous conversations that provided valuable background on the theory and practice of discontinuous Galerkin methods.

References

[1] A. Adcroft, R. Hallberg, M. Harrison, A finite volume discretization of the pressure gradient force using analytic integration, Ocean Modelling 22 (2008) 106–113.

[2] M. Ainsworth, Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods, J. Comput. Phys. 198 (2004) 106–130.

[3] A. Arakawa, V.R. Lamb, Computational design of the basic dynamical processes of the UCLA general circulation model, Methods in Computational Physics 17 (1977) 173–265.

[4] U.M. Ascher, L.R. Petzold, Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations, SIAM, Philadelphia, 1998.

[5] P.E. Bernard, E. Deleersnijder, V. Legat, J.F. Remacle, Dispersion analysis of discontinuous Galerkin schemes applied to Poincaré, Kelvin, and Rossby waves, J. Sci. Comput. 34 (2008) 26–47.

[6] R. Bleck, An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, Ocean Modelling 4 (2002) 55–88.

[7] B. Cockburn, C.W. Shu, Runge-Kutta discontinuous Galerkin methods for convection-dominated problems, J. Sci. Comput. 16 (2001) 173–261.

[8] B. Cushman-Roisin, Introduction to Geophysical Fluid Dynamics, Prentice-Hall, Englewood Cliffs, NJ, 1994.

[9] C. Dawson, E.J. Kubatko, J.J. Westerink, C. Trahan, C. Mirabito, C. Michoski, N. Panda, Discontinuous Galerkin methods for modeling hurricane storm surge, Advances in Water Resources 34 (2011) 1165–1176.

[10] J.K. Dukowicz, Mesh effects for Rossby waves, J. Comput. Phys. 119 (1995) 188–194.

[11] A.E. Gill, Atmosphere-Ocean Dynamics, Academic Press, San Diego, 1982.

[12] F.X. Giraldo, T. Warburton, A high-order triangular discontinuous Galerkin oceanic shallow water model, Int. J. Numer. Meth. Fluids 56 (2008) 899–925.

[13] S. Gottlieb, D.I. Ketcheson, C.W. Shu, High order strong stability preserving time discretizations, J. Sci. Comput. 38 (2009) 251–289.

[14] S.M. Griffies, Fundamentals of Ocean Climate Models, Princeton University Press, Princeton, N.J., 2004.

[15] W. Guo, X. Zhong, J.M. Qiu, Superconvergence of discontinuous Galerkin and local discontinuous Galerkin methods: Eigen-structure analysis based on Fourier approach, J. Comput. Phys. 235 (2013) 458–485.

[16] R.L. Higdon, A two-level time-stepping method for layered ocean circulation models, J. Comput. Phys. 177 (2002) 59–94.

[17] R.L. Higdon, A two-level time-stepping method for layered ocean circulation models: further development and testing, J. Comput. Phys. 206 (2005) 463–504.

[18] R.L. Higdon, Numerical modelling of ocean circulation, Acta Numerica 15 (2006) 385–470.

[19] R.L. Higdon, Physical and computational issues in the numerical modeling of ocean circulation, in: C. Dawson, M. Gerritsen (Eds.), Computational Challenges in the Geosciences, Springer, 2013, in press.

[20] T. Kärnä, V. Legat, E. Deleersnijder, A baroclinic discontinuous Galerkin finite element model for coastal flows, Ocean Modelling 61 (2013) 1–20.

[21] E.J. Kubatko, C. Dawson, J.J. Westerink, Time step restrictions for Runge-Kutta discontinuous Galerkin methods on triangular grids, J. Comput. Phys. 227 (2008) 9697–9710.

[22] E.J. Kubatko, J.J. Westerink, C. Dawson, hp discontinuous Galerkin methods for advection dominated problems in shallow water flow, Comput. Methods Appl. Mech. Engrg. 196 (2006) 437–451.

[23] R.J. LeVeque, D.L. George, M.J. Berger, Tsunami modelling with adaptively refined finite volume methods, Acta Numerica 20 (2011) 211–289.

[24] R.D. Nair, H.W. Choi, H.M. Tufo, Computational aspects of a scalable high-order discontinuous Galerkin atmospheric dynamical core, Computers and Fluids 38 (2009) 309–319.

[25] R.D. Nair, S.J. Thomas, R.D. Loft, A discontinuous Galerkin global shallow water model, Monthly Weather Review 133 (2005) 876–888.

[26] T.D. Ringler, J. Thuburn, J.B. Klemp, W.C. Skamarock, A unified approach to energy conservation and potential vorticity dynamics for arbitrarily structured C-grids, J. Comput. Phys. 229 (2010) 3065–3090.

[27] R.J. Spiteri, S.J. Ruuth, A new class of optimal high-order strong-stability-preserving time discretization methods, SIAM J. Numer. Anal. 40 (2002) 469–491.

[28] Y. Xing, X. Zhang, C.W. Shu, Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations, Advances in Water Resources 33 (2010) 1476–1493.

Appendix A. Two horizontal dimensions and curvilinear coordinates

Section 3 contains a derivation of the weak forms of the momentum and mass equations for the case where the horizontal dependence is quasi-one-dimensional. For future reference, we now generalize the derivations in Section 3 to the case of two horizontal dimensions.

Consider a hydrostatic and stratified fluid on a rotating spheroid, and assume that the horizontal coordinates are arbitrary orthogonal curvilinear coordinates. The vertical coordinate is a generalized coordinate s, as described in Sections 2 and 3. Partial differential equations for the conservation of mass, momentum, and tracers in this case are derived in [18]. Here we derive weak forms for the momentum and mass equations in this setting.

The present discussion uses notation similar to that used in [18], and more details can be found there. Denote the rotating spheroid by Σ, and assume that all or part of Σ is parameterized with coordinates $\mathbf{x} = (x_1, x_2)$. These

coordinates need not have units of length; for example, x_1 and x_2 could be angles that represent latitude and longitude, although such coordinates would not be used in a global ocean model due to the convergence of coordinate lines at the north pole. Let m_1 and m_2 be metric coefficients corresponding to x_1 and x_2, respectively, so that $m_1(\mathbf{x})dx_1$ and $m_2(\mathbf{x})dx_2$ represent elements of length along the surface Σ. Also let $\mathbf{i}(\mathbf{x},t)$ and $\mathbf{j}(\mathbf{x},t)$ denote unit vectors tangent to Σ in the directions of increasing x_1 and x_2, respectively, and assume $\mathbf{i}\cdot\mathbf{j}=0$. The (Eulerian) horizontal velocity of a fluid on the spheroid Σ is $\mathbf{u}(\mathbf{x},s,t) = u(\mathbf{x},s,t)\mathbf{i}(\mathbf{x},t) + v(\mathbf{x},s,t)\mathbf{j}(\mathbf{x},t)$, where $u = m_1\dot{x}_1$ and $v = m_2\dot{x}_2$.

For the sake of integration, let $E(\mathbf{x}) = m_1(\mathbf{x})m_2(\mathbf{x})$; since the coordinates are orthogonal, the element of area on Σ is then $E(\mathbf{x})d\mathbf{x} = m_1m_2dx_1dx_2$. Let D denote a region on Σ, and let $\tilde{D}$ denote its parameterization in terms of the coordinates $\mathbf{x} = (x_1, x_2)$. The regions $\tilde{D}$ and D can be regarded as a parameter domain and a physical domain, respectively.

The u-component of the momentum equation, from [18], is a generalization of equation (2) and can be written in the form

$$\begin{aligned} &\frac{\partial}{\partial t}\big[u(-p_s)\big] \;+\; \mathrm{div}\big[\mathbf{u}u(-p_s)\big] \;+\; \frac{\partial}{\partial s}\Big[\dot{s}u(-p_s)\Big] \;-\; \tilde{f}v(-p_s) \\ =\;& -\frac{1}{m_1}\frac{\partial P}{\partial x_1}\big(\mathbf{x}, z(\mathbf{x},s,t), t\big)\, gz_s \;+\; g\frac{\partial \tau_u}{\partial s}, \end{aligned} \tag{A.1}$$

Here,

$$\tilde{f}(\mathbf{x}) \;=\; f(\mathbf{x}) \;+\; \mathbf{j}\cdot\left(u\frac{1}{m_1}\frac{\partial \mathbf{i}}{\partial x_1} + v\frac{1}{m_2}\frac{\partial \mathbf{i}}{\partial x_2}\right), \tag{A.2}$$

where $f(\mathbf{x})$ is the Coriolis parameter; the remaining terms in (A.2) relate to the curvature of the spheroid Σ and the nature of its parameterization. The v-component of the momentum equation is an analogue of (A.1). The equation for conservation of mass is

$$\frac{\partial}{\partial t}\big(-p_s\big) \;+\; \mathrm{div}\big[\mathbf{u}(-p_s)\big] \;+\; \frac{\partial}{\partial s}\Big[\dot{s}(-p_s)\Big] \;=\; 0. \tag{A.3}$$

In equations (A.1) and (A.3), the symbol "div" denotes the divergence in two horizontal dimensions in curvilinear coordinates. For a vector-valued function on the parameter domain $\tilde{D}$ having the form $\mathbf{F}(\mathbf{x}) = F_1(\mathbf{x})\mathbf{i} + F_2(\mathbf{x})\mathbf{j}$, this quantity is defined by

$$\mathrm{div}\,\mathbf{F} = \frac{1}{m_1m_2}\left[\frac{\partial}{\partial x_1}\big(m_2F_1\big) + \frac{\partial}{\partial x_2}\big(m_1F_2\big)\right].$$

The action of this quantity on the physical domain D is obtained by multiplying by the area element $E(\mathbf{x})d\mathbf{x} = m_1m_2dx_1dx_2$ to obtain the divergence theorem

$$\int_{\tilde{D}} (\mathrm{div}\,\mathbf{F})\; E(\mathbf{x})d\mathbf{x} \;=\; \int_{\partial D} \mathbf{F}\cdot\mathbf{n}\; dS, \tag{A.4}$$

assuming that the function $\mathbf{F}$ is sufficiently smooth and the boundaries of $\tilde{D}$ and D are piecewise smooth. The right side of (A.4) is an integral on the boundary

∂D of the physical domain D, with values of $\mathbf{F}$ taken from the corresponding points on $\partial\tilde{D}$. The symbol $\mathbf{n}$ denotes the unit outward normal vector to ∂D, and dS refers to arclength along ∂D. If ϕ is a scalar-valued function on $\tilde{D}$, then (A.4) implies

$$\int_{\partial D} \phi \mathbf{F}\cdot\mathbf{n}\; dS \;=\; \int_{\tilde{D}} \phi(\operatorname{div}\mathbf{F})\; E(\mathbf{x})d\mathbf{x} + \int_{\tilde{D}} \mathbf{F}\cdot\nabla\phi\; E(\mathbf{x})d\mathbf{x}, \tag{A.5}$$

where

$$\nabla\phi = \left(\frac{1}{m_1}\frac{\partial\phi}{\partial x_1}\right)\mathbf{i} + \left(\frac{1}{m_2}\frac{\partial\phi}{\partial x_2}\right)\mathbf{j}.$$

Equation (A.5) is used below during integrations by parts.

Now regard the regions $\tilde{D}$ and D as grid cells in parameter space and physical space, respectively. Partition the fluid domain vertically with coordinate surfaces defined by $s = s_0, s_1, \ldots, s_R$, with $s_0 > s_1 > \ldots > s_R$, and let

$$\tilde{V}_r \;=\; \tilde{D}\times[s_r, s_{r-1}] \;=\; \{(\mathbf{x},s) : \mathbf{x}\in\tilde{D},\; s_r < s < s_{r-1}\}.$$

To obtain a weak form for the u-component of the momentum equation, multiply (A.1) by a test function ψ defined in $\tilde{D}$, and integrate on $\tilde{V}_r$ with respect to the measure $E(\mathbf{x})d\mathbf{x}ds$. As in Section 3.1, the test function is assumed to be independent of s, and the integration with respect to s produces quantities $u_r(\mathbf{x},t)$ and $\Delta p_r(\mathbf{x},t)$. For the divergence term, use (A.5) to integrate by parts. The resulting weak form is

$$\begin{aligned}
&\int_{\tilde{D}} \left\{ \frac{\partial}{\partial t}\left(u_r\Delta p_r\right) - \tilde{f}_r v_r \Delta p_r + \Big[\dot{s}u(-p_s)\Big]_{s=s_r}^{s=s_{r-1}} \right\} \psi(\mathbf{x})\; E(\mathbf{x})\, d\mathbf{x} \\
+ \;& \int_{\partial D} \Big[\mathbf{u}_r\left(u_r\Delta p_r\right)\Big]\cdot\mathbf{n}\;\psi\; dS \;-\; \int_{\tilde{D}} \Big[\mathbf{u}_r\left(u_r\Delta p_r\right)\Big]\cdot\nabla\psi\; E(\mathbf{x})\, d\mathbf{x} \\
= \;& \Pi_u(r,\psi) \;+\; g\int_{\tilde{D}} \Big\{(\tau_u)_{r-1}(\mathbf{x},t) - (\tau_u)_r(\mathbf{x},t)\Big\}\;\psi(\mathbf{x})\; E(\mathbf{x})\, d\mathbf{x},
\end{aligned} \tag{A.6}$$

where $\Pi_u(r,\psi)$ is the pressure term discussed below. Equation (A.6) is a generalization of the weak form (8) for the case that was developed in Section 3.1.

The pressure term in equation (A.6) is

$$\begin{aligned}
\Pi_u(r,\psi) \;&=\; -\int_{\tilde{D}} \frac{1}{m_1}\left[\int_{s_r}^{s_{r-1}} \frac{\partial P}{\partial x_1}\Big(\mathbf{x}, z(\mathbf{x},s,t), t\Big)\; g z_s\, ds\right]\psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x} \\
&=\; -g\int_{\tilde{D}} \frac{1}{m_1}\left[\int_{z_r(\mathbf{x},t)}^{z_{r-1}(\mathbf{x},t)} \frac{\partial P}{\partial x_1}(\mathbf{x},z,t)\; dz\right]\psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x}.
\end{aligned} \tag{A.7}$$

Here, $z_r(\mathbf{x},t) = z(\mathbf{x},s_r,t)$ and $z_{r-1}(\mathbf{x},t) = z(\mathbf{x},s_{r-1},t)$ denote the lower and upper elevations associated with the parameter domain $\tilde{V}_r$. In the first line of (A.7), the notation $\frac{\partial P}{\partial x_1}\big(\mathbf{x}, z(\mathbf{x},s,t), t\big)$ does not refer to a composite function, for which the chain rule would apply; instead, the function $\partial P/\partial x_1$ is evaluated at

the location $(\mathbf{x}, z(\mathbf{x}, s, t), t)$. The second line of (A.7) is obtained with a change of variable in the integration over s.

In analogy to (10), let

$$H_r(\mathbf{x}, t) \;=\; g \int_{z_r(\mathbf{x},t)}^{z_{r-1}(\mathbf{x},t)} P(\mathbf{x}, z, t)\; dz$$

for all $\mathbf{x} \in \tilde{D}$. Then, by using an analogue of (14), equation (A.7) can be written as

$$\begin{aligned}
\Pi_u(r, \psi) \;&=\; -\int_{\tilde{D}} \frac{1}{m_1(\mathbf{x})}\; \frac{\partial H_r}{\partial x_1}\; \psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x} \\
&+\; g \int_{\tilde{D}} p_{r-1}(\mathbf{x}, t)\; \frac{1}{m_1} \frac{\partial z_{r-1}}{\partial x_1}\; \psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x} \\
&-\; g \int_{\tilde{D}} p_r(\mathbf{x}, t)\; \frac{1}{m_1} \frac{\partial z_r}{\partial x_1}\; \psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x}, \qquad \text{(A.8)}
\end{aligned}$$

where $p_{r-1}(\mathbf{x}, t) = p(\mathbf{x}, s_{r-1}, t)$ and $p_r(\mathbf{x}, t) = p(\mathbf{x}, s_r, t)$ denote the pressures at the top and bottom of layer r, respectively. In the integral in the first term on the right side of (A.8), the derivative of H_r can be removed by observing

$$\begin{aligned}
&\int_{\tilde{D}} \frac{1}{m_1}\; \frac{\partial H_r}{\partial x_1}\; \psi(\mathbf{x})\, E(\mathbf{x})\, d\mathbf{x} \\
=\;& \int_{\tilde{D}} \nabla H_r \cdot \big(\psi(\mathbf{x})\mathbf{i}\big)\; E(\mathbf{x})\, d\mathbf{x} \\
=\;& \int_{\partial D} H_r\, \psi\; \mathbf{i} \cdot \mathbf{n}\; dS \;\; - \;\; \int_{\tilde{D}} H_r\; \mathrm{div}\big(\psi(\mathbf{x})\mathbf{i}\big)\; E(\mathbf{x})\, d\mathbf{x} \\
=\;& \int_{\partial D} H_r\, \psi\; \mathbf{i} \cdot \mathbf{n}\; dS \;\; - \;\; \int_{\tilde{D}} H_r \frac{\partial}{\partial x_1} \Big(m_2(\mathbf{x})\psi(\mathbf{x})\Big)\, d\mathbf{x}. \qquad \text{(A.9)}
\end{aligned}$$

The weak form of the u-component of the momentum equation is then equation (A.6), with the pressure term $\Pi_u(r, \psi)$ obtained by substituting the last line of (A.9) into the first term on the right side of (A.8).

The weak form of the mass equation (A.3) is obtained with a derivation that is similar to that of the momentum equation, except that it is less complicated, and the result is

$$\begin{aligned}
&\int_{\tilde{D}} \left\{ \frac{\partial}{\partial t}\, (\Delta p_r) + \Big[\dot{s}(-p_s)\Big]_{s=s_r}^{s=s_{r-1}} \right\}\; \psi(\mathbf{x})\; E(\mathbf{x})\, d\mathbf{x} \\
+\;& \int_{\partial D} (\mathbf{u}_r \Delta p_r) \cdot \mathbf{n}\; \psi\; dS \;\; - \;\; \int_{\tilde{D}} (\mathbf{u}_r \Delta p_r) \cdot \nabla\psi\; E(\mathbf{x})\, d\mathbf{x} \\
=\;& 0.
\end{aligned}$$