Energy-Conserving Galerkin Approximations for Quasigeostrophic Dynamics

by

Matthew Watwood

B.S., University of Denver, 2015

M.S., University of Denver, 2015

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Department of Applied Mathematics

2020

Committee Members:

Ian Grooms, Chair

Keith Julien

Aneesh Subramanian

Watwood, Matthew (M.S., Applied Mathematics)

Energy-Conserving Galerkin Approximations for Quasigeostrophic Dynamics

Thesis directed by Prof. Ian Grooms

A method is presented for constructing energy-conserving Galerkin approximations in the vertical coordinate of the full quasigeostrophic model with active surface buoyancy. The derivation generalizes the approach of Rocha **et al.** (2016) [30] to allow for general bases. Details are then presented for a specific set of bases: Legendre polynomials for potential vorticity and a recombined Legendre basis from Shen (1994) [32] for the streamfunction. The method is tested in the context of linear baroclinic instability calculations, where it is compared to the standard second-order finite-difference method and to a Chebyshev collocation method. The Galerkin scheme is quite accurate even for a small number of degrees of freedom , and growth rates converge much more quickly with increasing for the Galerkin scheme than for the finite-difference scheme. The Galerkin scheme is at least as accurate as finite differences and can in some cases achieve the same accuracy as the finite difference scheme with ten times fewer degrees of freedom. The energy-conserving Galerkin scheme is of comparable accuracy to the Chebyshev collocation scheme in most linear stability calculations, but not in the Eady problem where the Chebyshev scheme is significantly more accurate. Finally the three methods are compared in the context of a simplified version of the nonlinear equations: the two-surface model with zero potential vorticity. The Chebyshev scheme is the most accurate, followed by the Galerkin scheme and then the finite difference scheme. All three methods conserve energy with similar accuracy, despite not having any a priori guarantee of energy conservation for the Chebyshev scheme. Further nonlinear tests with non-zero potential vorticity to assess the merits of the methods will be performed in a future work.

Acknowledgements

Contents

Figures

Figure

Chapter 1

Introduction

The well-known quasigeostrophic (QG) model describes the dynamics of extratropical oceanic and atmospheric circulations [26, 29, 42], characterized by the dominant roles of background stratification, leading to hydrostatic balance, and planetary rotation, resulting in geostrophic balance. In the case of a fluid bounded above and below by flat rigid surfaces, where the vertical velocity must vanish, the dynamics are governed by the evolution of three quantities: the quasigeostrophic potential vorticity (PV) $q(x, y, z, t)$ and the buoyancy at the top and bottom surfaces, $b^+(x, y, t)$ and $b^-(x, y, t)$ respectively. (The superscripts $^+$ and $^-$ henceforth denote evaluation of a quantity at the top and bottom surfaces of the domain, respectively.) The three quantities q and $b^\pm$ evolve according to

$$\partial_t b^+ + \boldsymbol{u}^+ \cdot \nabla b^+ = 0 \tag{1.1a}$$

$$\partial_t q + \boldsymbol{u} \cdot \nabla q + \beta v = 0 \tag{1.1b}$$

$$\partial_t b^- + \boldsymbol{u}^- \cdot \nabla b^- = 0. \tag{1.1c}$$

The equations are set in a linear tangent plane to the sphere at latitude θ. The parameter β is $(f/R)\cos(\theta)$ where R is the radius of the Earth and f is twice the rotation rate of the Earth. Hereafter the symbol ∇ is to be understood as horizontal only. The velocity is also horizontal,

$\boldsymbol{u} \cdot \nabla = u\partial_x + v\partial_y$, and the velocity is incompressible $\nabla \cdot \boldsymbol{u} = 0$. The velocity is obtained from the PV and surface buoyancies by solving the following elliptic equation for the streamfunction ψ

$$f_0\partial_z\psi = b^+ \text{ at } z = H \tag{1.2a}$$

$$\nabla^2\psi + \partial_z\left(S(z)\partial_z\psi\right) = q \tag{1.2b}$$

$$f_0\partial_z\psi = b^- \text{ at } z = 0 \tag{1.2c}$$

and then setting $u = -\partial_y\psi$, $v = \partial_x\psi$. The streamfunction ψ can either be set to 0 on the side boundaries, or they can be periodic. For periodic boundaries the spatially-constant part of the streamfunction is not determined by (1.2), but this component of ψ has no impact on the dynamics and can therefore be set to any desired value. The function $S(z)$ is $f_0^2/N^2(z)$ where f_0 is the local Coriolis parameter $f_0 = f\sin(\theta)$ and $N(z) > 0$ is the Brunt-Väisälä frequency, also known as the buoyancy frequency.

There are two common simplifications of this system that set either q and β, or $b^\pm$ to zero. Both of these simplifications are exact solutions of the full equations. Setting $b^\pm = 0$ leads to considerable simplifications in the analysis of the system and in the development of numerical methods. For example, solutions of the system with $b^\pm = 0$ are known to be regular and to be a regular asymptotic limit of the Navier-Stokes equations [3], and standard finite-difference approximations are known to be convergent [7].

The other simplification sets $q = \beta = 0$. The surface-QG model (sQG) is obtained by further setting $b^+ = 0$ and replacing the boundary condition $\partial_z\psi = 0$ at $z = H$ with the condition $\partial_z\psi \to 0$ as $z \to \infty$. Rigorous mathematical analysis of the inviscid sQG model is considerably more difficult than the model with $b^\pm = 0$. A connection of the sQG model to the three-dimensional Euler equations was made in [9], and the study of well-posedness of the sQG model is ongoing (e.g. [8]; for related results in a model with dissipation see [25, 28]). As a result of this connection between

the full unsimplified QG model and the sQG model global regularity of the full QG system remains an open problem, though strong solutions are known to exist and be unique for a finite time horizon [2].

Surface buoyancy has a significant impact on the dynamics of the upper ocean [24] and on the atmosphere near the tropopause [41], and the simplified system with $b^\pm = 0$ is unable to model these dynamics. Surface-QG is able to model the impact of surface buoyancy on atmospheric and oceanic dynamics, but it is only an exact solution of the full system (1.1) when $\beta = 0$. In addition, several studies have used sQG theory to infer ocean subsurface velocities using only surface buoyancy but these studies have found that the assumption $q = 0$ prevents accurate reconstruction except near the ocean surface [23, 19, 22]. To model the interplay of surface and interior dynamics one needs the full unsimplified QG model.

The primary difficulty in constructing discretizations of the full QG model is the discretization of the vertical coordinate z. The equations do not have a particularly unusual form and any one of a variety of classical methods could be used, but particular attention is paid in the community to whether a discretization is energy-conserving. This is because simulations are often used to study the energetics of the system, e.g. the transfer of energy between horizontal and vertical scales [34, 35, 31], using integrations over long time scales. The classical second order finite difference discretization found in [29, 7, 42] conserves energy and is the standard method for simulations of the full system [34, 35, 17, 21, 27, 43, 31, 33, 16]. Higher-order alternatives are therefore desirable for the purpose of achieving equal accuracy with less cost, or higher accuracy at equal cost.

For a fully discrete system to be energy-conserving naturally depends also on the discretization of the horizontal directions. Energy-conserving horizontal discretizations are available [1]; our focus is on discretization of the vertical coordinate.

The main alternative to the standard second order finite difference discretization of the vertical coordinate is a Galerkin approach based on [10]. The operator $\partial_z(S(z)\partial_z\cdot)$ with homogeneous Neumann boundary conditions admits a set of eigenfunctions called 'baroclinic modes' that form an

orthogonal basis for $L^2(0,1)$. A finite number of these modes can be used as a basis for an energy-conserving Galerkin approximation of the QG equations, but the straightforward application clearly assumes that $b^\pm = 0$ since that is the boundary condition satisfied by the basis functions. It was recently shown in [30] how the same basis could be used in a way that does not require $b^\pm = 0$, and that still conserves energy. This approach is counter-intuitive in that it generates an approximate solution with $\partial_z \psi = 0$ at the surfaces, but with $b^\pm \neq 0$; nevertheless the approximation to ψ still converges absolutely and uniformly as the number of basis functions increases. Unfortunately these basis functions are not practically useful except in the case where $S(z)$ is a constant, in which case the modes are just Fourier modes. (Precisely this Fourier baroclinic mode basis was recently used in [37], but for the simplified QG model with $b^\pm = 0$.)

It was proposed in [41] to simply augment the baroclinic mode basis with auxiliary functions that enable satisfaction of the inhomogeneous Neumann conditions, but this method does not conserve energy [41, 30]. An alternative orthogonal basis was developed in [36]. This basis enables both satisfaction of the inhomogeneous Neumann boundary conditions and conservation of energy in a Galerkin approximation, but the basis does not enable separation of variables in the solution of the elliptic equation and is more useful for analysis of observational data than for high-resolution simulations of the nonlinear dynamics.

This article presents an energy-conserving Galerkin approximation scheme for the vertical coordinate of the full QG system that generalizes the approach in [30] so that it can be used with any appropriate basis while allowing active surface buoyancy. We immediately specialize to a global polynomial basis based on Legendre polynomials. Legendre polynomials are a convenient basis because energy is defined using an un-weighted L^2 norm squared, and the Legendre polynomials are orthogonal with respect to the un-weighted L^2 inner product. The paper is organized as follows. Our main result on the construction of an energy-conserving Galerkin approximation is found in chapter 2. Implementation details and a specific choice of polynomial basis are presented in chapter 3. The method is tested and compared to the standard finite diffence method and to a

non-energy-conserving Chebyshev collocation method in the context of linear baroclinic instability calculations in chapter 4. In chapter 5 the new Galerkin method is compared to the standard finite difference method and to the Chebyshev collocation scheme in fully-nonlinear simulations of idealized two-surface dynamics with $\beta = q = 0$. Results are discussed and conclusions are offered in chapter 6.

Chapter 2

Energy-Conserving Galerkin Approximations

The QG equations conserve energy in the form

$$E = \frac{1}{2} \int_\Omega |\nabla \psi|^2 + S(z)(\partial_z \psi)^2$$

where $\int_\Omega$ represents an integral over the whole domain $(x, y, z) \in \Omega$. The first term corresponds to kinetic energy and the second to available potential energy. The proof of energy conservation is straightforward: Equation (1.1b) is multiplied by $-\psi$, followed by integration over the volume. Careful use of integration by parts together with eq. (1.2) and the surface buoyancy equations (1.1a) and (1.1c) yield the desired result. The boundary conditions at the side boundaries can be assumed to be either impenetrable or periodic.

Suppose that q will be represented as a linear combination of basis functions $p_n^q(z)$

$$q_\mathcal{N} = \sum_{n=1}^{\mathcal{N}} \breve{q}_n(x, y, t) p_n^q(z). \tag{2.1}$$

The notation $\mathcal{N}$ serves to distinguish the number of basis functions $\mathcal{N}$ from the Brunt-Väisälä frequency N. The notation for the coefficients $\breve{q}_n$ follows [30]. Similarly, suppose that ψ will be represented as a linear combination of basis functions $p_n^\psi(z)$, which can be different from the basis used for q:

$$\psi_\mathcal{N} = \sum_{n=1}^{\mathcal{N}} \breve{\psi}_n(x, y, t) p_n^\psi(z). \tag{2.2}$$

The approximation to ψ is used to evolve surface buoyancy as follows

$$\partial_t b^\pm + \boldsymbol{u}_\mathcal{N}^\pm \cdot \nabla b^\pm = 0.$$

There are now two approximation problems. The first is related to PV inversion: given $q_\mathcal{N}$ and $b^\pm$, how does one obtain coefficients $\breve{\psi}_n$ for $\psi_\mathcal{N}$? The second is related to PV evolution: given $q_\mathcal{N}$ and $\psi_\mathcal{N}$, how does one obtain the tendencies $\partial_t \breve{q}_n$? The way that these questions are answered determines the kind of approximation being made, as well as whether the scheme conserves energy.

However these questions are answered one can always define residuals r_q and r_t related to the two approximations, along with residuals $r_b^\pm$ related to the boundary conditions

$$r_q = q_\mathcal{N} - \nabla^2 \psi_\mathcal{N} - \partial_z \left(S(z) \partial_z \psi_\mathcal{N} \right) \tag{2.3a}$$

$$r_t = \partial_t q_\mathcal{N} + \boldsymbol{u}_\mathcal{N} \cdot \nabla q_\mathcal{N} + \beta v_\mathcal{N} \tag{2.3b}$$

$$r_b^\pm = b^\pm - f_0 \partial_z \psi_\mathcal{N}^\pm. \tag{2.3c}$$

It is important to note that in the above equations $q_\mathcal{N}$ and $\psi_\mathcal{N}$ are not Galerkin coefficients (which are denoted $\breve{q}_n$ and $\breve{\psi}_n$); they are instead the full Galerkin approximations to q and ψ given by (2.1) and (2.2). If the boundary conditions on the PV inversion are exactly satisfied then $r_b^\pm = 0$. The approximate potential vorticity $q_\mathcal{N}$ is not materially conserved unless $r_t = 0$ and $\beta = 0$; this is true regardless of the basis functions chosen, including finite elements or the baroclinic modes of [30]. A Galerkin approximation to the PV inversion would choose the coefficients $\breve{\psi}_n$ according to the condition that the residual r_q be orthogonal to the span of the basis functions p_n^ψ.

An energy-conserving discretization should conserve energy in the following form

$$E_\mathcal{N} = \frac{1}{2} \int_\Omega |\nabla \psi_\mathcal{N}|^2 + S(z)(\partial_z \psi_\mathcal{N})^2.$$

One can obtain an exact evolution equation for $E_\mathcal{N}$ by, for example, multiplying eq. (2.3b) by $-\psi_\mathcal{N}$ and integrating over the volume. One eliminates $q_\mathcal{N}$ using eq. (2.3a), and then performs integrations by parts using boundary conditions eq. (2.3c). The result is

$$\frac{\mathrm{d}E_\mathcal{N}}{\mathrm{d}t} = -\int_x \left[S(H)\psi_\mathcal{N}^+ \partial_t r_b^+ - S(0)\psi_\mathcal{N}^- \partial_t r_b^- \right] + \int_\Omega \psi_\mathcal{N}(\partial_t r_q - r_t). \tag{2.4}$$

The first integral on the right hand side of this equation is taken over the horizontal upper and lower surfaces; this has been indicated by the subscript x on the integral: $\int_x$.

Energy conservation can evidently be achieved quite simply as follows. First apply the usual Galerkin condition to the PV inversion by requiring r_q to be L^2-orthogonal to the basis functions $p_n^\psi(z)$; this eliminates the term $\int_\Omega \psi_\mathcal{N} \partial_t r_q$ in eq. (2.4). Next, satisfy the boundary conditions exactly so that $r_b^\pm = 0$, eliminating the first term on the right hand side of eq. (2.4). Finally, apply a Petrov-Galerkin condition to determine the evolution $\partial_t q_\mathcal{N}$ by requiring r_t to be L^2-orthogonal to the span of p_n^ψ, eliminating the last term on the right hand side of eq. (2.4). (This latter is a Petrov-Galerkin condition because the residual is made orthogonal to a different subspace than the one in which the approximation is sought. If $p_n^q = p_n^\psi$ then this is just another Galerkin condition.)

The problem is that there are only $2\mathcal{N}$ degrees of freedom — one each for $\breve{\psi}_n$ and $\breve{q}_n$ — while the above recipe yields $2\mathcal{N} + 2$ conditions. This difficulty was avoided in [30] by making use of a clever reformulation of the PV inversion proposed by Bretherton [4]. The solution to the PV inversion eq. (1.2) is the same as the solution to the following reformulated problem [4]

$$\partial_z \psi = 0 \text{ at } z = H \tag{2.5a}$$

$$\nabla^2 \psi + \partial_z \left(S(z)\partial_z \psi \right) = q - \frac{f_0}{N^2(z)} b^+ \delta(z - H) + \frac{f_0}{N^2(z)} b^- \delta(z) \tag{2.5b}$$

$$\partial_z \psi = 0 \text{ at } z = 0. \tag{2.5c}$$

The equivalence of these two formulations can be obtained through the Green's function formulation

of the solution, as shown in C. There is a new residual associated with this reformulation of the inversion, defined to be

$$r_q^B = q_\mathcal{N} - \frac{f_0}{N^2(z)}b^+\delta(z - H) + \frac{f_0}{N^2(z)}b^-\delta(z) - \nabla^2\psi_\mathcal{N} - \partial_z\left(S(z)\partial_z\psi_\mathcal{N}\right). \qquad (2.6)$$

The superscript B in r_q^B stands for 'Bretherton.' Using this reformulated problem to derive an evolution equation for the discretized energy yields a deceptively similar equation:

$$\frac{\mathrm{d}E_\mathcal{N}}{\mathrm{d}t} = -\int_x \left[S(H)\psi_\mathcal{N}^+\partial_t r_b^+ - S(0)\psi_\mathcal{N}^-\partial_t r_b^-\right] + \int_\Omega \psi_\mathcal{N}(\partial_t r_q^B - r_t). \qquad (2.7)$$

Further simplifications are possible though, since the basis functions can now be assumed to satisfy homogeneous Neumann boundary conditions $\partial_z p_n^\psi = 0$, consistent with the boundary conditions of the reformulated PV inversion. These boundary conditions imply that

$$r_b^\pm = b^\pm \quad \Rightarrow \quad \partial_t r_b^\pm = -\boldsymbol{u}_\mathcal{N}^\pm \cdot \nabla b^\pm. \qquad (2.8)$$

This implies that the first integral on the right hand side of the energy budget eq. (2.7) is zero since $\int_x \psi_\mathcal{N}\boldsymbol{u}_\mathcal{N} \cdot \nabla b = 0$. As a consequence the energy budget takes the form

$$\frac{\mathrm{d}E_\mathcal{N}}{\mathrm{d}t} = \int_\Omega \psi_\mathcal{N}(\partial_t r_q^B - r_t). \qquad (2.9)$$

Energy conservation can now be achieved through the use of a Galerkin condition on the reformulated PV inversion (r_q^B L^2-orthogonal to p_n^ψ), and a Petrov-Galerkin condition on the evolution tendency (r_t L^2-orthogonal to p_n^ψ). This is essentially a re-derivation of the result in [30] using an arbitrary basis instead of the baroclinic mode basis. This derivation enables the use of practical algorithms based on finite element bases, spline bases, or polynomial bases in cases where the baroclinic mode basis of [30] is unavailable or unwieldy.

Energy conservation in eq. (2.9) requires orthogonality in the standard, un-weighted L^2 inner product. It is convenient to use Legendre polynomials because they are orthogonal with respect to this inner product, and a method based on Legendre polynomials is developed in chapter 3.

Chebyshev polynomials are not orthogonal with respect to the standard L^2 inner product; they are instead orthogonal with respect to a weighted L^2 inner product with weight function $w(z) = 1/\sqrt{z(H-z)}$ for $z \in [0, H]$. They are therefore less convenient than Legendre polynomials in this setting.

We note as a brief aside that the Galerkin condition on the PV inversion is equivalent to choosing $\psi_{\mathcal{N}}$ to minimize a semi-norm of the error. To wit, let ψ_* be the true solution to the reformulated PV inversion with $q = q_{\mathcal{N}}$ and note that in general ψ_* cannot be exactly described by an approximation of the form eq. (2.2); then the $\psi_{\mathcal{N}}$ that minimizes the following semi-norm of the error

$$\|\psi_* - \psi_{\mathcal{N}}\|_q^2 = \int |\nabla(\psi_* - \psi_{\mathcal{N}})|^2 + S(z)\left(\partial_z(\psi_* - \psi_{\mathcal{N}})\right)^2 \tag{2.10}$$

is the same as the $\psi_{\mathcal{N}}$ that sets the residual r_q^B L^2-orthogonal to the span of the basis functions p_n^ψ. This was not precisely clear in [30] where the discussion could be misconstrued to suggest that the Galerkin condition is equivalent to minimizing the L^2 norm of the error.

Energy is not the only sign-definite quadratic quantity conserved by the full quasigeostrophic system (1.1); it also conserves enstrophy when $\beta = 0$. Enstrophy is half the volume integral of the square of the potential vorticity; in the approximation (2.1) it is

$$Z_{\mathcal{N}} = \frac{1}{2} \int_\Omega q_{\mathcal{N}}^2.$$

Its evolution in the approximate system is derived by multiplying eq. (2.3b) by $q_{\mathcal{N}}$ and integrating over the volume. The result, assuming $\beta = 0$, is

$$\frac{dZ_{\mathcal{N}}}{dt} = \int_\Omega q_{\mathcal{N}} r_t. \tag{2.11}$$

The energy-conserving method proposed above imposes the condition that r_t be L^2-orthogonal to the span of the basis functions p_n^ψ. The above expression indicates that approximate enstrophy

will not be conserved in the energy-conserving scheme unless the same basis functions are used for both ψ and q, i.e. $p_n^\psi = p_n^q$. The energy-conserving method requires $\partial_z p_n^\psi = 0$ at the boundaries, but there is no a priori reason to impose the boundary condition $\partial_z q = 0$ by using $p_n^\psi = p_n^q$. One is therefore faced with the choice of using $p_n^\psi = p_n^q$, which leads to an enstrophy-conserving method but potentially degrades the approximation of q, or using $p_n^\psi \neq p_n^q$, which will not conserve enstrophy but may allow a more accurate approximation of q. In the following section we continue the analysis under the more general assumption that p_n^ψ is not necessarily equal to p_n^q, leaving open the option to choose $p_n^\psi = p_n^q$ if desired.

Chapter 3

A Legendre basis

This section begins with general considerations associated with implementation using an arbitrary basis, and continues to consideration of a specific basis using Legendre polynomials. There are two problems to deal with: (i) computing $\psi_{\mathcal{N}}$ from $q_{\mathcal{N}}$ and $b^{\pm}$, and (ii) computing $\partial_t q_{\mathcal{N}}$.

First consider the implementation of the Galerkin condition on the reformulated PV inversion eq. (2.5). Assume that the basis $p_n^{\psi}(z)$ satisfies homogeneous Neumann conditions. The condition that the residual in eq. (2.6) be L^2-orthogonal to the basis functions p_n^{ψ} leads to an $\mathcal{N} \times \mathcal{N}$ linear system of the following form

$$\nabla^2 \mathbf{M} \boldsymbol{\psi} - \mathbf{L} \boldsymbol{\psi} = \mathbf{B} \mathbf{q} - \frac{f_0}{N^2(H)} b^+ \mathbf{p}^+ + \frac{f_0}{N^2(0)} b^- \mathbf{p}^- \tag{3.1}$$

where the vectors $\boldsymbol{\psi}$ and $\mathbf{q}$ contain the Galerkin coefficients $\breve{\psi}_n$ and $\breve{q}_n$, the vectors $\mathbf{p}^+$ and $\mathbf{p}^-$ have elements

$$\mathbf{p}_n^+ = p_n^{\psi}(H) \qquad \mathbf{p}_n^- = p_n^{\psi}(0), \tag{3.2}$$

and the matrices have the following elements

$$\mathbf{M}_{i,j} = \int_0^H p_i^{\psi}(z) p_j^{\psi}(z) \mathrm{d}z, \tag{3.3a}$$

$$\mathbf{L}_{i,j} = \int_0^H S(z)(\partial_z p_i^{\psi}(z))(\partial_z p_j^{\psi}(z)) \mathrm{d}z, \tag{3.3b}$$

$$\mathbf{B}_{i,j} = \int_0^H p_i^\psi(z) p_j^q(z) \mathrm{d}z. \tag{3.3c}$$

The discrete PV inversion system eq. (3.1) can be diagonalized by a change of basis. The mass matrix $\mathbf{M}$ is symmetric positive definite and has a Cholesky factorization

$$\mathbf{M} = \mathbf{G}\mathbf{G}^T.$$

The matrix $\mathbf{G}^{-1}\mathbf{L}\mathbf{G}^{-T}$ is symmetric positive semi-definite, and has an orthogonal eigenvalue decomposition

$$\mathbf{G}^{-1}\mathbf{L}\mathbf{G}^{-T} = \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}^T$$

where $\mathbf{Q}$ is real orthogonal and $\boldsymbol{\Sigma}$ is diagonal. With these matrix decompositions the discrete PV inversion can be diagonalized as follows

$$\nabla^2\mathbf{M}\psi - \mathbf{L}\psi = \mathbf{G}\mathbf{Q}(\nabla^2\mathbf{I} - \boldsymbol{\Sigma})\mathbf{Q}^T\mathbf{G}^T\psi. \tag{3.4}$$

The PV inversion has to be solved repeatedly during a time integration of the PV equations. An efficient implementation first computes the matrices $\mathbf{G}$, $\mathbf{Q}$, and $\boldsymbol{\Sigma}$, then uses the diagonalization above to split the PV inversion into a set of independent two-dimensional elliptic inversions. Once the matrix factorizations have been computed, the cost to invert the PV is $\mathcal{O}(\mathcal{N}^2)$ (plus $\mathcal{N}$ times the cost of each two-dimensional elliptic inversion). This approach, of decomposing the three-dimensional PV inversion into a series of independent two-dimensional PV inversions, is employed by several widely-used codes including [34, 17, 21, 13, 16]. The grid sizes in common current use are small enough that several two-dimensional PV inversions fit within local memory in a single node of a distributed-memory machine, which means that cross-node communication is limited to the change of vertical coordinate. Furthermore, there are often between several hundred and one thousand degrees of freedom in each horizontal direction, while $\mathcal{N}$ is usually less than 100 and only occasionally in the low hundreds for codes that use the standard second-order finite

difference approximation. The Galerkin method developed here should require smaller $\mathcal{N}$ than the finite difference approximation, so in practice $\mathcal{N}$ should be expected to be relatively small and the $\mathcal{O}(\mathcal{N}^2)$ cost to convert the three-dimensional inversion to a set of two-dimensional inversions should be significantly smaller than the cost of a single two-dimensional inversion. Nevertheless, it may be true in some cases that a full multigrid approach applied directly to eq. (3.1) will be more efficient than the diagonalization described above (cf. [31]).

Each column of the matrix $\mathbf{G}^{-T}\mathbf{Q}$ contains coefficients of a function in the basis p_n^ψ. These functions together form a basis that diagonalizes the discrete PV inversion operator. These basis functions that diagonalize the discrete PV operator are approximations to the baroclinic modes of [30] that diagonalize the continuous PV operator.

Next consider the problem of computing $\partial_t q_\mathcal{N}$ via the Petrov-Galerkin condition that r_t in eq. (2.3b) be L^2-orthogonal to the p_n^ψ basis. This condition yields a system of the form

$$\mathbf{B}\left(\partial_t \mathbf{q}\right) = \dot{\mathbf{q}} \tag{3.5}$$

where the elements of the right hand side are

$$(\dot{\mathbf{q}})_n = -\int_0^H p_n^\psi(\boldsymbol{u}_\mathcal{N} \cdot \nabla q_\mathcal{N} + \beta v_\mathcal{N})\mathrm{d}z. \tag{3.6}$$

Fortunately, for basis functions consisting of polynomials or piecewise polynomials these integrals can be computed exactly, either analytically or via appropriate quadratures. Since this system needs to be solved repeatedly during a time integration, an LU decomposition of the matrix $\mathbf{B}$ can be computed once before starting the integration. In the specific basis developed in the next section, the matrix $\mathbf{B}$ has nonzero entries only on the diagonal and on the second super-diagonal. The cost to solve eq. (3.5) in this case is $3\mathcal{N} - 4$ floating point operations.

3.1 A Legendre basis

The foregoing analysis applies to any set of basis functions p_n^q and p_n^ψ with the assumption $\partial_z p_n^\psi = 0$ at the boundaries. Note that as in [30] this does not imply that $b^\pm = 0$. This section considers a specific choice of the two basis sets. First, since orthogonality is of necessity defined using the L^2 inner product, it is convenient to let $p_n^q(z) = L_{n-1}(z)$ where $L_k(z)$ is the kth order Legendre polynomial, rescaled to the interval $z \in [0, H]$.

Legendre polynomials do not satisfy $\partial_z L_k(z) = 0$ at the boundaries, and therefore cannot be used for p_n^ψ. Shen [32] constructs the following functions

$$\phi_k(z) = L_k(z) - \frac{k(k+1)}{(k+2)(k+3)} L_{k+2}(z). \tag{3.7}$$

These functions form a basis for polynomials with homogeneous Neumann conditions, and we set $p_n^\psi(z) = \phi_{k-1}(z)$.

With this choice of p_n^q and p_n^ψ the **B** matrix is upper triangular with upper bandwidth 2. The mass matrix **M** is pentadiagonal; in fact, an even-odd permutation converts the **M** matrix to a block-diagonal matrix with tridiagonal blocks. The elements of the matrices **M** and **B** are known analytically. The two-norm condition number of **M** is plotted in fig. 3.1 for $\mathcal{N}$ from 100 to 1000. The condition number is just under 10^7 for $\mathcal{N} = 1000$; a condition number of 10^7 is moderate for a banded symmetric positive definite matrix in double precision arithmetic, and $\mathcal{N} = 1000$ is much larger than would be used in most applications. For example, the maximum $\mathcal{N}$ used in the finite-difference simulations of [27, 31] was 128.

The structure of the **L** matrix depends on the stratification $S(z)$ and is in general dense. (For a finite element basis the **L** matrix would be sparse.) The elements of **L** can in general be computed using quadrature. In the computations described in chapter 4 the elements of **L** are computed using Gauss-Legendre quadrature.

With $\mathcal{N}$ Legendre basis functions $q_\mathcal{N}$ is a polynomial in z of degree $\leq \mathcal{N} - 1$, while $\psi_\mathcal{N}$ (and
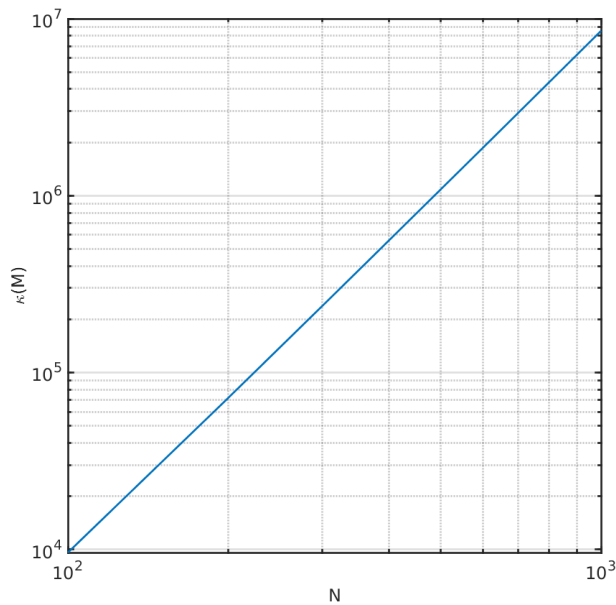
Figure 3.1: The two-norm condition number of the mass matrix $\mathbf{M}$ for $\mathcal{N}$ from 100 to 1000.

hence $\boldsymbol{u}_{\mathcal{N}}$) is a polynomial in z of degree $\leq \mathcal{N} + 1$. The elements of the vector $\dot{\mathbf{q}}$ are therefore integrals of polynomials of degree at most $3\mathcal{N} + 1$. These can be evaluated exactly (up to roundoff error) using Gauss-Legendre quadrature with $1.5\mathcal{N} + 1$ nodes.

Chapter 4

Linear Baroclinic Instability

This section makes a preliminary assessment of the accuracy of the energy-conserving Legendre Galerkin scheme described in the preceding section by applying it to the linear quasigeostrophic baroclinic instability problem. Any configuration of the form

$$\bar{\psi} = -y\bar{u}(z), \quad \bar{q}(z) = -y\frac{\mathrm{d}}{\mathrm{d}z}\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right), \quad \bar{b}^+ = -yf_0\left.\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right|_{z=H}, \quad \bar{b}^- = -yf_0\left.\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right|_{z=0}$$

is an exact solution of the fully nonlinear QG equations eq. (1.1). The linearization of the QG equations about an equilibrium of this form is

$$\partial_t b^+ + \bar{u}^+ \partial_x b^+ - \left(f_0\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)_{z=H} v^+ = 0 \tag{4.1a}$$

$$\partial_t q + \bar{u}\partial_x q + (\partial_y \bar{q} + \beta)v = 0 \tag{4.1b}$$

$$\partial_t b^- + \bar{u}^- \partial_x b^- - \left(f_0\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)_{z=0} v^- = 0. \tag{4.1c}$$

Linear instability about an equilibrium of this form is called 'baroclinic instability.' The linear equations vary only in z, so it is convenient to Fourier transform in the horizontal direction, and to assume exponential growth in time with the form $e^{-\mathrm{i}k_x ct}$ where k_x is the wavenumber in the x

direction and c is the wave phase speed. Also replacing $f_0 \mathrm{d}\bar{u}/\mathrm{d}z$ by $-\partial_y \bar{b}$ on the boundaries leads to equations of the form

$$\bar{u}^+ \hat{b}^+ + (\partial_y \bar{b}^+)\hat{\psi}^+ = c\hat{b}^+ \tag{4.2a}$$

$$\bar{u}\hat{q} + (\partial_y \bar{q} + \beta)\hat{\psi} = c\hat{q} \tag{4.2b}$$

$$\bar{u}^- \hat{b}^- + (\partial_y \bar{b}^-)\hat{\psi}^- = c\hat{b}^- \tag{4.2c}$$

where $\hat{q}$ is the Fourier transform of q and $\hat{b}^\pm$ is the Fourier transform of $b^\pm$. The standard energy-conserving second-order finite difference discretization of the linear stability problem is described in B. We also use a standard, non-energy-conserving Chebyshev collocation method to discretize the linear stability problem; Chebyshev collocation methods for linear problems are described in a variety of places including [38, 6].

4.1   Galerkin Discrete Linear Baroclinic Instability

To discretize this system (eq. (4.2)) according to the methods described in the previous section, one makes Galerkin approximations to both the equilibrium state and the perturbations, then one imposes Galerkin and Petrov-Galerkin conditions on the residuals. First consider how to construct the appropriate Galerkin approximation to the equilibrium state. Although the equilibrium can be completely described by $\bar{u}(z)$, the correct approach within the method described in the foregoing section is to first produce a Galerkin approximation to $\partial_y \bar{q}$ and then invert to find $\bar{u}$. The expansion coefficients in the approximation of $\partial_y \bar{q}$ are arranged into the vector $\bar{\mathbf{q}}_y$, whose entries are

$$(\bar{\mathbf{q}}_y)_n = \frac{\int_0^H p_n^q(z)(\partial_y \bar{q})\mathrm{d}z}{\int_0^H (p_n^q(z))^2 \mathrm{d}z}.$$

(This expression assumes that the basis functions $p_n^q(z)$ are L^2-orthogonal, which is true for the Legendre basis considered here. For a non-orthogonal basis one would have to solve a linear system to find $\bar{\mathbf{q}}_y$.) Once this vector is available, the coefficients in the Galerkin approximation of $\bar{u}$ are obtained by solving the following system

$$\mathbf{L}\bar{\mathbf{u}} = -\mathbf{B}\bar{\mathbf{q}}_y + \frac{f_0}{N^2(H)}(\partial_y \bar{b}^+)\mathbf{p}^+ - \frac{f_0}{N^2(0)}(\partial_y \bar{b}^-)\mathbf{p}^-. \qquad (4.3)$$

The vectors $\mathbf{p}^\pm$ are defined in eq. (3.2).

Unfortunately the matrix $\mathbf{L}$ is singular: its first row and column are zero because $p_1^\psi(z) = 1$ and $\partial_z p_1^\psi = 0$. Fortunately the right hand side is always compatible: the first entry of the right hand side is always zero as well. This statement is substantiated in A. The first entry of $\bar{\mathbf{u}}$ is not constrained by the linear system above; it corresponds to the depth-independent component of $\bar{u}(z)$. The first Galerkin coefficient of $\bar{u}$ can simply be set to

$$\bar{\mathbf{u}}_1 = \frac{\int_0^H p_1^\psi(z)\bar{u}\mathrm{d}z}{\int_0^H (p_1^\psi(z))^2\mathrm{d}z}.$$

The remaining entries of $\bar{\mathbf{u}}$ are obtained by solving the lower-right $\mathcal{N}-1 \times \mathcal{N}-1$ block of eq. (4.3).

Next consider how to construct an appropriate Galerkin approximation to the perturbations about the equilibrium. The Galerkin approximations to the perturbations $\hat{q}$ and $\hat{\psi}$ will have coefficients stored in the vectors $\mathbf{q}$ and $\boldsymbol{\psi}$, respectively. These coefficients are related through the Fourier transform of eq. (3.1), which is

$$-(k_x^2 + k_y^2)\mathbf{M}\boldsymbol{\psi} - \mathbf{L}\boldsymbol{\psi} = \mathbf{B}\mathbf{q} - \frac{f_0}{N^2(H)}b^+\mathbf{p}^+ + \frac{f_0}{N^2(0)}b^-\mathbf{p}^-. \qquad (4.4)$$

With Galerkin approximations to the equilibrium and perturbations in hand, one next inserts these approximations into eq. (4.2b) and requires the residual to be orthogonal to p_n^ψ. One also

inserts $\bar{u}_{\mathcal{N}}$ in place of $\bar{u}$ in eq. (4.2a) and eq. (4.2c), yielding the following linear system

$$\bar{u}_{\mathcal{N}}^+\hat{b}^+ + (\partial_y\bar{b}^+)\hat{\psi}_{\mathcal{N}}^+ = c\hat{b}^+ \tag{4.5a}$$

$$\overline{\mathbf{U}}\mathbf{q} + (\overline{\mathbf{Q}}_y + \beta\mathbf{M})\boldsymbol{\psi} = c\mathbf{B}\mathbf{q} \tag{4.5b}$$

$$\bar{u}_{\mathcal{N}}^-\hat{b}^- + (\partial_y\bar{b}^-)\hat{\psi}_{\mathcal{N}}^- = c\hat{b}^-. \tag{4.5c}$$

Note that $\hat{\psi}_{\mathcal{N}}^\pm = \boldsymbol{\psi}\cdot\mathbf{p}^\pm$, where the vectors $\mathbf{p}^\pm$ are defined in eq. (3.2). With this notation the linear system eq. (4.5) can be written

$$\bar{u}_{\mathcal{N}}^+\hat{b}^+ + (\partial_y\bar{b}^+)\mathbf{p}^+ \cdot \boldsymbol{\psi} = c\hat{b}^+ \tag{4.6a}$$

$$\overline{\mathbf{U}}\mathbf{q} + (\overline{\mathbf{Q}}_y + \beta\mathbf{M})\boldsymbol{\psi} = c\mathbf{B}\mathbf{q} \tag{4.6b}$$

$$\bar{u}_{\mathcal{N}}^-\hat{b}^- + (\partial_y\bar{b}^-)\mathbf{p}^- \cdot \boldsymbol{\psi} = c\hat{b}^-. \tag{4.6c}$$

The matrices $\overline{\mathbf{U}}$ and $\overline{\mathbf{Q}}_y$ have the following entries

$$\overline{\mathbf{U}}_{ij} = \int_0^H p_i^\psi(z)p_j^q(z)\bar{u}_{\mathcal{N}}(z)\mathrm{d}z \tag{4.7}$$

$$\left(\overline{\mathbf{Q}}_y\right)_{ij} = \int_0^H p_i^\psi(z)p_j^\psi(z)(\partial_y\bar{q}_{\mathcal{N}}(z))\mathrm{d}z. \tag{4.8}$$

To obtain an eigenvalue problem, the vector $\boldsymbol{\psi}$ can be eliminated using eq. (4.4)

$$\boldsymbol{\psi} = -\left((k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}\right)^{-1}\mathbf{B}\mathbf{q} + \frac{f_0}{N^2(H)}b^+\psi^+ - \frac{f_0}{N^2(0)}b^-\psi^-. \tag{4.9}$$

For notational convenience the following vectors have been defined

$$\psi^\pm = \left((k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}\right)^{-1}\mathbf{p}^\pm. \tag{4.10}$$

The matrix $(k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}$ is invertible as long as $k_x^2 + k_y^2 \neq 0$ since $\mathbf{M}$ is positive definite and $\mathbf{L}$ is positive semi-definite.

Eliminating ψ leads to a generalized eigenvalue problem of the form

$$\left[\begin{array}{c|c|c} a_{11} & \mathbf{a}_{12}^T & a_{13} \\ \hline \mathbf{a}_{21} & \mathbf{A}_{22} & \mathbf{a}_{23} \\ \hline a_{31} & \mathbf{a}_{32}^T & a_{33} \end{array}\right] \left(\begin{array}{c} \hat{b}^+ \\ \mathbf{q} \\ \hat{b}^- \end{array}\right) = c \left[\begin{array}{c|c|c} 1 & \mathbf{0}^T & 0 \\ \hline 0 & \mathbf{B} & 0 \\ \hline 0 & \mathbf{0}^T & 0 \end{array}\right] \left(\begin{array}{c} \hat{b}^+ \\ \mathbf{q} \\ \hat{b}^- \end{array}\right) \tag{4.11}$$

where the matrix on the left hand side has the following entries

$$a_{11} = \bar{u}_\mathcal{N}^+ + \frac{f_0}{N^2(H)}(\partial_y \bar{b}^+)\mathbf{p}^+ \cdot \psi^+ \tag{4.12a}$$

$$\mathbf{a}_{12} = -\frac{f_0}{N^2(H)}(\partial_y \bar{b}^+)\mathbf{B}^T \left((k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}\right)^{-1}\mathbf{p}^+ \tag{4.12b}$$

$$a_{13} = -\frac{f_0}{N^2(H)}(\partial_y \bar{b}^+)\mathbf{p}^+ \cdot \psi^- \tag{4.12c}$$

$$\mathbf{a}_{21} = (\overline{\mathbf{Q}}_y + \beta\mathbf{M})\psi^+ \tag{4.12d}$$

$$\mathbf{A}_{22} = \overline{\mathbf{U}} - (\overline{\mathbf{Q}}_y + \beta\mathbf{M})\left((k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}\right)^{-1}\mathbf{B} \tag{4.12e}$$

$$\mathbf{a}_{23} = -(\overline{\mathbf{Q}}_y + \beta\mathbf{M})\psi^- \tag{4.12f}$$

$$a_{31} = \frac{f_0}{N^2(0)}(\partial_y \bar{b}^-)\mathbf{p}^- \cdot \psi^+ \tag{4.12g}$$

$$\mathbf{a}_{32} = -\frac{f_0}{N^2(0)}(\partial_y \bar{b}^-)\mathbf{B}^T \left((k_x^2 + k_y^2)\mathbf{M} + \mathbf{L}\right)^{-1}\mathbf{p}^- \tag{4.12h}$$

$$a_{33} = \bar{u}_{\mathcal{N}}^- - \frac{f_0}{N^2(0)}(\partial_y \bar{b}^-)\mathbf{p}^- \cdot \boldsymbol{\psi}^-. \tag{4.12i}$$

To solve a specific linear baroclinic instability problem, one chooses external parameters f_0, β, and $N^2(z)$, and an equilibrium state $\bar{u}(z)$. Then one computes eigenvalues of the generalized eigenvalue problem, typically over some range of values of k_x and k_y; eigenvalues c with positive imaginary part are associated with linearly unstable solutions. Code to set up and solve the Galerkin and finite-difference linear stability problems is available in [45].

4.2   The Eady Problem

The classical Eady problem is defined by constant $N^2(z)$, linear velocity $\bar{u}(z)$, and $\beta = 0$. In this case the baroclinic modes of [30] are simply Fourier modes and are tractable analytically and computationally. The exact linear perturbation equations eq. (4.2) are also analytically solvable in the Eady problem (see, e.g. [42, Chapter 6]), which makes for a good test problem. This section sets $N^2(z) = f_0^2 = 1$, $H = 1$, and $\bar{u} = z$. The most unstable solutions are found along the axis $k_y = 0$, so the generalized eigenvalue problem is solved for a range of k_x.

Figure 4.1 shows the results of the linear Eady problem with $\mathcal{N} = 7$, compared to the analytical results from [42, Chapter 6]. The growth rates (upper left) and wave speeds (lower left) as a function of k_x are extremely well reproduced with only $\mathcal{N} = 7$ basis functions, to the point where the plots are indistinguishable. (Note that the streamfunction of the most unstable mode, $k_x \approx 1.6$, is approximated with a polynomial of degree 9.) The amplitude (upper right) and phase (lower right) of the eigenfunction corresponding to the most unstable mode are also very accurate using only $\mathcal{N} = 7$. The approximate eigenfunction has $\partial_z \hat{\psi} = 0$ on the boundaries, while the exact does not; nevertheless, the approximate eigenfunction still converges pointwise to the true eigenfunction with increasing $\mathcal{N}$ (not shown).
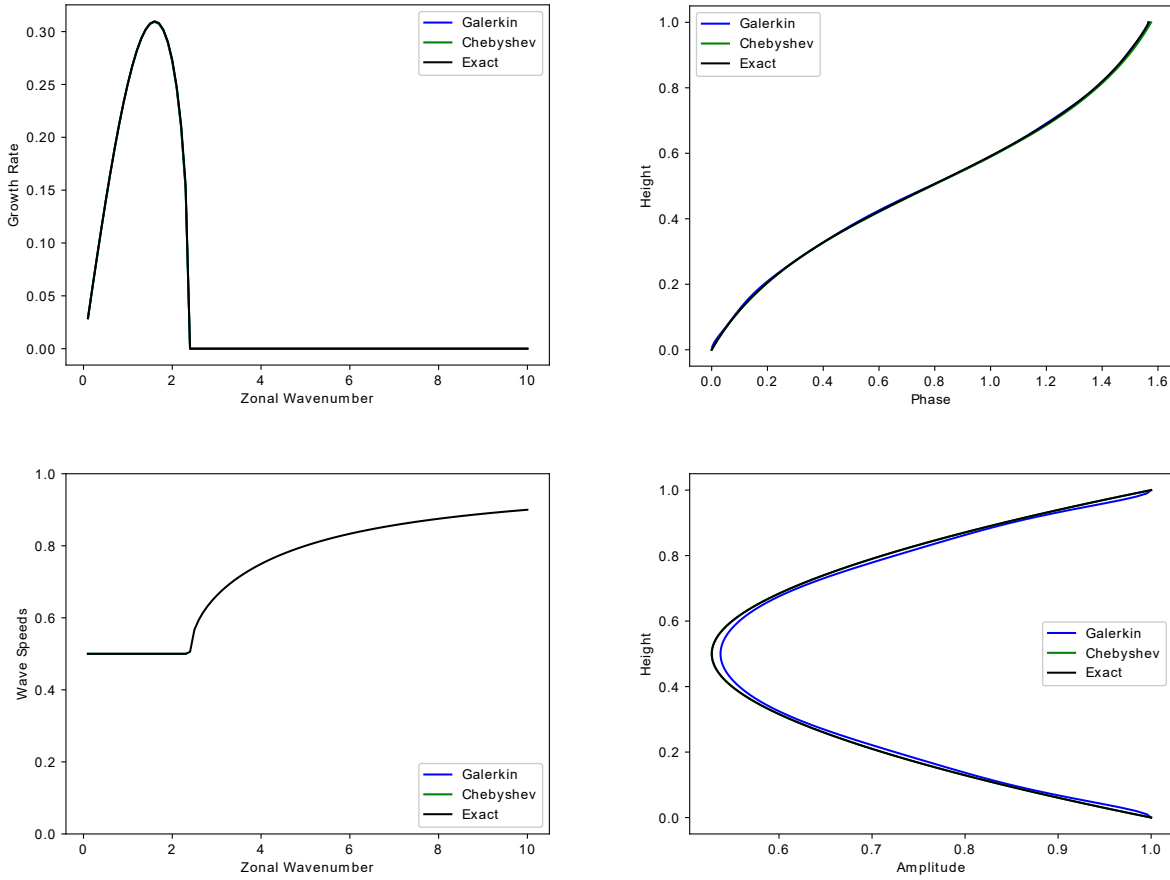
Figure 4.1: Comparing Galerkin (blue) and Chebyshev (green) with $\mathcal{N} = 7$ to exact results (black) in the Eady Problem with $k_y = 0$. Upper left: Growth rates as a function of k_x. Lower left: Wave speeds (real part of the eigenvalue c) as a function of k_x. Upper right: Complex phase as a function of z for the eigenfunction associated with the fastest-growing mode. Lower right: Amplitude as a function of z for the eigenfunction associated with the fastest-growing mode, normalized to 1 at $z = 0$. The Galerkin and Chebyshev methods are so accurate that the results are indistinguishable from exact in the above plots, except for the Galerkin method in the lower-right panel.

Figure 4.2 shows the absolute value of the error in the growth rate of the most unstable mode as a function of $\mathcal{N}$ for the Galerkin approximation, the finite-difference discretization (see B), and the Chebyshev discretization. Note that in the finite-difference approximation there are $\mathcal{N}$ degrees of freedom, while in the Galerkin approximation there are $\mathcal{N} + 2$ degrees of freedom: one for each Galerkin coefficient and one for each surface buoyancy. The error in the Galerkin approximation decreases as $\mathcal{O}(\mathcal{N}^{-3})$, while the error in the finite difference approximation decreases
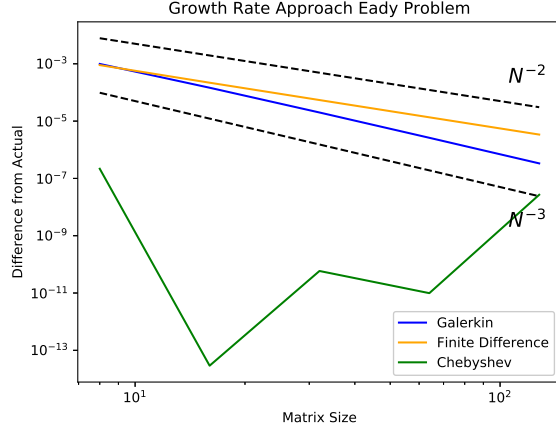
Figure 4.2: Error in the growth rate of the fastest growing mode as a function of $\mathcal{N}$ for the Galerkin (blue), Chebyshev (green), and finite-difference (orange) methods.

only quadratically. The Chebyshev method exhibits spectral accuracy as expected, with far greater accuracy than the other two methods; accuracy reaches a plateau with increasing $\mathcal{N}$, which may be a result of the ill-conditioning of Chebyshev differentiation matrices [11, 38, 6]. Problems of this sort are avoided by the Chebyshev Galerkin methods of [20].

Convergence of spectral methods can be limited by lack of smoothness in the functions being approximated, but the eigenfunctions of the Eady problem are entire functions expressible as a sum of hyperbolic sine and cosine functions [42]. The fact that the Galerkin method converges algebraically rather than exponentially is therefore presumably due to the mismatch between the homogeneous boundary conditions satisfied by the basis functions p_n^ψ and the inhomogeneous boundary conditions satisfied by the true eigenfunctions. Since this mismatch only occurs in the value of the derivative on the boundary the approximate eigenfunctions still converge to the true eigenfunctions with increasing $\mathcal{N}$ (as in [30]), though the rate of convergence is slower than it would be if the boundary conditions matched.

4.3   The Phillips Problem

The instability in the Eady problem is driven by interacting edge waves, and is therefore of a type not often seen in the atmosphere or ocean. Another classical linear baroclinic instability problem that, unlike the Eady problem, is observed in the oceans occurs when the potential vorticity gradient $d\bar{q}/dy$ changes sign in the interior of the fluid. The canonical representation of this kind of instability is the 'Phillips' problem which is distinguished by an equilibrium velocity that has zero shear at the top and bottom surface, and a single sign change in the potential vorticity in the interior. We construct a 'Phillips' problem of this type as follows

$$f_0 = N(z) = H = 1, \ \ \beta = 3.1, \ \ \ \bar{u} = -\pi^{-1}\cos(\pi z), \ \ \bar{q} = \pi\cos(\pi z)y. \quad (4.13)$$

The total potential vorticity gradient is $3.1 + \pi\cos(\pi z)$. The negative potential vorticity gradient near the top boundary has small amplitude, and as a result the equilibrium is only slightly above the threshold for instability. There is only a small range of wavenumbers near $k_x = 3$ that are unstable, as shown in the left panel of fig. 4.3.

The center and right panels of fig. 4.3 show the absolute value of the error in the growth rate at $k_x = 3$ as a function of $\mathcal{N}$ for the Galerkin (center), Chebyshev (center), and finite-difference (right) methods. The center panel uses a logarithmic scale on the growth rate axis and a linear scale on the $\mathcal{N}$ axis to show that the Galerkin and Chebyshev methods are converging exponentially rather than algebraically. Although both methods converge exponentially, the Galerkin method is more accurate than the Chebyshev method. The right panel uses a logarithmic scaling on both axes to show that the finite-difference method is converging quadratically, as usual. Exponential convergence is expected for the Galerkin method in this case, since the eigenfunctions are smooth and have the same homogeneous Neumann boundary conditions as the basis functions p_n^ψ. Note that the accuracy of the finite difference scheme with $\mathcal{N} = 256$ can be achieved by the Galerkin scheme with ten times fewer degrees of freedom.
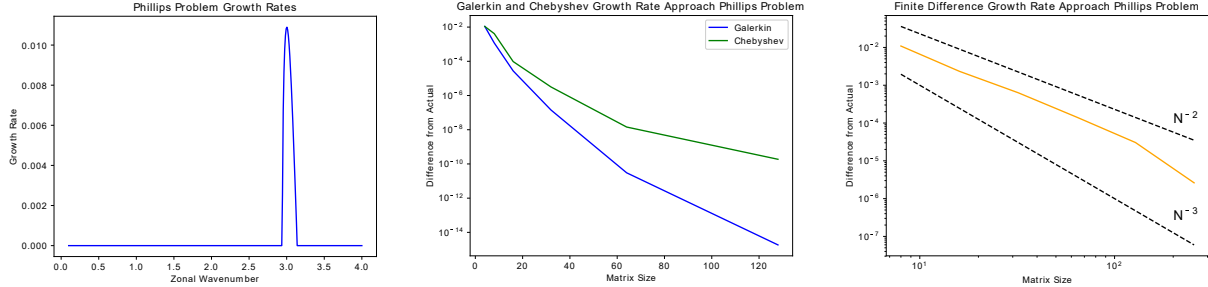
Figure 4.3: Growth rates in the Phillips problem. Left: Growth rates versus k_x for the Galerkin method with $\mathcal{N} = 256$. Center: Error in the growth rate at $k_x = 3$ as a function of $\mathcal{N}$ for the Galerkin (blue) and Chebyshev (green) methods. Right: Error in the growth rate at $k_x = 3$ as a function of $\mathcal{N}$ for the finite-difference method.

4.4 A Charney-Type Problem

Another common type of baroclinic instability in the ocean is driven by the interaction of an edge wave with a Rossby wave in the interior of the fluid [39]. A canonical problem describing this kind of instability is the Charney problem, but the canonical Charney problem is posed in a semi-infinite domain with no upper surface. A Charney-type problem more relevant to the ocean is defined by having a nonzero shear $\mathrm{d}\bar{u}/\mathrm{d}z$ at the top surface, zero shear at the bottom surface, and a constant potential vorticity gradient $\mathrm{d}\bar{q}/\mathrm{d}y$ in the interior. We construct such a Charney-type problem with exponential rather than constant stratification to demonstrate the ability of the Galerkin method to handle non-constant stratification. The equilibrium state is defined as follows

$$f_0 = \beta = H = 1, \quad N^2(z) = e^{6z-6}, \quad \bar{u} = \frac{1}{54}\left(3e^{6z-6}(6z-1) - 2 - e^{-6}\right), \quad \bar{q} = -2y. \tag{4.14}$$

Both the stratification and the velocity are surface-intensified (the velocity $\bar{u}(z)$ is shown in the lower right panel of fig. 4.4), which leads to surface-intensification of the unstable linear eigenfunctions.

Figure 4.4 shows the results of the linear Charney-type problem. The upper left panel of fig. 4.4 shows the growth rate as a function of k_x for $\mathcal{N} = 32$ and 256 using the Galerkin method; the finite difference and Chebyshev methods are extremely similar (not shown). The upper right panel of fig. 4.4 shows the absolute value of the error in the growth rate of the most unstable mode
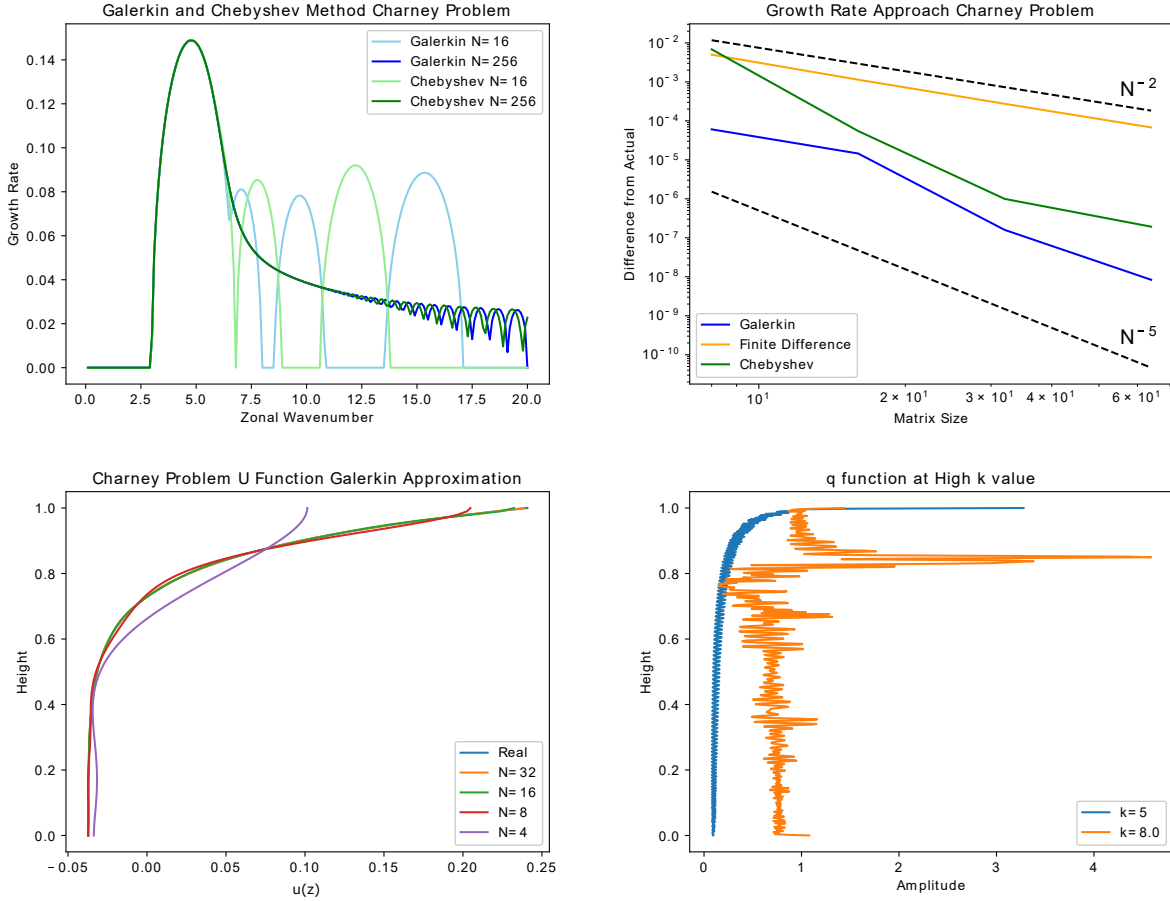
Figure 4.4: Growth rates in the Charney-type problem. Upper left: Growth rates versus k_x for the Galerkin (blue) and Chebyshev (green) methods with $\mathcal{N} = 32$ (pale) and $\mathcal{N} = 256$ (dark). Upper right: Error in the growth rate of the fastest-growing mode as a function of $\mathcal{N}$ for the Galerkin (blue), Chebyshev (green), and finite-difference (orange) methods. Lower left: Background velocity $\bar{u}(z)$ (blue) and Galerkin approximations $\bar{u}_{\mathcal{N}}(z)$ for $\mathcal{N} = 4$, 8, 16, and 32. Lower right: Amplitude of the eigenfunction $|\hat{q}(z)|$ at $k_x = 5$ (blue) and 8 (orange) using $\mathcal{N} = 256$.

as a function of $\mathcal{N}$ for Galerkin, Chebyshev, and finite-difference methods. The growth rate in the Galerkin and Chebyshev methods converge approximately quintically ($\mathcal{O}(\mathcal{N}^{-5})$) while the finite-difference method converges approximately quadratically. Unlike the Eady problem, the Galerkin approximation is more accurate than both the Chebyshev and the finite difference approximations for the entire range of $\mathcal{N}$. The accuracy of the finite difference scheme with $\mathcal{N} = 256$ can be achieved with the Galerkin scheme with ten times fewer degrees of freedom.

A distinctive feature of the Charney-type problem is the presence of weak instability at

small scales (large k_x), as shown in the upper left panel of fig. 4.4. Representation of the small-scale instabilities clearly requires large $\mathcal{N}$; this is true in both the Galerkin and finite-difference methods, which behave similarly for large k_x (not shown). These small-scale unstable modes result from the interaction of a Rossby wave in a thin layer near the upper surface with an edge wave propagating along the surface. The lower-right panel of fig. 4.4 shows the eigenfunction structure $|\hat{q}_{\mathcal{N}}(z)|$ at both $k_x = 5$ and $k_x = 8$, both computed using $\mathcal{N} = 256$; the near-surface layer is evident in the eigenfunction at $k_x = 8$. (See also fig. 7 of [30].) Even at $\mathcal{N} = 256$ these modes are clearly poorly resolved. As a result of the near-surface nature of the instability, the instability is especially sensitive to the representation of the equilibrium background velocity $\bar{u}(z)$ near the boundary. The lower left panel of fig. 4.4 shows the background velocity $\bar{u}$ along with the Galerkin approximations $\bar{u}_{\mathcal{N}}(z)$ for $\mathcal{N} = 4$, 8, 16, and 32. Convergence of $\bar{u}_{\mathcal{N}}(z)$ to $\bar{u}(z)$ is slow near the upper boundary because the basis functions satisfy $\partial_z p_n^{\psi} = 0$ at the boundary, while the equilibrium background profile $\bar{u}(z)$ has $\partial_z \bar{u} \neq 0$ at the boundary. This slow convergence of $\bar{u}_{\mathcal{N}}(z)$ near the boundary is ultimately why the small-scale (high k_x) instability shown in the upper left panel of fig. 4.4 converges slowly.

Chapter 5

Nonlinear Simulations

This section makes a preliminary assessment of the accuracy of the energy-conserving Galerkin scheme described in chapter 3 by using it in fully-nonlinear simulations. Assessment of the method in comparison with other competing methods using simulations of the full system eq. (1.1) will be postponed to a future work; this section uses a simplified exact solution of the full system with $q = \beta = 0$. We further specify $f_0 = N = H = 1$, which leads to the following system of two-dimensional partial differential equations

$$\partial_t b^+ + \boldsymbol{u}^+ \cdot \nabla b^+ = 0, \quad \partial_z \psi = b^+ \text{ at } z = 1 \tag{5.1a}$$

$$\left(\nabla^2 + \partial_z^2\right) \psi = 0 \tag{5.1b}$$

$$\partial_t b^- + \boldsymbol{u}^- \cdot \nabla b^- = 0, \quad \partial_z \psi = b^- \text{ at } z = 0. \tag{5.1c}$$

This system is very close to the nonlinear Eady model [40] except that it lacks both a mean shear and dissipation terms. This simplified version of the nonlinear Eady model is used here because it conserves energy exactly, unlike the full nonlinear Eady model where interaction with the mean shear generates energy, and is therefore an apt test case for comparing our energy-conserving
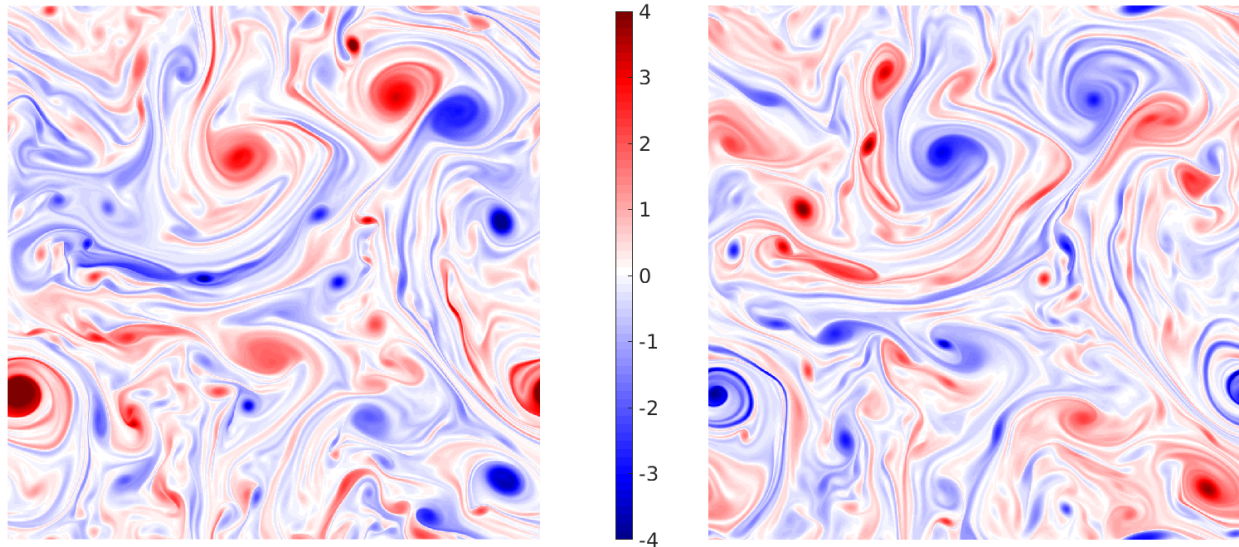
Figure 5.1: The state of b^+ (left) and b^- (right) used to both estimate inversion accuracy and to initialize simulations tracking energy changes.

method. The inversion from $b^\pm$ to $\psi^\pm$ can be solved analytically by means of a Fourier transform

$$\begin{pmatrix} \hat{\psi}^+ \\ \hat{\psi}^- \end{pmatrix} = \frac{1}{k} \begin{bmatrix} \coth(k) & -\operatorname{csch}(k) \\ \operatorname{csch}(k) & -\coth(k) \end{bmatrix} \begin{pmatrix} \hat{b}^+ \\ \hat{b}^- \end{pmatrix} \tag{5.2}$$

where $k^2 = k_x^2 + k_y^2$. The three methods used in the preceding section are also used here to solve for $\psi^\pm$. The Galerkin method uses eq. (3.1), the finite difference method is described in B, and the Chebyshev collocation method is described in several places including [38, 6].

The domain is a periodic square of width 8π, and the advection terms in eq. (5.1a) and eq. (5.1c) are discretized using a dealiased Fourier spectral method with 1024 points in each direction; this method is energy-conserving provided that the vertical discretization is also energy-conserving. Time integration is achieved via a fourth-order semi-implicit method as described in [15]. This time integration method is not energy-conserving. Energy conservation typically requires an implicit method, but energy non-conservation due to time integration errors is usually small, and is small in our simulations described below. For a wide class of systems energy conservation is required in the spatial discretization but not in the temporal discretization in order to achieve time-asymptotic accuracy in the representation of a system's ergodic invariant measure [12, 44].

The code used here is a slight adaptation of the publicly-available code [14].

We begin by spinning up the system from random initial conditions using the exact analytical inversion. The simulation was stopped when b^+ and b^- reached realistic values shown in fig. 5.1. We used this configuration of surface buoyancy to compare the inversion accuracy of the three approximate methods: Galerkin, Chebyshev, and finite-difference. We computed the value of ψ^+ and $\boldsymbol{u}^+$ using the exact inversion and the three approximate methods, and then computed the Fourier transform of the error in $\boldsymbol{u}^+$. The one-dimensional spectrum of the velocity error, i.e. the error kinetic energy spectrum, for each method is shown in fig. 5.2 for different values of $\mathcal{N}$; for comparison, the true kinetic energy spectrum is also shown. The Galerkin method (center panel) is clearly more accurate than the finite-difference method (left panel) for a fixed value of $\mathcal{N}$; the error also decreases faster with increasing $\mathcal{N}$ for the Galerkin method than for the finite difference method. On the other hand, the Chebyshev method is so much more accurate than the other two methods that there is essentially no comparison. This echoes the results from the Eady linear stability problem, which also has constant stratification $S(z) = 1$. In the linear stability problem with non-constant stratification the Chebyshev method was slightly less accurate than the Galerkin method, so the extreme accuracy of the Chebyshev method in fig. 5.2 may not be indicative of its performance generally. All three methods have small relative errors at large horizontal scales, because these horizontal scales are associated with nearly depth-independent vertical structures. In all three methods the relative errors are largest at small horizontal scales, where the vertical structure is strongly surface-intensified.

We next assessed the impact of energy conservation by running four simulations from the initial condition shown in fig. 5.1: one with the exact inversion and one with each of the three approximate inversions. These simulations were run for 50 time units (which corresponds to about 40 eddy turnover times), and the total energy was tracked. The finite difference simulation used 128 levels, the Galerkin simulation used $\mathcal{N} = 16$, and Chebyshev simulations were run with $\mathcal{N} = 4, 8$, and 16. In all simulations the energy changed by less than 1% over the course of the simulation,
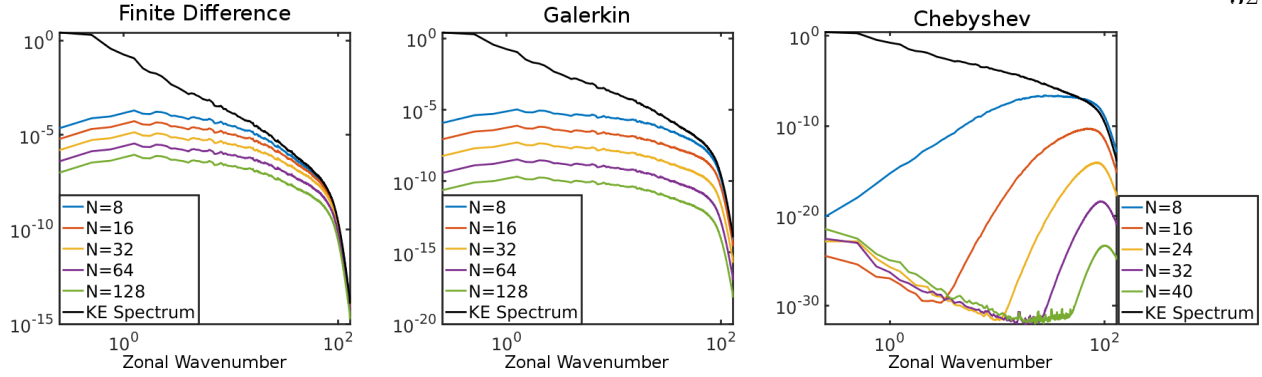
Figure 5.2: Error kinetic energy spectra at the top surface for the three methods: finite difference (left), Galerkin (center), and Chebyshev (right). Note (i) that the vertical axis scaling is different in each panel, and (ii) that the values of $\mathcal{N}$ for the Chebyshev method are smaller than for the other two methods. The kinetic energy spectrum is shown in black in each panel for comparison.

excepting the Chebyshev simulation with $\mathcal{N} = 4$, which only changed by 1.3%. The degree of energy conservation for all three methods was thus near perfect, despite the Chebyshev method not being guaranteed to conserve energy; the observed changes in energy are attributable to the fact that the time integration scheme does not conserve energy. Further fully nonlinear simulations with non-constant stratification and $q \neq 0$ are needed to fully assess the relative merits of the three methods. These are left for future work.

Chapter 6

Conclusions

This article presents an energy-conserving Galerkin approximation scheme for the vertical direction of the full QG system with active surface buoyancy. The scheme generalizes the Galerkin scheme of [30]. The method in [30] uses Sturm-Liouville eigenfunctions as a basis to approximate both the potential vorticity q and the streamfunction ψ, but these functions are in most cases computationally intractable. The scheme presented here generalizes the method of [30] to allow an arbitrary basis for q and any basis for ψ that satisfies homogeneous Neumann boundary conditions at the top and bottom surfaces. Attention is then focused on a Legendre basis for q, and a recombined Legendre basis from [32] for ψ. Energy is defined using an unweighted L^2 norm, based on an unweighted L^2 inner product; Legendre polynomials were used because their orthogonality with respect to the unweighted L^2 inner product make them particularly convenient. Chebyshev polynomials are used more commonly than Legendre polynomials in spectral methods partly because of the existence of a fast transform for Chebyshev polynomials, but Chebyshev polynomials are orthogonal with respect to a weighted L^2 inner product and are therefore less convenient than Legendre polynomials in the current setting.

The method was tested and compared to the standard energy-conserving second-order finite-difference method and to a Chebyshev collocation method in the context of linear stability calculations. In these calculations the Galerkin scheme converged much faster with respect to increasing $\mathcal{N}$ (vertical resolution) than the finite difference scheme. The eigenvalues computed with the finite difference scheme converged quadratically, while those computed with the Galerkin scheme con-

verged either at fifth order (in cases with nonzero surface buoyancy) or exponentially (with zero surface buoyancy). In some cases the Galerkin scheme was able to achieve comparable accuracy to the finite difference method using ten times fewer grid points. In one case the Chebyshev scheme was very significantly more accurate than the Galerkin scheme, but in two other cases the Galerkin scheme was similar to and slightly more accurate than the Chebyshev method.

The method was then compared to the finite-difference and Chebyshev methods in the context of fully-nonlinear simulations in the simplified setting of $q = \beta = 0$, and with constant stratification. In this simplified setting the only dependent variables are surface buoyancies making the dynamics two-dimensional, though the surface buoyancies are still coupled by a three-dimensional elliptic inversion. All three methods conserved energy to extremely high accuracy, despite the fact that the Chebyshev scheme is not guaranteed to do so a priori. The Galerkin method was significantly more accurate than the finite difference method, but was not nearly as accurate as the Chebyshev method.

The scheme presented here is ultimately intended for use in a fully nonlinear, fully three-dimensional setting with nonzero potential vorticity q. The linear stability computations and simplified nonlinear simulations presented here only give limited insight into the accuracy of the scheme for the intended application. The scheme presented here will be tested in a fully nonlinear and three-dimensional setting in future work. Finite element bases rather than a global polynomial basis could also be explored as a means of enhancing sparsity in future work.

Bibliography

[1] A Arakawa. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. J. Comput. Phys., 1(1):119–143, 1966.

[2] AF Bennett and PE Kloeden. The periodic quasigeostrophic equations: existence and uniqueness of strong solutions. P. Roy. Soc. Edinb. A, 91(3-4):185–203, 1982.

[3] Alfred J Bourgeois and J Thomas Beale. Validity of the quasigeostrophic model for large-scale flow in the atmosphere and ocean. SIAM J. Math. Anal., 25(4):1023–1068, 1994.

[4] FP Bretherton. Critical layer instability in baroclinic flows. Q. J. Roy. Meteor. Soc., 92(393):325–334, 1966.

[5] Jörn Callies, Glenn Flierl, Raffaele Ferrari, and Baylor Fox-Kemper. The role of mixed-layer instabilities in submesoscale turbulence. J. Fluid Mech., 788:5–41, 2016.

[6] C Canuto, MY Hussaini, A Quarteroni, and TA Zhang. Spectral Methods Fundamentals in Single Domains. Springer-Verlag Berlin Heidelberg, 2006.

[7] T Colin. The Cauchy problem and the continuous limit for the multilayer model in geophysical fluid dynamics. SIAM J. Math. Anal., 28(3):516–529, 1997.

[8] P Constantin and HQ Nguyen. Global weak solutions for SQG in bounded domains. Comm. Pure Appl. Math., 2017.

[9] Peter Constantin, Andrew J Majda, and Esteban Tabak. Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar. Nonlinearity, 7(6):1495, 1994.

[10] Glenn R Flierl. Models of vertical structure and the calibration of two-layer models. Dyn. Atmos. Oceans, 2(4):341–381, 1978.

[11] Bengt Fornberg. A Practical Guide to Pseudospectral Methods. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 1996.

[12] S. Gottlieb, F. Tone, C. Wang, X. Wang, and D. Wirosoetisno. Long time stability of a classical efficient scheme for two-dimensional Navier–Stokes equations. SIAM J. Numer. Anal., 50(1):126–150, 2012.

[13] Ian Grooms. A Gaussian-product stochastic Gent–McWilliams parameterization. Ocean Model., 106:27–43, 2016.

[14] Ian Grooms. QG_DNS. `https://github.com/iangrooms/QG_DNS`, 2019.

[15] Ian Grooms and Andrew J Majda. Stochastic superparameterization in quasigeostrophic turbulence. J. Comput. Phys., 271:78–98, 2014.

[16] Ian Grooms and Louis-Philippe Nadeau. The effects of mesoscale ocean–atmosphere coupling on the quasigeostrophic double gyre. Fluids, 1(4):34, 2016.

[17] Andrew Mc C Hogg, William K Dewar, Peter D Killworth, and Jeffrey R Blundell. A quasi-geostrophic coupled model (Q-GCM). Mon. Weather Rev., 131(10):2261–2278, 2003.

[18] John K Hunter and Bruno Nachtergaele. Applied Analysis. World Scientific, 2001.

[19] Jordi Isern-Fontanet, Guillaume Lapeyre, Patrice Klein, Bertrand Chapron, and Matthew W Hecht. Three-dimensional reconstruction of oceanic mesoscale currents from surface information. J. Geophys. Res.-Oceans, 113(C9), 2008.

[20] Keith Julien and Mike Watson. Efficient multi-dimensional solution of PDEs using Chebyshev spectral methods. J. Comput. Phys., 228(5):1480–1503, 2009.

[21] SA Karabasov, Pavel S Berloff, and VM Goloviznin. CABARET in the ocean gyres. Ocean Model., 30(2-3):155–168, 2009.

[22] Patrice Klein, Guillaume Lapeyre, Guillaume Roullet, Sylvie Le Gentil, and Hideharu Sasaki. Ocean turbulence at meso and submesoscales: connection between surface and interior dynamics. Geophys. Astro. Fluid, 105(4-5):421–437, 2011.

[23] JH LaCasce and A Mahadevan. Estimating subsurface horizontal and vertical velocities from sea-surface temperature. J. Marine Res., 64(5):695–721, 2006.

[24] G Lapeyre and P Klein. Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. J. Phys. Ocean., 36(2):165–176, 2006.

[25] Guillaume Lapeyre. Surface quasi-geostrophy. Fluids, 2(1):7, 2017.

[26] AJ Majda. Introduction to PDEs and Waves for the Atmosphere and Ocean. American Mathematical Society, 2003.

[27] M Jeroen Molemaker, James C McWilliams, and Xavier Capet. Balanced and unbalanced routes to dissipation in an equilibrated eady flow. J. Fluid Mech., 654:35–63, 2010.

[28] Matthew D Novack and Alexis F Vasseur. Global in time classical solutions to the 3D quasi-geostrophic system for large initial data. Commun. Math. Phys., 358(1):237–267, 2018.

[29] J Pedlosky. Geophysical Fluid Dynamics. Springer-Verlag New York, second edition edition, 1987.

[30] Cesar B Rocha, William R Young, and Ian Grooms. On Galerkin approximations of the surface active quasigeostrophic equations. J. Phys. Ocean., 46(1):125–139, 2016.

[31] G Roullet, JC McWilliams, X Capet, and MJ Molemaker. Properties of steady geostrophic turbulence with isopycnal outcropping. J. Phys. Ocean., 42(1):18–38, 2012.

[32] Jie Shen. Efficient spectral-Galerkin method I. Direct solvers of second-and fourth-order equations using Legendre polynomials. SIAM J. Sci. Comput., 15(6):1489–1505, 1994.

[33] IV Shevchenko and PS Berloff. Multi-layer quasi-geostrophic ocean dynamics in eddy-resolving regimes. Ocean Model., 94:1–14, 2015.

[34] K Shafer Smith and Geoffrey K Vallis. The scales and equilibration of midocean eddies: Freely evolving flow. J. Phys. Ocean., 31(2):554–571, 2001.

[35] K Shafer Smith and Geoffrey K Vallis. The scales and equilibration of midocean eddies: Forced–dissipative flow. J. Phys. Ocean., 32(6):1699–1720, 2002.

[36] K Shafer Smith and Jacques Vanneste. A surface-aware projection basis for quasigeostrophic flow. J. Phys. Ocean., 43(3):548–562, 2013.

[37] Benjamin A Storer, Francis J Poulin, and Claire Ménesguen. The dynamics of quasigeostrophic lens-shaped vortices. J. Phys. Ocean., 48(4):937–957, 2018.

[38] L. Trefethen. Spectral Methods in MATLAB. Society for Industrial and Applied Mathematics, 2000.

[39] Ross Tulloch, John Marshall, Chris Hill, and K Shafer Smith. Scales, growth rates, and spectral fluxes of baroclinic instability in the ocean. J. Phys. Ocean., 41(6):1057–1076, 2011.

[40] Ross Tulloch and K Shafer Smith. A note on the numerical representation of surface dynamics in quasigeostrophic turbulence: Application to the nonlinear Eady model. J. Atmos. Sci., 66(4):1063–1068, 2009.

[41] Ross Tulloch and K Shafer Smith. Quasigeostrophic turbulence with explicit surface dynamics: Application to the atmospheric energy spectrum. J. Atmos. Sci., 66(2):450–467, 2009.

[42] Geoffrey K Vallis. Atmospheric and oceanic fluid dynamics. Cambridge University Press, 2017.

[43] Antoine Venaille, Geoffrey K Vallis, and K Shafer Smith. Baroclinic turbulence in the ocean: Analysis with primitive equation and quasigeostrophic simulations. J. Phys. Ocean., 41(9):1605–1623, 2011.

[44] Xiaoming Wang. Numerical algorithms for stationary statistical properties of dissipative dynamical systems. Discrete Cont. Dyn.-A, 36(8):4599–4618, 2016.

[45] Matthew Watwood and Ian Grooms. QG-Galerkin. `https://github.com/mwatwood-cu/QG-Galerkin`, 2018.

Appendix A

Compatibility of the system eq. (4.3)

The first entry of the right hand side of eq. (4.3) is a sum of three components, and this appendix demonstrates that the sum of these three components is zero. The proof relies on the fact that $p_1^\psi(z) = 1$, $p_j^q(z) = L_{j-1}(z)$, and $p_1^q(z) = 1$.

The first component of the right hand side of eq. (4.3) is the first entry of the vector $-\mathbf{B}\bar{\mathbf{q}}_y$. The first row of the matrix $\mathbf{B}$ has elements

$$\int_0^H p_1^\psi(z)p_j^q(z)\mathrm{d}z = \int_0^H p_j^q(z)\mathrm{d}z = H\delta_{1j}$$

where δ_{ij} is the Kronecker delta. The first entry of $-\mathbf{B}\bar{\mathbf{q}}_y$ is thus simply the first entry of $\bar{\mathbf{q}}_y$ multiplied by $-H$.

The first entry of $\bar{\mathbf{q}}_y$ is

$$\frac{\int_0^H p_1^q(z)\partial_y\bar{q}(z)\mathrm{d}z}{\int_0^H (p_1^q(z))^2\mathrm{d}z} = -\frac{1}{H}\int_0^H \frac{\mathrm{d}}{\mathrm{d}z}\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{d}z$$

$$= \frac{1}{H}\left[S(0)\left.\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right|_{z=0} - S(H)\left.\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right|_{z=H}\right]$$

$$= \frac{1}{H}\left[\frac{f_0}{N^2(H)}\partial_y\bar{b}^+ - \frac{f_0}{N^2(0)}\partial_y\bar{b}^-\right].$$

Multiplying this by $-H$ results in an expression that exactly cancels the remaining two components on the right hand side of eq. (4.3).

Appendix B

Finite Difference Discretization

This section recalls the standard finite-difference discretization of the QG equations, which can be found in, e.g., [29] and [42]. A derivation with careful treatment of surface buoyancy and unequal spacing can be found in [16].

Let $\Delta_z = 1/\mathcal{N}$ be the grid spacing where $\mathcal{N}$ is the number of vertical levels. Both ψ and q are tracked at $\mathcal{N}$ points starting at $z_1 = \Delta_z/2$ and ending at $z_{\mathcal{N}} = 1 - \Delta_z/2$. The finite difference approximation to $\nabla^2 \psi + \partial_z(S(z)\partial_z \psi)$ at an interior point z_k ($k \neq 1, \mathcal{N}$) is

$$\left(\nabla^2 \psi + \partial_z(S(z)\partial_z \psi)\right)|_{z=z_k} \approx \nabla^2 \psi_k + \frac{1}{\Delta_z}\left[S_k \frac{\psi_{k+1} - \psi_k}{\Delta_z} - S_{k-1}\frac{\psi_k - \psi_{k-1}}{\Delta_z}\right] = q_k \qquad \text{(B.1)}$$

where $S_k = S(k\Delta_z)$. At the boundaries we have the following approximations

$$\left(\nabla^2 \psi + \partial_z(S(z)\partial_z \psi)\right)|_{z=z_1} \approx \nabla^2 \psi_1 + \frac{1}{\Delta_z}\left[S_1 \frac{\psi_2 - \psi_1}{\Delta_z} - \frac{f_0}{N^2(0)}b^-\right] = q_1 \qquad \text{(B.2)}$$

$$\left(\nabla^2 \psi + \partial_z(S(z)\partial_z \psi)\right)|_{z=z_{\mathcal{N}}} \approx \nabla^2 \psi_{\mathcal{N}} + \frac{1}{\Delta_z}\left[\frac{f_0}{N^2(H)}b^+ - S_{\mathcal{N}-1}\frac{\psi_{\mathcal{N}} - \psi_{\mathcal{N}-1}}{\Delta_z}\right] = q_{\mathcal{N}}. \qquad \text{(B.3)}$$

As discussed in [16], if one defines

$$Q_1 = q_1 + \frac{f_0}{\Delta_z N^2(0)}b^- = \nabla^2 \psi_1 + \frac{1}{\Delta_z}\left[S_1 \frac{\psi_2 - \psi_1}{\Delta_z}\right], \qquad \text{(B.4a)}$$

$$Q_{\mathcal{N}} = q_{\mathcal{N}} - \frac{f_0}{\Delta_z N^2(H)}b^+ = \nabla^2 \psi_{\mathcal{N}} - \frac{1}{\Delta_z}\left[S_{\mathcal{N}-1}\frac{\psi_{\mathcal{N}} - \psi_{\mathcal{N}-1}}{\Delta_z}\right] \qquad \text{(B.4b)}$$

Then the fully nonlinear system dynamics are controlled entirely by the following system

$$\partial_t Q_1 + J[\psi_1, Q_1] + \beta \partial_x \psi_1 = 0 \tag{B.5a}$$

$$\partial_t q_k + J[\psi_k, Q_k] + \beta \partial_x \psi_k = 0, \quad k = 2, \ldots, \mathcal{N} - 1 \tag{B.5b}$$

$$\partial_t Q_\mathcal{N} + J[\psi_\mathcal{N}, Q_\mathcal{N}] + \beta \partial_x \psi_\mathcal{N} = 0. \tag{B.5c}$$

The only caveat is that by evolving this system one knows Q_1 and $Q_\mathcal{N}$ but not $b^\pm$ or q_1 and $q_\mathcal{N}$, but the dynamics of ψ_k are completely controlled by the above system: eq. (B.5) for the dynamics and eq. (B.1) and eq. (B.4) for the PV inversion.

The discrete version of the linear stability problem is straightforward in the finite difference approximation. One can start with eq. (4.2) and then discretize as described above. The discrete finite difference problem takes the form of the following generalized eigenvalue problem

$$\left[\overline{\mathbf{U}}_{FD} \left((k_x^2 + k_y^2)\mathbf{I} + \mathbf{L}_{FD} \right) - \left(\overline{\mathbf{Q}}_{y,FD} + \beta \mathbf{I} \right) \right] \psi = c \left[(k_x^2 + k_y^2)\mathbf{I} + \mathbf{L}_{FD} \right] \psi.$$

The matrix $\overline{\mathbf{U}}_{FD}$ is diagonal with diagonal elements $\bar{u}(z_k)$. The matrix $\mathbf{L}_{FD}$ is tridiagonal with the form

$$\mathbf{L}_{FD} = \frac{1}{\Delta_z^2} \begin{bmatrix} S_1 & -S_1 & 0 & \cdots & & 0 \\ -S_1 & S_1 + S_2 & -S_2 & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ & & -S_{k-1} & S_{k-1} + S_k & -S_k & \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & & 0 & -S_{\mathcal{N}-1} & S_{\mathcal{N}-1} \end{bmatrix}$$

The matrix $\overline{\mathbf{Q}}_{y,FD}$ is also diagonal. If one defines a vector $\bar{\boldsymbol{u}}$ whose elements are $\bar{u}(z_k)$, the diagonal elements of $\overline{\mathbf{Q}}_{y,FD}$ are the elements of the vector $\mathbf{L}_{FD}\bar{\boldsymbol{u}}$. It is interesting to note that in the Galerkin method the approximate velocity profile $\bar{u}_{\mathcal{N}}(z)$ is derived from the potential vorticity gradient and the surface bouyancy gradients, while in the finite-difference approximation the potential vorticity gradient is derived from the velocity profile.

Appendix C

Bretherton's Formulation

This section formally demonstrates the equivalence of the original PV inversion problem eq. (1.2) with Bretherton's [4] reformulation eq. (2.6). The goal is to show that despite having imposed homogeneous Neumann boundary conditions on the streamfunction in Bretherton's formulation, the presence of Dirac delta distributions on the right hand side ensures that the actual solution satisfies the same inhomogeneous boundary conditions as the original problem. In the context of the discretized problem, the approximate solution will satisfy homogeneous Neumann boundary conditions on both boundaries for any $\mathcal{N}$. The discrete approximation is nevertheless still expected to converge *pointwise* to the true solution as $\mathcal{N} \to \infty$.

We begin by giving the Fourier transform of the Bretherton problem eq. (2.6)

$$-k^2\hat{\psi} + \partial_z\left(S(z)\partial_z\hat{\psi}\right) = \hat{q} - \frac{f_0\hat{b}^+}{N^2(z)}\delta(z-H) + \frac{f_0\hat{b}^-}{N^2(z)}\delta(z), \qquad \partial_z\hat{\psi} = 0 \text{ at } z = 0, H. \qquad \text{(C.1)}$$

Away from the boundaries the above equation is exactly the same as the original formulation, so the question of equivalence of the two formulations centers on the behavior of $\partial_z\hat{\psi}$ on the boundaries. The Green's function formulation of the solution is

$$\hat{\psi}(z;k) = \int_0^H g(z,s;k)\left(\hat{q}(s) - \frac{f_0\hat{b}^+}{N^2(s)}\delta(s-H) + \frac{f_0\hat{b}^-}{N^2(s)}\delta(s)\right)\mathrm{d}s$$

$$= \int_0^H g(z,s;k)\hat{q}(s)\mathrm{d}s - \frac{f_0\hat{b}^+}{N^2(H)}g(z,H;k) + \frac{f_0\hat{b}^-}{N^2(0)}g(z,0;k) \qquad \text{(C.2)}$$

where $g(z, s; k)$ the the Green's function. The derivative of $\hat{\psi}$ with respect to z is formally

$$\partial_z \hat{\psi}(z; k) = \int_0^H \partial_z g(z, s; k) \hat{q}(s) ds - \frac{f_0 \hat{b}^+}{N^2(H)} \partial_z g(z, H; k) + \frac{f_0 \hat{b}^-}{N^2(0)} \partial_z g(z, 0; k). \qquad \text{(C.3)}$$

The Green's function can be written in the form [18, Chapter 10]

$$g(z, s; k) = \frac{1}{S(s)(w_1(s)w_2'(s) - w_1'(s)w_2(s))} \begin{cases} w_1(z)w_2(s) & 0 \leq z \leq s \\ \\ w_1(s)w_2(z) & s \leq z \leq 1 \end{cases} \qquad \text{(C.4)}$$

where w_1 and w_2 are functions satisfying

$$-k^2 w_i(z) + \partial_z \left(S(z) \partial_z w_i(z) \right) = 0 \text{ for } i = 1, 2 \qquad \text{(C.5a)}$$

$$w_1'(0) = 0, \quad w_1'(H) = 1 \qquad \text{(C.5b)}$$

$$w_2'(0) = 1, \quad w_2'(H) = 0 \qquad \text{(C.5c)}$$

and the notation w_1' denotes the derivative of w_1. Consider $\partial_z g(z, s; k)$ at $z = 0$ for any $s \neq 0$

$$[\partial_z g(z, s; k)]_{z=0} = \frac{w_1'(0)w_2(s)}{S(s)(w_1(s)w_2'(s) - w_1'(s)w_2(s))} = 0. \qquad \text{(C.6)}$$

A similar manipulation shows that $\partial_z g(z, s; k) = 0$ at $z = H$ for any $s \neq H$. This shows that the contribution to $\partial_z \hat{\psi}$ from the integral in eq. (C.3) is zero on the boundaries $z = 0, H$. Now consider $\partial_z g(z, 0; k)$. At $s = 0$ the Green's function is

$$g(z, 0; k) = \frac{w_2(z)}{S(0)} \qquad \text{(C.7)}$$

and the derivative is

$$\partial_z g(z, 0; k) = \frac{w_2'(z)}{S(0)}. \qquad \text{(C.8)}$$

At $z = 0$ we have $\partial_z g(z, 0; k) = S(0)^{-1} = N^2(0)/f_0^2$, while at $z = H$ we have $\partial_z g(z, 0; k) = 0$. A similar argument shows that at $z = H$ $\partial_z g(z, H; k) = -S(H)^{-1} = -N^2(H)/f_0^2$ while at $z = 0$ $\partial_z g(z, H; k) = 0$. Plugging these expressions into eq. (C.3) and evaluating at $z = 0, H$ yields

$$\partial_z \hat{\psi}|_{z=0} = f_0 b^- \text{ and } \partial_z \hat{\psi}|_{z=H} = f_0 b^+. \tag{C.9}$$

(The foregoing argument is a more general version of the proof carried out in appendix A of [40] for constant $N^2(z)$.) This seems to contradict the homogeneous Neumann boundary conditions imposed at the beginning of this section. The paradox is explained by realizing that the presence of Dirac delta distributions on the right hand side of the reformulated problem implies that eq. (C.1) can only be understood in a weak sense, i.e. if you multiply eq. (C.1) by any sufficiently smooth test function and integrate from $z = 0$ to H the result should be true. The homogeneous boundary conditions that appear in eq. (C.1) are used in the construction of the Green's function, and guarantee that the solution will satisfy homogeneous Neumann boundary conditions whenever $b^\pm = 0$.