CONTINUOUS AND DISCRETE BAROCLINIC MODES IN CONTINUOUSLY VARYING STRATIFICATION*

RACHEL ROBEY† AND IAN GROOMS†

Abstract. We study the behavior of baroclinic modes in a continuously stratified fluid and their discrete representation in a layer model. The modes are shown to rapidly approach simple sinusoidal behavior under a Charney-coordinate transform. We propose a corresponding grid scheme to near-optimally preserve the oscillating structure in the discrete modes. The discrete modal representation and analysis are relevant for quasi-geostrophic models and are also shown to apply to the primitive equations using a common discretization scheme.

Key words. geophysical fluid dynamics, quasi-geostrophic, primitive equation, baroclinic modes, vertical grid

MSC codes. 86A05, 76U05, 76F45, 76M45, 76M99

1. Introduction.

This paper is concerned with the baroclinic modal structures that arise in continuously stratified fluids and the discrete representations of the modes and dynamics in the context of ocean models.

The baroclinic modes arise as the eigenfunctions of the Sturm-Liouville problem,

$$\frac{d}{dz}\left(\frac{f^2}{N^2(z)}\frac{d}{dz}\phi_m\right)+\lambda_m^2\phi_m=0, \tag{1.1}$$

with boundary conditions $\frac{d}{dz}\phi_m(0)=0$, $\frac{d}{dz}\phi_m(-H)=0$, where $f=2\Omega\sin\theta_0$ is the Coriolis parameter at latitude θ_0. The buoyancy frequency, $N(z)$, is a measure of stratification given by

$$N(z)=\left(-\frac{g}{\rho_0}\frac{\partial\rho}{\partial z}\right)^{1/2}, \tag{1.2}$$

where ρ is the density, g is the acceleration due to gravity, and ρ_0 is a constant reference density, typically in the vicinity of $1,035$ kg m^{-3} for global ocean models. Consistent with stable and nonvanishing stratification, we take the buoyancy frequency to be positive and bounded away from zero, $N^2(z)\geq c_N>0$, for a positive constant, c_N.

Homogeneous Neumann boundary conditions reflect a rigid, flat-bottom geometry. Setting a homogeneous Dirichlet condition $\phi_m(-H)=0$ at the bottom produces a related set of modes, sometimes called 'surface modes' [6], which are appropriate to situations with a rough bottom boundary. The baroclinic modes appear in the analysis of the linearized Boussinesq, hydrostatic primitive equations underlying most large-scale ocean models as well as the classical quasi-geostrophic (QG) approximated system often used in studies of extratropical, mesoscale eddy dynamics. Both systems are briefly reviewed in section 2.

In stratified flow, the vertical dimension plays a unique role. For the ocean, it is often treated in geopotential or isopycnal coordinates. We introduce a family

*Submitted to the editors DATE.

Funding: This material is based upon work partially supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific 25 Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0021110, and partially supported by the U.S. National Science Foundation under Award Number 1912332.

†Department of Applied Mathematics, University of Colorado, Boulder, CO (rachel.robey@colorado.edu, ian.grooms@colorado.edu).

of coordinates (1.3) based on the buoyancy frequency, in which the strength of the dependency is scaled via a parameter α.

$$\xi_\alpha(z) = \frac{1}{\overline{\xi}_\alpha} \int_{-H}^{z} N^\alpha(s) ds. \tag{1.3}$$

Here, $\overline{\xi}_\alpha$ is a normalization constant such that $\xi_\alpha \in [0, 1]$ is dimensionless. Scaled versions of the geopotential and isopycnal coordinates are recovered with $\alpha = 0$ and $\alpha = 2$, respectively. Within this family, we make heavy use of the intermediate coordinate with $\alpha = 1$,

$$\xi_c(z) = \frac{1}{\overline{\xi}_c} \int_{-H}^{z} N(s) ds. \tag{1.4}$$

The subscript c acknowledges the introduction of this coordinate by J. Charney who presented it in differential form [4, eq. 19],

$$d\xi_c = \frac{N}{f} dz. \tag{1.5}$$

Known properties of the baroclinic modes for the continuous problem are reviewed in section 3 along with a discussion of characterizations, including the WKB approximation, and their applicability. We note the independent appearance of the stretched Charney coordinate, ξ_c, in the approximation and that under the coordinate transformation even low-order modes have near-constant wavenumber. Section 4 explores the form of the baroclinic modes under standard layer-based discretizations. We show that discrete modes display oscillatory behavior mirroring that of the continuous modes, among other similar properties.

Although the Charney coordinate has appeared incidentally across analytic treatments, it has not, to our knowledge, been leveraged to numerical advantage. Given the sinusoidal behavior of the continuous modes under the coordinate transform and constraints on the discrete modes, we propose and test a vertical grid with layers equispaced in the Charney coordinate. Section 5 presents results for a continuously varying oceanic stratification showing that a Charney coordinate grid is close to optimal in preserving the structure and non-linear interactions of the baroclinic modes.

2. Review of primitive and quasi-geostrophic systems.

The primitive equations model rotating stratified flow under the hydrostatic and Boussinesq approximations and are frequently used in oceanic modeling. Equations (2.1)-(2.4) comprise a standard expression of the system:

$$\partial_t \boldsymbol{u} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{u} + f\hat{\boldsymbol{z}} \times \boldsymbol{u} = -\frac{1}{\rho_0} \nabla p, \tag{2.1}$$

$$\nabla \cdot \boldsymbol{v} = 0, \tag{2.2}$$

$$\partial_z p = -\rho g, \tag{2.3}$$

$$\partial_t \rho + (\boldsymbol{v} \cdot \nabla)\rho = 0. \tag{2.4}$$

Here, $\boldsymbol{u} = (u, v)^T$ is the horizontal velocity and $\boldsymbol{v} = (u, v, w)^T$ is the three-dimensional velocity; f is the Coriolis parameter, p is the pressure perturbation, ρ is the density perturbation, and g is gravity. For simplicity, we omit viscous and diffusive terms and assume a linear equation of state and no diabatic effects.

The baroclinic modes, i.e. the normal vertical modes, arise from the linearization of the primitive equations around a state of rest and fixed stratification varying only in z, (e.g., [19, sect. 3.4.1]). The vertical and horizontal coordinates of the linearized system may be separated provided the vertical differential operator,

$$\frac{d}{dz}\left(\frac{1}{N^2(z)}\frac{d}{dz}\phi_m\right)+\frac{1}{c_m^2}\phi_m=0, \tag{2.5}$$

admits a basis of eigenfunctions, which constitute a natural set of vertical modes for the pressure and horizontal velocity. Taking $\partial_z\phi = 0$ at $z = -H, 0$ reflects stress-free boundary conditions, though other physically-meaningful boundary conditions can be applied [6]. Expanding $\boldsymbol{u}$ and p in terms of the eigenfunction basis (see section 3) evolution of mode m in the linearized system becomes,

$$\partial_t\frac{\hat{p}_m}{\rho_0}+c_m^2\nabla\cdot\hat{\boldsymbol{u}}_m=0. \tag{2.6}$$

The form clearly identifies c_m as a wave speed, which is associated with an m-th baroclinic mode gravity wave. The corresponding deformation radius, L_m, may be related to the wave speed by [10],

$$L_m=\sqrt{\frac{c_m^2}{f^2+2\beta c_m}}. \tag{2.7}$$

where $\beta = \partial f/\partial y$ is the meridional gradient of the Coriolis parameter.

From the primitive equations, a further simplified, quasi-geostrophic system may be obtained by asymptotic analysis [19]. The approximation is relevant for studies of dynamics in extratropical regions of the ocean, rich in mesoscale eddies. The QG system on a β-plane is given by,

$$\partial_t q+J[\psi,q]=0, \tag{2.8}$$

$$\nabla^2\psi+\partial_z\left(\frac{f^2}{N^2(z)}\partial_z\psi\right)+\beta y=q. \tag{2.9}$$

where q is the potential vorticity and ψ the streamfunction with $u = -\partial_y\psi$ and $v = \partial_x\psi$. The advection term is written using a Jacobian operator, $J[a, b] = \partial_x a\partial_y b - \partial_y a\partial_x b$. The vertical differential operator in (2.9) yields the vertical modes eigenproblem for the QG system (1.1); the modes correspond to those of the primitive equations (2.5) and eigenvalues reflect extratropical ($f \neq 0$, $f^2 \gg \beta c_m$) deformation radii,

$$\lambda_m^{-1}=L_m\approx\frac{c_m}{f}. \tag{2.10}$$

3. Continuous baroclinic modes. Well-developed Sturm-Liouville theory establishes a known structure in the modal solutions of (1.1), which are briefly reviewed [18, 19]. The vertical modes are eigenfunctions of the Sturm-Liouville problem and are guaranteed to form an infinite set of solutions, $\{\phi_m\}_{m=0}^{\infty}$, which comprise an orthonormal basis for $L^2([-H, 0])$. That is,

$$\int_{-H}^{0}\phi_n\phi_m dz=\delta_{nm}, \tag{3.1}$$

where δ_{nm} is a Kronecker delta and any L^2 function may be expanded in $\{\phi_m\}$. The eigenvalues are non-positive (preemptively denoted as $-\lambda_m^2$), distinct, countably infinite, and strictly increasing in magnitude to infinity.

It is trivial to confirm the existence of the constant, barotropic mode $\phi_0(z) = 1/\sqrt{H}$ with eigenvalue $\lambda_0 = 0$. The depth-varying, baroclinic modes ($m \geq 1$) are characterized by increasing oscillations such that ϕ_m has exactly m simple roots in $[-H, 0]$. Furthermore, due to Sturm's separation theorem [17, Lemma 5.21] the positions of the roots of subsequent solutions ϕ_m and ϕ_{m+1} interlace, i.e., between two roots of ϕ_{m+1} there must be a root of ϕ_m.

Important physical attributes due to the eigensystem include the vertical structure imparted by the modes; the deformation radius, L_m, associated with each mode (2.10); and the interaction between vertical modes.

The interaction is most straightforward to see in the QG system. Expanding the potential vorticity and streamfunction in the baroclinic mode basis, $\psi = \sum_{m=0}^{\infty} \hat{\psi}_m \phi_m(z)$ and $q = \sum_{m=0}^{\infty} \hat{q}_m \phi_m(z)$, the evolution equation for mode n becomes,

$$\partial_t \hat{q}_n + \sum_{\ell,m} \Theta_{\ell mn} J[\hat{\psi}_\ell, \hat{q}_m] + \beta \partial_x \hat{\psi}_n = 0. \tag{3.2}$$

Interactions between modes occur in triads, mediated by triple interaction coefficient,

$$\Theta_{\ell mn} = \int_{-H}^{0} \phi_\ell(z)\phi_m(z)\phi_n(z)dz. \tag{3.3}$$

The interaction describes the efficiency of the energetic interactions and transfers between vertical scales. The interaction coefficient is clearly invariant under permutations of the indices and interactions involving the barotropic mode are simple, i.e., $\Theta_{0mn} = \delta_{mn}/\sqrt{H}$ and similarly for $\Theta_{\ell 0n}$ and $\Theta_{\ell m0}$.

3.1. Approximations of modal behavior.

For a general stratification, $N(z)$, the eigensystem cannot be explicitly solved analytically. Approximations and reformulations can be useful to illuminate characteristics of the modes and deformation radii.

3.1.1. Asymptotic WKB approximation.

The Sturm-Liouville eigenproblem (1.1) lends itself to the asymptotic WKB approximation [3, sect. 10.1]. The method was employed by Chelton et al. [5, Appendix A] to estimate ocean vertical modes and the corresponding eigenvalues to gain insight into their behavior. The authors derive the approximate solution for the vertical velocity baroclinic modes; the corresponding form for the horizontal velocity baroclinic modes studied here is found in [7, Appendix B] by exploiting the relationship between the two kinds of baroclinic modes.

The WKB method may also be applied directly to the horizontal velocity baroclinic modes, as presented here. As noted, $\{\phi_m\}$ are solutions to a Sturm-Liouville problem, ensuring $\lambda_m \to \infty$ as $m \to \infty$. Let $S(z) = f^2/N^2(z)$ and take $\epsilon^2 = \lambda_m^{-2}$ in (1.1), yielding

$$-\epsilon^2 \frac{d}{dz}\left(S(z)\frac{d\phi_m}{dz}\right) = \phi_m. \tag{3.4}$$

Following the WKB method, we assume a solution of the form

$$\phi_m(z) = \exp[T(z)/\delta]. \tag{3.5}$$

Substituting (3.5) into (3.4) and simplifying yields,

$$-\frac{\epsilon^2}{\delta^2}\left[S(T')^2 + \delta(S'T' + ST''(z))\right] = 1. \tag{3.6}$$

In order to achieve dominant balance with the $\mathcal{O}(1)$ right-hand side, we choose $\delta = \epsilon$. Making this substitution and expanding an asymptotic series, $T(z) = T_0(z) + \epsilon T_1(z) + \mathcal{O}(\epsilon^2)$,

$$S[T_0']^2 + \epsilon[S'T_0' + ST_0'' + 2ST_0'T_1'] = -1 + \mathcal{O}(\epsilon^2). \tag{3.7}$$

Equating $\mathcal{O}(1)$ terms and recalling the definition of S, T_0 must satisfy the differential equation,

$$[T_0']^2 = -\frac{1}{S(z)} = -\frac{N^2(z)}{f^2}. \tag{3.8}$$

This yields a solution,

$$T_0 = \pm \mathrm{i} \int_0^z \frac{N(s)}{f} ds, \tag{3.9}$$

where without loss of generality we impose $T_0(0) = 0$. Equating the $\mathcal{O}(\epsilon)$ terms in (3.7) and using the definition for S and (3.9) for T_0,

$$2ST_0'\left(\frac{S'}{2S} + \frac{T_0''}{2T_0'} + T_1'\right) = 0. \tag{3.10}$$

Noting that $2ST_0'$ is not equivalently zero, we obtain the following differential equation for T_1,

$$T_1'(z) = -\frac{1}{2}\left(\frac{S'}{S} + \frac{T_0''}{T_0'}\right) = -\frac{1}{2}\frac{d}{dz}\ln(ST_0') = \frac{d}{dz}\ln\left[(1 \mp \mathrm{i})\sqrt{\frac{N(z)}{2f}}\right]. \tag{3.11}$$

The rearranged equation then simply yields,

$$T_1(z) = \ln\left[(1 \mp \mathrm{i})\sqrt{\frac{N(z)}{2f}}\right], \tag{3.12}$$

up to an additive constant that has no effect on the final form of the solution. Substituting (3.9) and (3.12) back into the ansatz (3.5),

$$\begin{aligned} \phi_m(z) &\approx \exp\left(\frac{1}{\epsilon}T_0 + T_1\right) \\ &= \exp\left(\pm\frac{\mathrm{i}}{\epsilon}\int_0^z \frac{N(s)}{f}ds\right)(1 \mp \mathrm{i})\sqrt{\frac{N(z)}{2f}}. \end{aligned} \tag{3.13}$$

Linear combinations of the two solutions yield a real form,

$$\phi_m(z) \approx N^{1/2}(z)\left[a\sin\left(\frac{1}{\epsilon}\int_0^z \frac{N(s)}{f}ds\right) + b\cos\left(\frac{1}{\epsilon}\int_0^z \frac{N(s)}{f}ds\right)\right]. \tag{3.14}$$

To simplify the application of the Neumann boundary conditions, we note that the derivative is dominated by the sines and cosines at order $\mathcal{O}(\epsilon^{-1})$, with the derivative of $N^{1/2}$ only $\mathcal{O}(1)$. The top boundary, $\phi'(0) = 0$, requires $a = 0$. The bottom boundary, $\phi'(-H) = 0$, recovers an approximation for λ_m,

$$\lambda_m = \frac{1}{\epsilon} \approx \frac{m\pi f}{\int_{-H}^{0} N(s)ds}. \tag{3.15}$$

Simplifying, we obtain the following approximation, which may be expressed in terms of the Charney coordinate, ξ_c, (1.4).

$$\phi_m(z) \approx N^{1/2}(z)\cos\left(\lambda_m \int_0^z \frac{N(s)}{f} ds\right) = N^{1/2}(z)\cos\left[m\pi\xi_c(z)\right]. \tag{3.16}$$

The small parameter for the asymptotic approximation depended only on the expected growth of λ_m; the modes are guaranteed to approach the cosine approximation as $m \to \infty$ for any stratification $N(z) \geq c_N > 0$. It is, however, the lowest modes (small m), outside the asymptotic regime, that are most studied and most dynamically relevant. In this case, a small parameter must be instead be identified with respect to the behavior of $N(z)$. In the derivation of the vertical velocity baroclinic modes, Chelton et al. [5] suggest conditions on N under which the approximation is valid: the scaled rate of change of N should be small compared to the scaled wavenumber. Given typical oceanic stratification (see, e.g., Figure 1), N does vary slowly in the bulk of the water column, but routinely violates such an assumption due to the strong stratification present in the pyncocline.

Despite the breakdown of the underlying assumptions, the WKB approximation has still empirically proven quite effective, and been used with some success, for low-order baroclinic modes (e.g., [5, 7, 16]). Of particular interest for our purposes is the introduction of the Charney coordinate (1.4) in the solution of T_0. The leading order solution for ϕ_m, (3.14), has constant wavenumber in the transformed coordinate, regardless of the particular linear boundary condition.

3.1.2. Liouville integral form. Identification of the stretched Charney coordinate using the WKB approach, however, hinges on having made the scaling assumptions and approximations, which we know may not universally hold. An alternate approach to characterizing the behavior of the baroclinic modes involves the reformulation of (1.1) into Liouville normal form [11, 14]. Expanding (1.1) into standard form,

$$\phi_m'' + \left(-2\frac{N'}{N}\right)\phi_m' + \lambda_m^2 \frac{N^2}{f^2}\phi_m = 0. \tag{3.17}$$

Following the Liouville transformation (see supplement SM1, [17, Problem 5.13]), we define a transformed mode, $\eta(\xi)$, with

$$\phi(z) = N^{1/2}(z)\eta(\xi(z)), \tag{3.18}$$

along with a stretched coordinate, ξ, which is exactly the Charney coordinate, ξ_c,

$$\xi(z) = \xi_c(z) = \frac{1}{\overline{\xi}} \int_{-H}^{z} N(t)dt, \qquad \overline{\xi} = \int_{-H}^{0} N(t)dt. \tag{3.19}$$

Under the transformation, the differential equation for the transformed mode, $\eta(\xi)$, takes the Liouville normal form,

$$\eta''(\xi) + \left[\kappa^2 - \gamma(\xi)\right]\eta(\xi) = 0, \qquad \xi \in (0,1), \tag{3.20}$$

$$\eta'(\xi) + \frac{\overline{\xi}}{2}\frac{\frac{dN}{dz}(\xi)}{N^2(\xi)}\eta(\xi) = 0, \qquad \xi = 0, 1. \tag{3.21}$$

For conciseness, we have defined

$$\gamma(\xi(z)) = \frac{\overline{\xi}^2}{4N^2}\left(\frac{5}{N^2}\left(\frac{dN}{dz}\right)^2 - \frac{2}{N}\frac{d^2N}{dz^2}\right). \tag{3.22}$$

The solutions to the transformed problem may be related to the original eigenpairs using (3.18) and $\kappa = \lambda_m \overline{\xi}/f$. It is guaranteed that $\kappa^2 \geq 0$ since $\overline{\xi}/f$ is a real constant and it is known $\lambda_m \geq 0$ from the original formulation.

The Liouville normal form is typically used in investigations of asymptotic behavior and convergence, but also yields, without approximation, an illuminating implicit integral form ([11], see supplement SM2)

$$\eta(\xi) = \cos(\kappa\xi) + \frac{c_0}{\kappa}\sin(\kappa\xi) + \frac{1}{\kappa}\int_0^{\xi}\gamma(s)\eta(s)\sin[\kappa(\xi - s)]ds, \tag{3.23}$$

where $c_0 = -\frac{1}{2}\overline{\xi}(\frac{dN}{dz}N^{-2})(-H)$ is a constant arising from the boundary conditions.

The asymptotic behavior as $\lambda_m \to \infty$ may be recovered with the corresponding limit $\kappa \to \infty$; the second two terms in (3.23) vanish and the first cosine term dominates. Transforming back into ϕ_m,

$$\phi_m(z) \approx N^{1/2}(z)\cos(\kappa\xi(z)). \tag{3.24}$$

As with the WKB approximation, the application of the Neumann boundary conditions cannot be exact but is simplified by noting that the $\mathcal{O}(\kappa)$ sine term dominates the derivative. No mutable parameters exist for the top $z = 0$ boundary. The bottom boundary implies $\kappa = m\pi$. Therefore the eigenfunction and eigenvalue may be estimated using,

$$\phi_m(z) \approx N^{1/2}\cos(m\pi\xi_c(z)), \qquad \lambda_m = \frac{\kappa f}{\overline{\xi}_c} \approx \frac{m\pi f}{\overline{\xi}_c}. \tag{3.25}$$

The approximated solutions correspond to the WKB result, (3.15) and (3.16).

Compared to the WKB method, the Liouville transformation gives rise to the Charney coordinate independent of any scaling assumptions. The form (3.23) shows, without approximation, that the baroclinic modes may be understood as sinusoids of constant wavenumber in the Charney coordinate, modified by an integral correction term. Without relying on substantial decay of the correction term as κ grows, it is sufficient for our purposes that that the convolution inherits the phase of $\sin(\kappa\xi)$.

4. Discrete baroclinic modes.

Numerical discretization in large-scale ocean models is typically built on a grid of vertical layers. The continuous baroclinic modes are not represented directly in the discrete system; rather, the discrete system gives rise to its own, distinct set of modes equal to the number of layers. The discretization and placement of the layers in the grid impact the fidelity of the baroclinic mode representation.

4.1. Quasi-geostrophic modes. As in the continuous problem, the discrete baroclinic modes are analogously defined by an eigenproblem, based on the discrete representation of the vertical differential operator. A classical numerical scheme for the QG system with continuously-varying stratification uses a second-order, centered finite-volume scheme (e.g., [9]). Under this approach, for a general grid with $\mathcal{N}$ layers of depth $h_{1,\ldots,\mathcal{N}}$, the discrete vertical difference operator, $\mathbf{L}$, is given by,

$$\partial_z \left(\frac{f^2(z)}{N^2(z)} \partial_z \phi \right) \approx (\mathbf{L}\phi)_i = \frac{f^2}{h_i} \left(N^{-2}_{i-\frac{1}{2}} \frac{\phi_{i-1} - \phi_i}{\frac{1}{2}(h_{i-1} + h_i)} - N^{-2}_{i+\frac{1}{2}} \frac{\phi_i - \phi_{i+1}}{\frac{1}{2}(h_i + h_{i+1})} \right), \tag{4.1}$$

where the indices follow the ocean convention counting down from the surface and $N^{-2}_{i\pm 1/2}$ is the evaluation of $N^{-2}(z)$ at the layer interfaces, $z_{i\pm 1/2}$. The homogeneous Neumann boundary conditions are treated by letting the exact derivative $\partial_z \phi_m = 0$ at the top and bottom take the place of the finite difference approximation. Thus,

$$\mathbf{L}\boldsymbol{\phi} + \lambda^2 \boldsymbol{\phi} = 0, \tag{4.2}$$

is the discrete analog to the continuous eigenproblem (1.1).

Note that the QG stretching matrix derived from the rotating shallow water equations (e.g., [19, sect. 5.3.2], [12, sect. 6.16]), i.e.,

$$(\mathbf{L}_{\text{layer}}\phi)_1 = f^2 \left(\frac{\phi_1}{g h_1} + \frac{\phi_1 - \phi_2}{g'_1 h_1} \right),$$

$$(\mathbf{L}_{\text{layer}}\phi)_i = f^2 \left(\frac{\phi_{i-1} - \phi_i}{g'_{i-1} h_i} + \frac{\phi_i - \phi_{i+1}}{g'_i h_i} \right), \tag{4.3}$$

with $g'_k = g(\rho_{k+1} - \rho_k)/\rho_1$ the 'reduced gravity', is isomorphic to the finite volume formulation (4.1) in the limit $g \gg g'_1$ (e.g. [19, sect. 5.4.6]).

4.2. Equivalence with primitive equation modes. To demonstrate that the analysis of the discrete baroclinic modes performed here has relevance outside the setting of quasi-geostrophic dynamics, we make a connection between the discrete baroclinic modes of the primitive equations and those of QG. The approach is to discretize the primitive equations in the vertical direction only, and then to study the normal modes of the discrete system, linearized around a state of rest.

There are many ways to discretize the vertical coordinate in the primitive equations, so we focus on an approach shared by two modern ocean models: The Model for Prediction Across Scales - Ocean (MPAS-O; [13]) and the Modular Ocean Model, version 6 (MOM6; [1]). Both of these models use an arbitrary Lagrangian-Eulerian (ALE) approach to the vertical coordinate.

We will analyze the purely Lagrangian limit of this discretization, given below

$$\partial_t \mathbf{u}_k + \mathbf{u}_k \cdot \nabla \mathbf{u}_k + f \hat{\boldsymbol{z}} \times \mathbf{u}_k = -\frac{1}{\rho_0} \nabla p_k - \frac{g \rho_k}{\rho_0} \nabla z_k, \tag{4.4}$$

$$\partial_t h_k + \nabla \cdot (h_k \mathbf{u}_k) = 0, \tag{4.5}$$

$$\partial_t (h_k \rho_k) + \nabla \cdot (h_k \mathbf{u}_k \rho_k) = 0. \tag{4.6}$$

These equations use a Boussinesq approximation and assume a linear equation of state and no diabatic effects. The index k denotes layers, and the ocean modeling convention is that k increases downwards. The thickness of the k^{th} layer is denoted h_k.

The fact that this is a purely Lagrangian limit is reflected in the fact that the thickness evolution equation includes no source or sink terms associated with transport between adjacent layers. The pressure p_k and height z_k are obtained from the layer density ρ_k and thickness h_k via

$$p_k = g\left[\frac{1}{2}\rho_k h_k + \sum_{n=1}^{k-1}\rho_n h_n\right], \tag{4.7}$$

$$z_k = \frac{1}{2}h_k + \sum_{n=k+1}^{\mathcal{N}} h_n + z_{\mathcal{N}}, \tag{4.8}$$

where $z_{\mathcal{N}}$ is the depth of the lower boundary.

To find the normal modes, linearize around a state of rest $\mathbf{u}_k = 0$, thicknesses $h_k = H_k$, and densities $\rho_k = R_k$, where H_k and R_k are assumed to be positive and independent of the horizontal coordinates and of time. The density stratification is also assumed to be stable, i.e. $R_k > R_{k-1}$ for $k > 2$. In order for this state to be an equilibrium of the governing system, the lower boundary depth $z_{\mathcal{N}}$ must be independent of the horizontal coordinates, i.e. the lower boundary must be flat. The linear perturbation equations are

$$\partial_t \mathbf{u}_k + f\hat{\boldsymbol{z}} \times \mathbf{u}_k = -\frac{1}{\rho_0}\nabla p_k - \frac{gR_k}{\rho_0}\nabla z_k, \tag{4.9}$$

$$\partial_t h_k + H_k \nabla \cdot \mathbf{u}_k = 0, \tag{4.10}$$

$$\partial_t \rho_k = 0, \tag{4.11}$$

where the pressure and height perturbations are

$$p_k = g\left[\frac{1}{2}\left(R_k h_k + H_k \rho_k\right) + \sum_{n=1}^{k-1}\left(R_n h_n + H_n \rho_n\right)\right], \tag{4.12}$$

$$z_k = \frac{1}{2}h_k + \sum_{n=k+1}^{N} h_n. \tag{4.13}$$

To separate the vertical and horizontal directions in the perturbation equations it is convenient to define some notation: let

$$\mathbf{T} = \begin{bmatrix} \frac{1}{2} & & & 0 \\ 1 & \ddots & & \\ \vdots & & \ddots & \\ 1 & \cdots & 1 & \frac{1}{2} \end{bmatrix}, \tag{4.14}$$

$\mathbf{D}_R$ be a diagonal matrix with diagonal elements R_k, and $\mathbf{D}_H$ be a diagonal matrix with diagonal elements H_k. With this notation the relationship between interface heights and layer thicknesses (4.13) can be written

$$\boldsymbol{z} = \mathbf{T}^T \boldsymbol{h}, \tag{4.15}$$

and the pressure perturbation equations (4.12) can be written

$$\boldsymbol{p} = g\mathbf{T}\left(\mathbf{D}_R \boldsymbol{h} + \mathbf{D}_H \boldsymbol{\rho}\right). \tag{4.16}$$

The velocity (4.9) and thickness perturbation (4.10) equations can be written in matrix form as,

$$\partial_t \mathbf{u} + f\hat{\boldsymbol{z}} \times \mathbf{u} = -\frac{g}{\rho_0}\nabla\left[\mathbf{T}\left(\mathbf{D}_R\boldsymbol{h} + \mathbf{D}_H\boldsymbol{\rho}\right) + \mathbf{D}_R\mathbf{T}^T\boldsymbol{h}\right], \tag{4.17}$$

$$\partial_t \boldsymbol{h} + \mathbf{D}_H \nabla \cdot \mathbf{u} = 0. \tag{4.18}$$

We may eliminate both the thickness and density perturbations by taking the time derivative of (4.17) and using the time-independence of the density perturbations (4.11) and the evolution equation (4.18) for the thickness perturbations. The result is,

$$\partial_t^2 \mathbf{u} + f\hat{\boldsymbol{z}} \times \partial_t \mathbf{u} = \frac{g}{\rho_0}\left[\left(\mathbf{T}\mathbf{D}_R + \mathbf{D}_R\mathbf{T}^T\right)\mathbf{D}_H\right]\nabla \cdot \mathbf{u}. \tag{4.19}$$

We can separate variables in this wave equation for the horizontal velocity provided that the matrix,

$$\frac{g}{\rho_0}\left[\left(\mathbf{T}\mathbf{D}_R + \mathbf{D}_R\mathbf{T}^T\right)\mathbf{D}_H\right], \tag{4.20}$$

has a complete set of eigenvectors; these eigenvectors would then correspond to the vertical structure of the linear normal modes of the system.

If the eigenvalues of this matrix are λ, then the dispersion relation for the associated waves is

$$\omega = 0, \quad \pm\sqrt{f^2 + \lambda k^2}, \tag{4.21}$$

where k is the magnitude of the horizontal Fourier wavenumber vector.

The relationship between the vertical structure of the normal modes of the primitive equations and the baroclinic modes of the discrete QG system is provided by the following theorem:

THEOREM 4.1. *The matrix* (4.20) *is equal to* $-f^2\mathbf{L}_{layer}^{-1}$*, where* $\mathbf{L}_{layer}$ *is the QG stretching matrix* (4.3).

The proof is deferred to Appendix A.

The implication of Theorem 4.1 is that the discrete baroclinic modes from QG theory are the same as the vertical structure of the normal modes (for the horizontal velocity) of the linearized discrete primitive equations (4.9)–(4.13). The interaction coefficients associated with the nonlinear advection terms in (4.4) and (4.5) are exactly the same as the interaction coefficients for the QG nonlinear term.

4.3. Discrete mode properties. We analyze the properties of the discrete eigensystem based on $\mathbf{L}$, defined in (4.1); however, the analysis applies to the layered shallow water version of QG as well as prevalent discretizations of the primitive equations as shown in subsection 4.1 and subsection 4.2.

A rich structure of properties, in many ways mirroring those of the eigenpairs of the continuous Sturm-Liouville problem (1.1), can be derived for the discrete system (4.2). The commonalities, however, do not ensure accuracy; the choice of grid may in fact preclude the discrete eigenvectors from accurately representing the behavior of the continuous eigenfunctions. To begin, we can confirm that important properties of the operator spectrum are preserved.

THEOREM 4.2. *The matrix* $\mathbf{L}$ *has real, distinct eigenvalues and is negative semi-definite.*

Proof. Via a similarity matrix $P = \text{diag}[h_1^{1/2}, h_2^{1/2}, \cdots, h_{\mathcal{N}}^{1/2}]$, $\mathbf{L}$ is similar to a real, symmetric, irreducible tridiagonal matrix. It therefore has real, distinct eigenvalues. $\mathbf{L}$ is also diagonally dominant, with $|\ell_{i,i}| \geq |\ell_{i,i-1}| + |\ell_{i,i+1}|$. Applying Gershgorin's circle theorem, the eigenvalues must then lie along the non-positive real axis. ∎

LEMMA 4.3. *The matrix* $\mathbf{L}$ *is self-adjoint with respect the the weighted inner product* $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}^T \mathbf{H} \boldsymbol{b}$.

Proof. Note that we may write $\mathbf{L} = \mathbf{H}^{-1}\mathbf{S}$ where $\mathbf{H} = \text{diag}(h_1, h_2, \ldots, h_{\mathcal{N}})$ and $\mathbf{S}$ is symmetric and tridiagonal. Then it follows easily that,

$$\langle \mathbf{L}\boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \mathbf{L}^T \mathbf{H} \boldsymbol{y} = \boldsymbol{x}^T \mathbf{S} \boldsymbol{y} = \boldsymbol{x}^T \mathbf{H} \mathbf{L} \boldsymbol{y} = \langle \boldsymbol{x}, \mathbf{L}\boldsymbol{y} \rangle,$$

and thus $\mathbf{L}$ is self-adjoint. ∎

THEOREM 4.4. *The eigenvectors of* $\mathbf{L}$ *form an orthonormal basis of* $\mathbb{R}^{\mathcal{N}}$ *with respect to the weighted inner product* $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \mathbf{H} \boldsymbol{y}$.

Proof. By Lemma 4.3, the matrix $\mathbf{L}$ is self-adjoint with respect to the weighted inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \mathbf{H} \boldsymbol{y}$. Eigenvectors of a self-adjoint matrix corresponding to distinct eigenvalues are orthogonal with respect to the appropriate inner product and by Theorem 4.2, all eigenvalues of $\mathbf{L}$ are distinct. Therefore, all eigenvectors of $\mathbf{L}$ are mutually orthogonal and thus form a basis for $\mathbb{R}^{\mathcal{N}}$. ∎

As in the continuous case, the discrete modes display a particular structure. Borrowing concepts and results from studies of oscillatory matrices, we are able to concretely characterize oscillations in the discrete eigenvectors. $\mathbf{L}$ is an irreducible tridiagonal matrix, sometimes termed a *Jacobi matrix* following [8, Chap. 2.1]. Further, requiring stable stratification, $N^2(z) > 0$, $-\mathbf{L}$ is a *normal Jacobi matrix* of the form,

$$\begin{bmatrix} a_1 & -b_1 & & & \\ -c_1 & a_2 & -b_2 & & \\ & -c_2 & \ddots & \ddots & \\ & & \ddots & \ddots & -b_{n-1} \\ & & & -c_{n-1} & a_n \end{bmatrix}$$

with $a_i, b_i, c_i \geq 0$. Such matrices are a common starting point in studies of more general oscillatory matrices. From Gantmacher and Krein [8, Thm. 1 and Thm. 4],

THEOREM 4.5. *Given* $\mathbf{L}$ *a normal Jacobi matrix, the sequence of entries in the* m^{th} *eigenvector,* $\boldsymbol{\phi}_m$*, has exactly* m *changes in sign* ($m = 0, \ldots, \mathcal{N} - 1$).

THEOREM 4.6. *The sign-changes (nodes) of two successive eigenvectors alternate.*

Therefore the discrete modes have both the increasing oscillations and vertical complexity of the continuous modes, as well as a version of the interlacing properties. These constraints on the structure of the discrete modes may then be examined in relation to the behavior of the continuous modes we hope to represent.

The barotropic mode is well-handled, regardless of grid choice. As in the continuous problem, for $\mathbf{L}$ defined by (4.1) we can confirm the guaranteed existence of a constant barotropic mode associated with $\lambda_0 = 0$. The discrete analog for the interaction coefficient, using the $\mathbf{H}$-weighted inner product for the orthogonality condition, is given by,

(4.22) $$\Theta_{\ell mn} = \boldsymbol{\phi}_n^T \mathbf{H} \left(\boldsymbol{\phi}_m \circ \boldsymbol{\phi}_\ell \right),$$

where $\circ$ denotes elementwise multiplication. The barotropic mode is constant so that (4.22) reduces to the $\mathbf{H}$-weighted orthonormal relation between the remaining two modes, just as it does in the continuous problem, yielding the correct interaction coefficients as well. (For $\mathbf{L}_{\text{layer}}$ defined by (4.3) the barotropic mode has no sign changes, but is weakly non-constant.)

The behavior of the oscillating baroclinic modes, however, is not guaranteed to coincide with that of the continuous system and may be undermined by a poor choice of grid. Consider, for example, the highest order discrete mode, $\boldsymbol{\phi}_{\mathcal{N}-1}$. Two thin layers might be placed together such that the continuous mode, $\phi_{\mathcal{N}-1}$, has no roots between the centers of these two thin layers. The discrete form requires that the elements $\boldsymbol{\phi}_{\mathcal{N}-1}$ corresponding to these layers differ in sign, creating an oscillation where there should not be one. Because $\boldsymbol{\phi}_{\mathcal{N}-1}$ has this spurious sign change, it must omit one of the continuous oscillations that occurs elsewhere in the domain in order to maintain the appropriate number of total sign changes. Thus the discrete baroclinic modes may be subject to both elided and spurious oscillations. The interaction coefficient will also be duly impacted. In order to avoid these errors, the layers should be ideally placed to support oscillations in the correct positions.

5. Optimal grid spacing and diagnostics.

We propose a grid that is equispaced in the Charney coordinate (1.4) in order to near-optimally resolve the baroclinic mode structures. As discussed above, approximations (subsection 3.1) and heuristics suggest that the baroclinic modes rapidly approach modulated cosines in ξ_c. Furthermore, the behavior of the discrete modes (subsection 4.3) implies that the layers should be in accordance with the oscillations of the baroclinic modes to allow the sign-changes to reflect continuous behavior. Equispaced layers in the Charney coordinate are well-suited to resolving the cosine-like behavior and are a sensible generalization for the continuously varying case.

Note that as the order of the highest-represented mode increases, the cosine approximation should improve. That is, the lower-order modes for which the spacing is less optimal have the benefit of having many points between the roots. The spacing becomes more optimal for the high-order modes that have the fewest points to capture the denser oscillations.

Previous multi-layer QG studies have proposed related schemes for choosing the grid based the idea of resolving baroclinic modes. For example, Beckmann [2] placed $\mathcal{N}$ layer interfaces at the roots of the $\mathcal{N}^{\text{th}}$ continuous mode. Practically speaking, to assign $\mathcal{N}-1$ layer depths with this method requires first computing an accurate approximation to the $\mathcal{N}^{\text{th}}$ continuous mode, which requires using a temporary discretization with significantly more than $\mathcal{N}$ layers.

Roullet et al. [15, eq. 10] develop an alternative approach to specifying the layer thicknesses by requiring their discretization of the continuous QG elliptic operator (2.9) be well-balanced. Their prescription can be related to the differential form of the Charney coordinate (1.5) as follows:

$$\Delta\xi = \frac{N}{f}\Delta z \quad \implies \quad \Delta\xi^2 = \frac{N^2}{f^2}\Delta z^2 = -\frac{g}{\rho_0 f^2}\Delta\rho\Delta z. \tag{5.1}$$

The final equality of the expression above uses the finite-difference approximation $N^2 = -g\partial_z\rho/\rho_0 \approx -g\Delta\rho/(\rho_0\Delta z)$, and the final grid spacing Δz is obtained by requiring $\Delta\xi$ to be constant. Practically speaking, to assign $\mathcal{N}$ layer depths with this method requires solving an equality-constrained nonlinear optimization problem.

Neither of these QG approaches has been proposed as a basis for a vertical co-

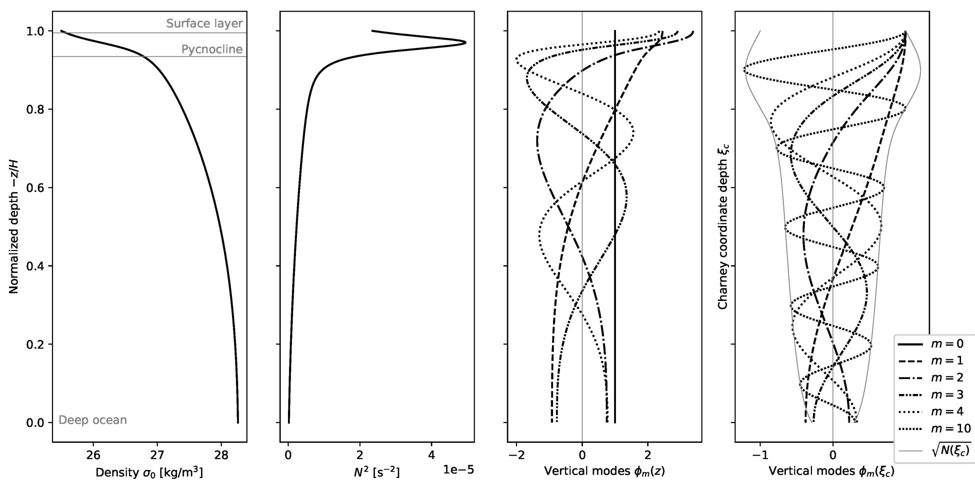

FIG. 1. *Density and buoyancy frequency profiles for a representative oceanic stratification profile with a strong subsurface pycnocline. The first few baroclinic modes, computed at high resolution, are shown in geopotential space and the stretched Charney coordinate along with the* $N^{1/2}(\xi_c)$ *envelope.*

ordinate in primitive-equation models, where the stratification varies slowly in the horizontal directions, because both would be impractical. In contrast, an equispaced grid in the Charney coordinate only requires integrating the profile, $N(z)$, which is typically already computed within the model for other purposes.

The importance of designing grids that accurately resolve the baroclinic modes is only now being discussed for global primitive equation modes. Stewart et al. [16] propose a geopotential grid in which they constrain the spacing using estimated roots of the first one or two baroclinic modes computed from global data with the WKB approximation. In a geopotential grid the layer depths are the same for all points on the globe, so to accommodate global variations in the stratification (and hence variations in the mode structure), the geopotential grid requires a significant number of layers – 50 to resolve the first baroclinic mode, and 25 more for each of the second and third modes.

5.1. Grid diagnostics with realistic ocean stratification. In order to test the behavior of different grids in an ocean-like setting, we define a realistic reference stratification profile,

$$\frac{N(z)}{f} = c_1 - c_2\frac{z}{H} + c_p\frac{w_p^2}{(\frac{z}{H} - z_p)^2 + w_p^2}. \tag{5.2}$$

The parameters $c_1 = 4$ and $c_2 = 22$ are chosen to ensure $N(z)$ positive and nonvanishing and to establish a gentle stratification at depth. Ocean profiles are typically dominated by the pycnocline, a region of strong stratification at or near the surface. The intensity, center, and depth of the pycnocline are controlled via the parameters $c_p = 45$, $z_p = -0.03$ and $w_p = 0.03$. The resulting density and buoyancy frequency profiles are shown in Figure 1. A high-resolution (512 equispaced layers in geopotential coordinates) reference computation of the baroclinic modes illustrates the expected leading barotropic mode and oscillating behavior of the baroclinic modes.

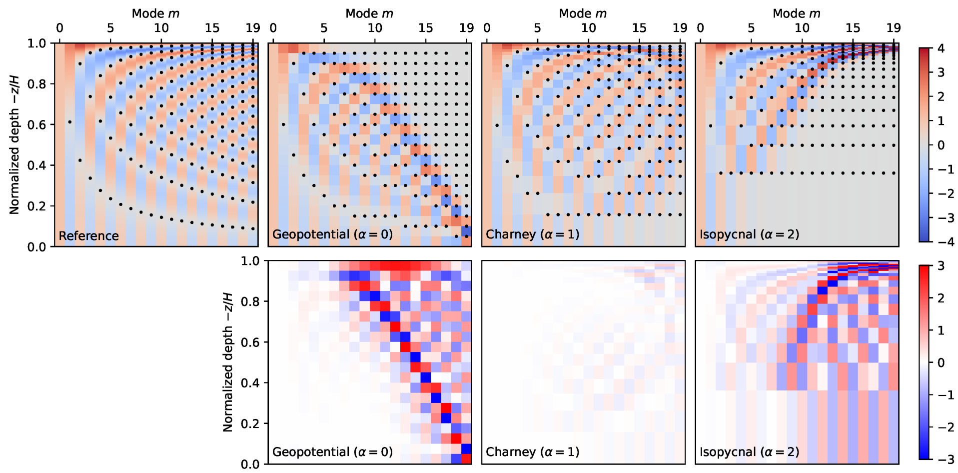

FIG. 2. *First twenty discrete baroclinic modes computed with high resolution and using twenty equispaced layers in geopopotential, Charney, and isopycnal coordinates. Black dots indicate sign changes in the eigenvector. Second row shows the corresponding error.*

We are primarily interested in the accuracy of the structure of the baroclinic modes, the corresponding baroclinic deformation radii, and the triad interactions represented by the interaction coefficient. Diagnostics on the performance of the grids focus on a limited resolution case. We compute the discrete baroclinic modes with twenty equispaced layers in geopotential, Charney, and isopycnal coordinates. As a reference, we use a high-resolution computation, which is well-converged for the first twenty modes and should be representative of the continuous behavior. This comparison primarily serves to illustrate the salutary properties of the Charney grid in determining a spacing; it is important to note that the geopotential and isopycnal grids used in ocean models are typically not equispaced. The Charney grid is seen to generalize behavior from the discrete representation of constant stratification case, in which the modes are all simple cosines, to the more complex, continuously-varying-stratification case (see supplement SM3).

We compare the structure of the low-resolution modes along with the pointwise error (Figure 2). As anticipated, the barotropic mode is captured across grids. The interlacing sign-changes of the modes is also apparent across grids. The first few baroclinic modes are fairly well resolved and display modest error. Considering the higher-order modes ($m \gtrsim 5$), a marked divergence emerges. Immediate, dramatic differences in the vertical structure are evident when using equispaced geopotential and isopycnal grids: the largest oscillations are trapped too low in the water column with geopotential grid and too high in the pycnocline with the isopycnal grid. The Charney coordinate grid allows for the oscillations to remain appropriately distributed and produces consistently low error throughout. The consistency confirms the theoretical prediction that a Charney-coordinate grid will require fewer layers to achieve resolution of the highest desired mode.

The normalized baroclinic deformation radii, H/λ_m, arising from the eigenvalues of the system and their corresponding errors are plotted in Figure 3. The barotropic mode is omitted; its zero eigenvalue and infinite radius are known to be reproduced for

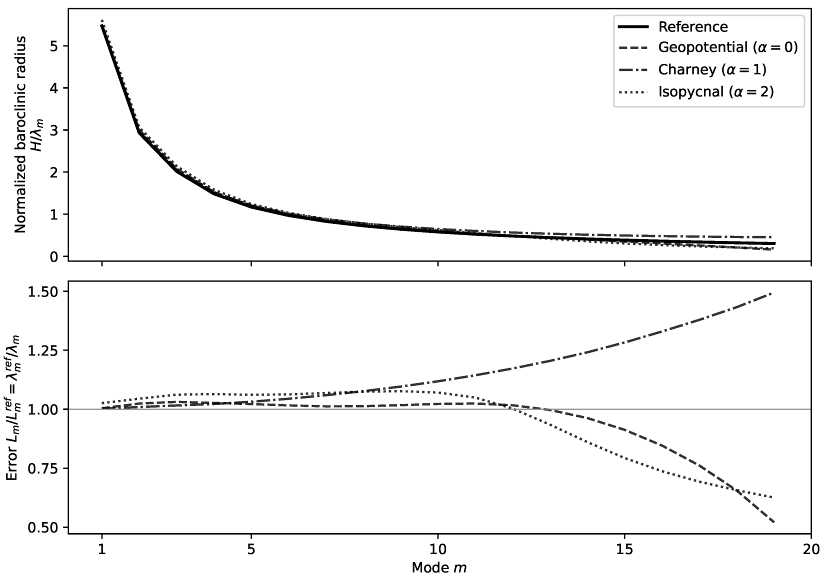

Fig. 3. *Normalized baroclinic radii computed using a high-resolution reference and with twenty equispaced layers in geopopotential, Charney, and isopycnal coordinates.*

any grid. The Charney grid outperforms the geopotential and isopycnal grids for the lowest modes and the error steadily increases with m, systematically overestimating the radius. The overestimation is consistent with behavior expected in a generalization of the discrete representation in a constant stratification case. The geopotential grid suffers in the lower modes but maintains lower error for longer before reaching similar magnitudes as the Charney grid at high m. The isopycnal grid performs worst of the three, overestimating the large, low-mode radii and underestimating the smaller, high-mode radii.

To assess the interaction coefficients (3.3), we use the discrete form (4.22) and examine slices with one index fixed (Figure 4). As noted, interactions with the barotropic mode reduce to orthogonality conditions reproduced across grids. For more complex triads, the characteristic diagonal structures present in the reference collapse under the geopotential and isopycnal grids. Interactions between the higher modes can fully misrepresent active and inactive triads, but even as low as the 4^{th} or 5^{th} mode, there are significant errors. The Charney grid, which successfully captures the structure of the baroclinic modes, also captures the interaction coefficients with much better accuracy throughout; the most significant errors are confined to the interactions with the very highest modes represented. With the Charney grid, the largest instances of error appear in the introduction of spurious negative interactions when $\ell + m + n = 2\mathcal{N}$, echoing aliasing effects that can be shown to be present in a constant stratification case.

The geopotential, isopycnal, and Charney grids may be understood as instances in a parameter space by leveraging the formulation of a coordinate family (1.3). We evaluate how the error varies in the transition between grids by testing equispaced grids that use coordinates with values $\alpha \in [0, 2]$ weighting the buoyancy frequency (Figure 5). The heuristic findings underscore the unique role of $\alpha = 1$ within the parameter space to minimize error in the representation of the baroclinic modes as

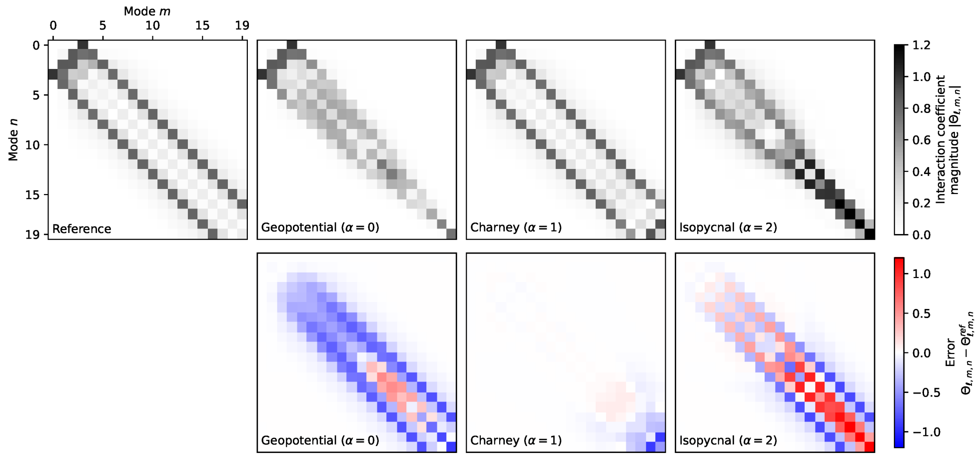

FIG. 4. *Slices of the interaction coefficient tensor $\Theta_{\ell mn}$ with $\ell = 3$ fixed. A high-resolution reference is shown, along with the values and errors computed using grids of twenty equispaced layers in geopopotential ($\alpha = 0$), Charney ($\alpha = 1$) and isopycnal ($\alpha = 2$) coordinates.*

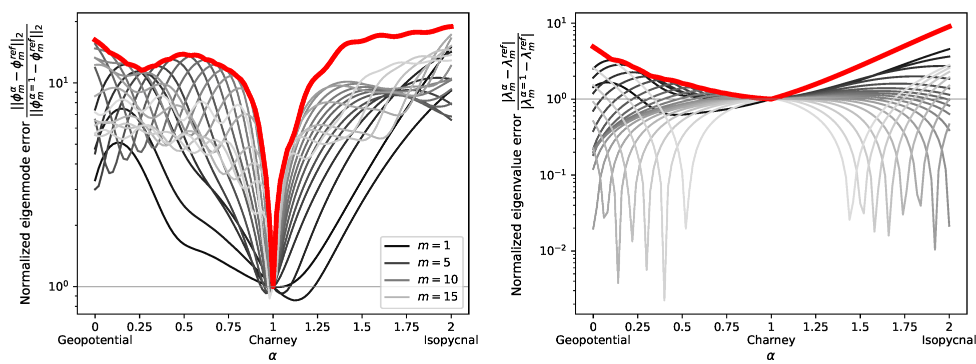

FIG. 5. *Normalized errors in the eigenmodes (left) and eigenvalues (right) computed using twenty-layer grids equispaced in varying α coordinates. Gray lines indicate individual modes with darker lines at lower modes. The thicker red line highlights the maximum over all m.*

well as the global error in the eigenvalues, indicating it is indeed an optimal choice.

6. Concluding remarks. In this paper we review the baroclinic modes governing natural vertical oscillations in the primitive and quasi-geostrophic systems along with their known properties. Given a continuously-varying stratification, the modal solutions cannot in general be analytically solved.

We present a derivation of the WKB approximation of the baroclinic modes for large eigenvalues and discuss the applicability for low modes. An alternate, unapproximated Liouville integral form of the modal solutions is also derived that coincides with the WKB approximation for large eigenvalues and suggests similar sinusoidal behavior may be present in the transformed Charney coordinate, independent of asymptotic scaling assumptions.

In numerical models, the underlying baroclinic modes are not typically explicitly represented but instead arise from the discrete stretching matrix for the vertical differential operator. Under common discretizations of the QG and primitive equation systems, we show the stretching matrices coincide. We further demonstrate properties of the discrete eigensystem, including the identification of oscillatory, sign-change requirements in the eigenvectors analogous to those of the continuous eigenfunctions.

Leveraging the analytical framework suggesting a special role of a Charney coordinate in regulating the phase of the modal oscillations, and constraints in the discrete behavior that demand layers coincide with the locations of the oscillations, we propose a new discrete grid approach with layers equispaced in the Charney coordinate. While the coordinate has appeared in analytic treatments [4, 5], its unique properties have not to our knowledge been exploited numerically.

The Charney grid is shown to near-optimally resolve the baroclinic modes and interaction coefficients, as well as the baroclinic radii with much improved accuracy compared to equispaced geopotential and isopycnal grids. With respect to existing approaches in QG literature, the Charney-coordinate approach to constructing a grid achieves favorable resolution properties, but is also efficient and computationally tractable enough for adaptation to primitive-equation models. Because the grid also responds to the local stratification, resolution can be achieved globally with fewer layers than required by a geopotential grid approach.

Building off of the theoretical basis for the Charney grid presented here, it remains to test its performance and impact on modeled dynamics in fully nonlinear QG simulations, and eventually to implement it as a vertical coordinate in a primitive-equation model.

Appendix A. Proof of Theorem 4.1. Let

$$\mathbf{M} = \frac{g}{\rho_0}\mathbf{P}\mathbf{D}_H \text{ where } \mathbf{P} = \mathbf{T}\mathbf{D}_R + \mathbf{D}_R\mathbf{T}^T. \tag{A.1}$$

The proof hinges on finding an explicit form for $\mathbf{P}^{-1}$, which is obtained by Gauss-Jordan elimination.

First note that

$$\mathbf{P} = \begin{bmatrix} R_1 & \cdots & \cdots & R_1 \\ \vdots & R_2 & \cdots & R_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_1 & R_2 & \cdots & R_N \end{bmatrix}. \tag{A.2}$$

Consider the application of elementary row operations to find the inverse by solving $\mathbf{PX} = \mathbf{I}$. The initial augmented matrix is

$$\left[\begin{array}{cccc|cccc} R_1 & \cdots & \cdots & R_1 & 1 & & & \\ \vdots & R_2 & \cdots & R_2 & & \ddots & & \\ \vdots & \vdots & \ddots & \vdots & & & \ddots & \\ R_1 & R_2 & \cdots & R_N & & & & 1 \end{array}\right]. \tag{A.3}$$

Eliminating below the first diagonal, followed by elimination to the right of the first diagonal, followed by a normalization of the first diagonal produces

$$
\text{(A.4)} \qquad \left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & \frac{1}{R_1}+\frac{1}{R_2^{(1)}} & -\frac{1}{R_2^{(1)}} & & \\
0 & R_2^{(1)} & \cdots & R_2^{(1)} & -1 & 1 & & \\
0 & \vdots & \ddots & \vdots & -1 & & \ddots & \\
0 & R_2^{(1)} & \cdots & R_N^{(1)} & -1 & & & 1
\end{array}\right]
$$

where

$$
\text{(A.5)} \qquad R_i^{(k)} = R_i - R_k > 0.
$$

Since the lower $(N-1)\times(N-1)$ submatrix on the left has the same form as the original matrix, we can almost proceed inductively, but not quite; the column of -1 on the right prevents this. However, we may proceed to eliminate below the diagonal in the second column, then eliminate to the right of the second diagonal, then normalize the second diagonal, which produces

$$
\text{(A.6)} \qquad \left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & \frac{1}{R_1}+\frac{1}{R_2^{(1)}} & -\frac{1}{R_2^{(1)}} & & \\
0 & 1 & 0 & 0 & -\frac{1}{R_2^{(1)}} & \frac{1}{R_2^{(1)}}+\frac{1}{R_3^{(2)}} & -\frac{1}{R_3^{(2)}} & \\
0 & 0 & R_3^{(2)} & \vdots & 0 & -1 & 1 & \\
0 & 0 & \cdots & R_N^{(2)} & 0 & -1 & & 1
\end{array}\right].
$$

At this point we can see the pattern repeating inductively and producing the final expression for $\mathbf{P}^{-1}$

$$
\text{(A.7)} \qquad \mathbf{P}^{-1} = \begin{bmatrix}
\frac{1}{R_1}+\frac{1}{R_2^{(1)}} & -\frac{1}{R_2^{(1)}} & & & \\
-\frac{1}{R_2^{(1)}} & \frac{1}{R_2^{(1)}}+\frac{1}{R_3^{(2)}} & -\frac{1}{R_3^{(2)}} & & \\
& \ddots & \ddots & \ddots & \\
& & -\frac{1}{R_{N-1}^{(N-2)}} & \frac{1}{R_{N-1}^{(N-2)}}+\frac{1}{R_N^{(N-1)}} & -\frac{1}{R_N^{(N-1)}} \\
& & & -\frac{1}{R_N^{(N-1)}} & \frac{1}{R_N^{(N-1)}}
\end{bmatrix}.
$$

To complete the proof we need to show that $-f^2\mathbf{M}^{-1} = \mathbf{L}$ where $\mathbf{L}$ is the discrete QG stretching matrix (4.3). Expanding,

$$
\text{(A.8)} \qquad -f^2\mathbf{M}^{-1} = -\frac{f^2\rho_0}{g}\mathbf{D}_H^{-1}\mathbf{P}^{-1} =
$$

$$
\begin{bmatrix}
-\frac{f^2}{gH_1}-\frac{f^2}{H_1 g_1'} & \frac{f^2}{H_1 g_1'} & & & \\
\frac{f^2}{H_2 g_1'} & -\frac{f^2}{H_2 g_1'}-\frac{f^2}{H_2 g_2'} & \frac{f^2}{H_2 g_2'} & & \\
& \ddots & \ddots & \ddots & \\
& & \ddots & \ddots & \frac{f^2}{H_{N-1} g_{N-1}'} \\
& & & \frac{f^2}{H_N g_{N-1}'} & -\frac{f^2}{H_{N-1} g_{N-1}'}
\end{bmatrix},
$$

this is clearly satisfied when the reduced gravities are $g_k' = gR_{k+1}^{(k)}/\rho_0 = g(R_{k+1} - R_k)/\rho_0$ and we assume $\rho_0 = R_1$. Further, under the identification

$$
\text{(A.9)} \qquad \frac{f^2}{g_{k-1}'} = f^2\frac{N_{k-1/2}^{-1}}{h_{k-1/2}}, \text{ i.e. } N^2(z_{k-1/2}) = \frac{g_{k-1}'}{\frac{1}{2}(h_{k-1}+h_k)}
$$

and in the limit $g \gg g'_1$ it is isomorphic to the discretization (4.1) of the continuous QG equations,

(A.10)
$$\mathbf{L} = f^2 \begin{bmatrix} -\frac{N_{3/2}^{-2}}{h_{3/2}h_1} & \frac{N_{3/2}^{-2}}{h_{3/2}h_1} & & & \\ \frac{N_{3/2}^{-2}}{h_{3/2}h_2} & \ddots & \ddots & & \\ & \ddots & \ddots & & \\ & \frac{N_{k-1/2}^{-2}}{h_{k-1/2}h_k} & -\left(\frac{N_{k-1/2}^{-2}}{h_{k-1/2}h_k} + \frac{N_{k+1/2}^{-2}}{h_{k+1/2}h_k}\right) & \frac{N_{k+1/2}^{-2}}{h_{k+1/2}h_k} & \\ & & \ddots & \ddots & \\ & & \ddots & \ddots & \frac{N_{\mathcal{N}-1/2}^{-2}}{h_{\mathcal{N}-1/2}h_{\mathcal{N}-1}} \\ & & & \frac{N_{\mathcal{N}-1/2}^{-2}}{h_{\mathcal{N}-1/2}h_{\mathcal{N}}} & -\frac{N_{\mathcal{N}-1/2}^{-2}}{h_{\mathcal{N}-1/2}h_{\mathcal{N}}} \end{bmatrix},$$

with $h_{i\pm\frac{1}{2}} = (h_i + h_{i\pm1})/2$.

The difference lies in the $(1,1)$ entry; the layer form assumes a free surface whereas the finite volume form imposes a stress-free boundary condition. The limit $g \gg g'_1$ is tantamount to saying that the variations in the height of the free upper surface are small, which is appropriate for ocean mesoscales.

REFERENCES

[1] A. Adcroft, W. Anderson, V. Balaji, C. Blanton, M. Bushuk, C. O. Dufour, J. P. Dunne, S. M. Griffies, R. Hallberg, M. J. Harrison, I. M. Held, M. F. Jansen, J. G. John, J. P. Krasting, A. R. Langenhorst, S. Legg, Z. Liang, C. McHugh, A. Radhakrishnan, B. G. Reichl, T. Rosati, B. L. Samuels, A. Shao, R. Stouffer, M. Winton, A. T. Wittenberg, B. Xiang, N. Zadeh, and R. Zhang, *The GFDL Global Ocean and Sea Ice Model OM4.0: Model Description and Simulation Features*, Journal of Advances in Modeling Earth Systems, 11 (2019), pp. 3167–3211, https://doi.org/10.1029/2019MS001726.

[2] A. Beckmann, *Vertical Structure of Midlatitude Mesoscale Instabilities*, Journal of Physical Oceanography, 18 (1988), pp. 1354–1371, https://doi.org/10.1175/1520-0485(1988)018⟨1354:VSOMMI⟩2.0.CO;2.

[3] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I*, Springer, New York, NY, 1999, https://doi.org/10.1007/978-1-4757-3069-2.

[4] J. G. Charney, *Geostrophic Turbulence*, Journal of the Atmospheric Sciences, 28 (1971), pp. 1087–1095, https://doi.org/10.1175/1520-0469(1971)028⟨1087:GT⟩2.0.CO;2.

[5] D. B. Chelton, R. A. deSzoeke, M. G. Schlax, K. E. Naggar, and N. Siwertz, *Geographical Variability of the First Baroclinic Rossby Radius of Deformation*, Journal of Physical Oceanography, 28 (1998), pp. 433–460, https://doi.org/10.1175/1520-0485(1998)028⟨0433:GVOTFB⟩2.0.CO;2.

[6] M. S. de La Lama, J. LaCasce, and H. K. Fuhr, *The vertical structure of ocean eddies*, Dynamics and Statistics of the Climate System, (2016), p. dzw001.

[7] R. Ferrari, S. M. Griffies, A. J. G. Nurser, and G. K. Vallis, *A boundary-value problem for the parameterized mesoscale eddy transport*, Ocean Modelling, 32 (2010), pp. 143–156, https://doi.org/10.1016/j.ocemod.2010.01.004.

[8] F. R. Gantmakher and M. G. Kreĭn, *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems*, AMS Chelsea Pub, Providence, R.I, rev. ed ed., 2002.

[9] I. Grooms and L.-P. Nadeau, *The effects of mesoscale ocean–atmosphere coupling on the quasigeostrophic double gyre*, Fluids, 1 (2016), p. 34.

[10] R. Hallberg, *Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects*, Ocean Modelling, 72 (2013), pp. 92–103, https://doi.org/10.1016/j.ocemod.2013.08.007.

[11] J. Liouville, *Second Mémoire sur le développement des fonctions ou parties de fonctions en séries dont les divers termes sont assujétis à satisfaire à une même équation différentielle du second ordre, contenant un paramètre variable.*, Journal de Mathématiques Pures et Appliquées, (1837), pp. 16–35.

[12] J. Pedlosky, *Quasigeostrophic Motion of a Stratified Fluid on a Sphere*, in Geophysical Fluid Dynamics, J. Pedlosky, ed., Springer US, New York, NY, 1979, pp. 314–422, https://doi.org/10.1007/978-1-4684-0071-7_6.

[13] T. Ringler, M. Petersen, R. L. Higdon, D. Jacobsen, P. W. Jones, and M. Maltrud, *A multi-resolution approach to global ocean modeling*, Ocean Modelling, 69 (2013), pp. 211–232, https://doi.org/10.1016/j.ocemod.2013.04.010.

[14] C. B. Rocha, W. R. Young, and I. Grooms, *On Galerkin Approximations of the Surface Active Quasigeostrophic Equations*, Journal of Physical Oceanography, 46 (2016), pp. 125–139, https://doi.org/10.1175/JPO-D-15-0073.1.

[15] G. Roullet, J. C. McWilliams, X. Capet, and M. J. Molemaker, *Properties of Steady Geostrophic Turbulence with Isopycnal Outcropping*, Journal of Physical Oceanography, 42 (2012), pp. 18–38, https://doi.org/10.1175/JPO-D-11-09.1.

[16] K. Stewart, A.McC. Hogg, S. Griffies, A. Heerdegen, M. Ward, P. Spence, and M. England, *Vertical resolution of baroclinic modes in global ocean models*, Ocean Modelling, 113 (2017), pp. 50–65, https://doi.org/10.1016/j.ocemod.2017.03.012.

[17] G. Teschl, *Ordinary differential equations and dynamical systems*, vol. 140, American Mathematical Soc., 2012.

[18] E. C. Titchmarsh, *Elgenfunction Expansions Associated With Second Order Differential Equations*, Oxford At The Clarendon Press., 1946.

[19] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*, Cambridge University Press, Cambridge, 2 ed., 2017, https://doi.org/10.1017/9781107588417.