Generated using the official AMS LaTeX template v5.0 two-column layout. This work has been submitted for publication. Copyright in this work may be transferred without further notice, and this version may no longer be accessible.

arXiv:submit/3563733 [physics.ao-ph] 18 Jan 2021

On the Discrete Normal Modes of Quasigeostrophic Theory

Houssam Yassin*

Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

Stephen M. Griffies

NOAA/Geophysical Fluid Dynamics Laboratory and
Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

ABSTRACT

The baroclinic modes of quasigeostrophic theory are incomplete and the incompleteness manifests as a loss of information in the projection process. The incompleteness of the baroclinic modes is related to the presence of two previously unnoticed stationary step-wave solutions of the Rossby wave problem with flat boundaries. These step-waves are the limit of surface quasigeostrophic waves as boundary buoyancy gradients vanish. A complete normal mode basis for quasigeostrophic theory is obtained by considering the traditional Rossby wave problem with prescribed buoyancy gradients at the lower and upper boundaries. The presence of these boundary buoyancy gradients activates the previously inert boundary degrees of freedom. These Rossby waves have several novel properties such as the presence of multiple equivalent barotropic modes, a finite number of modes with negative norms, and their vertical structures form a basis capable of representing any quasigeostrophic state. Using this complete basis, we are able to obtain a series expansion to the potential vorticity of Bretherton (with Dirac delta contributions). We compare the convergence and differentiability properties of these complete modes with various other modes in the literature. We also examine the quasigeostrophic vertical velocity modes and derive a complete basis for such modes as well. In the process, we introduce the concept of the quasigeostrophic phase space which we define to be the space of all possible quasigeostrophic states. A natural application of these modes is the development of a weakly non-linear wave-interaction theory of geostrophic turbulence that takes prescribed boundary buoyancy gradients into account.

1. Introduction

a. Background

The vertical decomposition of quasigeostrophic motion into normal modes plays an important role in bounded stratified geophysical fluids (e.g., Charney 1971; Flierl 1978; Fu and Flierl 1980; Wunsch 1997; Chelton et al. 1998; Smith and Vallis 2001; Tulloch and Smith 2009; Lapeyre 2009; Ferrari et al. 2010; Ferrari and Wunsch 2010; de La Lama et al. 2016; LaCasce 2017; Brink and Pedlosky 2019). Most prevalent are the traditional baroclinic modes (e.g., section 6.5.2 in Vallis 2017) that are the vertical structures of Rossby waves in a quiescent ocean with no topography or boundary buouyancy gradients. In a landmark contribution, Wunsch (1997) partitions the ocean's kinetic energy into the baroclinic modes and finds that the zeroth and first baroclinic modes dominate over most of the extratropical ocean. Additionally, Wunsch (1997) concludes that the surface signal primarily reflects the first baroclinic mode and, therefore, the motion of the thermocline.

However, the use of baroclinic modes has come under increasing scrutiny in recent years (Lapeyre 2009; Roullet et al. 2012; Scott and Furnival 2012; Smith and Vanneste 2012). Lapeyre (2009) observes that the vertical shear of the baroclinic modes vanishes at the boundaries, thus leading to the concomitant vanishing of the boundary buoyancy. Consequently, Lapeyre (2009) proposes that the baroclinic modes cannot be complete[1] due to their inability to represent boundary buoyancy. To supplement the baroclinic modes, Lapeyre (2009) includes a boundary-trapped exponential surface quasigeostrophic solution (see Held et al. 1995) and suggests that the surface signal primarily reflects, not thermocline motion, but boundary-trapped surface quasigeostrophic dynamics (see also Lapeyre 2017).

Appending additional functions to the collections of normal modes as in Lapeyre (2009) or Scott and Furnival (2012) does not result in a set of normal modes since the appended functions are not orthogonal to the original modes. It is only with Smith and Vanneste (2012) that a set of normal modes capable of representing arbitrary surface buoyancy is derived.

Yet it is not clear how the normal modes of Smith and Vanneste (2012) differ from the baroclinic modes or what these modes correspond to in linear theory. Indeed, Rocha et al. (2015), noting that the baroclinic series expansion of any sufficiently smooth function converges uniformly to the function itself, argues that the incompleteness of the baro-

*Corresponding author: Houssam Yassin, hyassin@princeton.edu

[1] A collection of functions is said to be complete in some function space $\mathcal{F}$ if this collection forms a basis of $\mathcal{F}$. Specifying the underlying function space $\mathcal{F}$ turn out to be crucial, as we see in section 2d.

clinic modes has been "overstated". Moreover, de La Lama et al. (2016) and LaCasce (2017), motivated by the observation that the leading empirical orthogonal function of Wunsch (1997) vanishes near the ocean bottom, propose an alternate set of modes—the surface modes—that have a vanishing pressure at the bottom boundary.

We thus have a variety of proposed normal modes and it is not clear how their properties differ. Are the baroclinic modes actually incomplete? What about the surface modes? What does completeness mean in this context? What modes should be used to project model/observational data? The purpose of this paper is to answer these questions.

b. Normal modes and eigenfunctions

A normal mode is a linear motion in which all components of a system move coherently at a single frequency. Mathematically, a normal mode has the form

$$\Phi_a(x,y,z)\,\mathrm{e}^{-\mathrm{i}\omega_a t} \tag{1}$$

where Φ_a describes the spatial structure of the mode and ω_a is its angular frequency. The function Φ_a is obtained by solving a differential eigenvalue problem and hence is an eigenfunction. The collection of all eigenfunctions $\{\Phi_a\}_a$ forms a basis of some function space relevant to the problem.

By an abuse of terminology, the spatial structure Φ_a—alone—is often called a normal mode (e.g., the term "Fourier mode" is often used for $\mathrm{e}^{\mathrm{i}kx}$ where k is a wavenumber). In linear theory, this misnomer is often benign as each Φ_a corresponds to a frequency ω_a. For example, given some initial condition $\Psi(x,y,z)$, we decompose Ψ as a sum of modes at $t=0$,

$$\Psi(x,y,z) = \sum_a c_a\,\Phi_a(x,y,z), \tag{2}$$

where the c_a are the Fourier coefficients, and the time evolution is then given by

$$\sum_a c_a\,\Phi_a(x,y,z)\,\mathrm{e}^{-\mathrm{i}\omega_a t}. \tag{3}$$

However, in non-linear theory, this abuse of terminology can be confusing. Given some spatial structure $\Psi(x,y,z)$ in a non-linear fluid, we can exploit the basis properties of the eigenfunctions Φ_a to decompose Ψ as in equation (2). Whereas in a linear fluid only wave motion of the form (1) is possible, a non-linear fluid permits a larger collection of non-linear solutions (e.g., non-linear waves and vortices) and so the linear wave solution (3) no longer follows from the decomposition (2).

For this reason, we call the linear solution (1) a *physical* normal mode to distinguish it from the spatial structure Φ_a, which is only an eigenfunction. Otherwise, we will use the terms "normal mode" and "eigenfunction" interchangeably to refer to the spatial structure Φ_a, as is prevalent in the literature.

Our strategy here is then the following. We find the *physical* normal modes [of the form (1)] to various Rossby wave problems and examine the basis properties of their constituent eigenfunctions Φ_a. Our goal is to find a collection of eigenfunctions (i.e., "normal modes" in the prevalent terminology) capable of representing every possible quasigeostrophic state.

c. The contents of this article

This article constitutes an examination of all collections of discrete (i.e., non-continuum[2]) quasigeostrophic normal modes. We include the baroclinic modes, the surface modes of de La Lama et al. (2016) and LaCasce (2017), the surface-aware mode of Smith and Vanneste (2012), as well as various generalizations. To study the completeness of a set of normal modes, one must first define the underlying space in question. From general considerations, we introduce in section 2 the quasigeostrophic phase space, defined as the space of all possible quasigeostrophic states. Subsequently, in section 3 we use the general theory of differential eigenvalue problems with eigenvalue dependent boundary condition, developed in Yassin (in prep.), to study Rossby waves in an ocean with prescribed boundary buoyancy gradients (e.g., topography, see section 2a). Intriguingly, in an ocean with no topography, we find that, in addition to the usual baroclinic modes, there are two additional stationary step-mode solutions that have not been noted before. The stationary step-modes are the limits of boundary-trapped surface quasigeostrophic waves as the boundary buoyancy gradient vanishes. Our study of Rossby waves then leads us examine all possible discrete collections of normal modes in section 4.

As shown in section 4, the baroclinic modes are incomplete, as argued by Lapeyre (2009), but in a rather subtle manner. The baroclinic modes are incomplete in the sense that one *loses* information after projecting a function onto the baroclinic modes. In contrast, modes such as those suggested by Smith and Vanneste (2012) are complete in the quasigeostrophic phase space so that projecting a function onto such modes provides an *equivalent* representation of the function.

2. Mathematics of the quasigeostrophic phase space

a. The potential vorticity

Consider a three-dimensional region $\mathcal{D}$ of the form

$$\mathcal{D} = \mathcal{D}_0 \times [z_1, z_2]\,, \tag{4}$$

[2]Continuum modes appear once a sheared mean-flow is present, e.g., Drazin et al. (1982), Balmforth and Morrison (1994, 1995), and Brink and Pedlosky (2019).

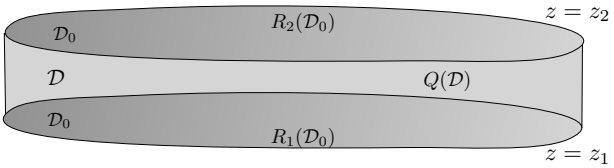

FIG. 1. Geometry of the quasigeostrophic problem as presented in section 2a. The domain is given by $\mathcal{D} = \mathcal{D}_0 \times [z_1, z_2]$ where $\mathcal{D}_0$ is the two-dimensional region representing the flat rigid lower and upper boundaries and z_1, z_2 are constants. We either have rigid side-walls or periodic boundary conditions. There is a volume potential vorticity density Q spread throughout the interior of $\mathcal{D}$ and there are surface potential vorticity densities R_1 and R_2 distributed on the lower and upper boundaries $\mathcal{D}_0$. In quasigeostrophic theory, the lower and upper boundaries are geometrically flat. Topography is included, not by modifying the geometry, but by prescribing a buoyancy gradient (or equivalently, a surface potential vorticity density gradient) at the appropriate boundary. In this figure, the vertical scale is enlarged for clarity; quasigeostrophic theory only applies when the aspect ratio H/L is small, where $H = z_2 - z_1$ and L is the horizontal scale of the domain.

as illustrated in figure 1; $\mathcal{D}_0$ is the rigid two-dimensional lower and upper boundaries and z_1, z_2 are constants. The horizontal boundaries are either rigid or periodic.

The state of a quasigeostrophic fluid in $\mathcal{D}$ is determined by a charge-like quantity known as the quasigeostrophic potential vorticity (Hoskins et al. 1985; Schneider et al. 2003). If the potential vorticity is distributed throughout the three-dimensional region $\mathcal{D}$, we are concerned with the volume potential vorticity density, Q, with Q related to the geostrophic streamfunction ψ by [e.g., section 5.4 of Vallis (2017)]

$$Q = f + \nabla^2\psi + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right). \tag{5}$$

Here, the latitude dependent Coriolis parameter is

$$f = f_0 + \beta y, \tag{6}$$

$N(z)$ is the prescribed background buoyancy frequency, ∇^2 is the horizontal Laplacian operator, and

$$\boldsymbol{u} = \hat{\boldsymbol{z}} \times \boldsymbol{\nabla}\psi \tag{7}$$

is the horizontal geostrophic velocity, $\boldsymbol{u} = (u, v)$. Additionally, the potential vorticity may be distributed over a two-dimensional region, say the lower and upper boundaries $\mathcal{D}_0$, to obtain surface potential vorticity densities R_1 and R_2. The surface potential vorticity densities are related to the streamfunction by

$$R_j = (-1)^{j+1}\left[g_j + \left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right)\bigg|_{z=z_j}\right] \tag{8}$$

where g_j is an imposed surface potential vorticity density at the lower or upper boundary and $j = 1, 2$. Furthermore, g_j corresponds to a prescribed buoyancy

$$b_j = \frac{N^2}{f_0} g_j \tag{9}$$

at the jth boundary [see equation (A6)].

The gradient of the surface potential vorticity density, R_j, is proportional to the buoyancy gradient at the jth boundary. A non-zero $\boldsymbol{\nabla} R_j$ signals outcrops or incrops of buoyancy on $\mathcal{D}_0$. In quasigeostrophy, boundary buoyancy can be either prescribed, as in the first term of equation (8), or induced by a geostrophic flow, as in the second term of equation (8). An example prescribed buoyancy gradient arises in the following manner. Namely, quasigeostrophy formally has zero topography so that the lower and upper boundaries are rigid and flat. However, topography can be included by proxy through prescribing a surface potential vorticity

$$g_j = f_0 h_j \tag{10}$$

where h_j represents infinitesimal topography at the jth boundary.

We can interpret a surface potential vorticity density R_j as an infinitesimally thin volume potential vorticity density (Bretherton 1966). This interpretation motivates the definition of the potential vorticity density *distribution*

$$\mathfrak{Q} = Q + \sum_{j=1}^{2} R_j\, \delta(z - z_j) \tag{11}$$

where $\delta(z)$ is the Dirac delta distribution with the dimension of inverse length. The potential vorticity in the volume $\mathcal{D}$ is then given by the volume integral

$$\int_{\mathcal{D}} \mathfrak{Q}\, \mathrm{d}V. \tag{12}$$

Knowledge of Q, R_1 and R_2 determines the streamfunction ψ through equations (5) and (8). After determining the geostrophic horizontal velocity by $\boldsymbol{u} = \hat{\boldsymbol{z}} \times \boldsymbol{\nabla}\psi$, the quasigeostrophic state is evolved using

$$\frac{\partial Q}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} Q = 0 \quad \text{for } z \in (z_1, z_2) \tag{13a}$$

$$\frac{\partial R_j}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} R_j = 0 \quad \text{for } z = z_j. \tag{13b}$$

Thus, the dynamical system is entirely determined by the material conservation of the volume potential vorticity density, Q, in the interior and material conservation of the surface potential vorticity densities, R_1 and R_2, at the lower and upper boundaries.

b. Defining the quasigeostrophic phase space

We define the quasigeostrophic phase space to be the space of all possible quasigeostrophic states, with a quasigeostrophic state determined by the potential vorticity densities, Q, R_1, and R_2. Equivalently, a quasigeostrophic

state is determined by a distribution $\mathfrak{Q}$ of the form given in equation (11). Note that the volume potential vorticity density, Q, is defined throughout the whole fluid region $\mathcal{D}$, so that $Q = Q(x, y, z, t)$. In contrast, the surface potential vorticity densities, R_1 and R_2, are only defined on the two-dimensional lower and upper boundary surfaces, $\mathcal{D}_0$, so that $R_j = R_j(x, y, t)$.

It will be useful to restate the previous paragraph with some mathematical precision. For that purpose, let $L^2[\mathcal{D}]$ be the space of square-integrable functions[3] in the fluid volume $\mathcal{D}$, and let $L^2[\mathcal{D}_0]$ be the space of square-integrable functions on the boundary area $\mathcal{D}_0$. Elements of $L^2[\mathcal{D}]$ are functions of three spatial coordinates whereas elements of $L^2[\mathcal{D}_0]$ are functions of two spatial coordinates. Hence, $Q \in L^2[\mathcal{D}]$ and $R_1, R_2 \in L^2[\mathcal{D}_0]$.

Define the space $\mathscr{P}$ by

$$\mathscr{P} = L^2[\mathcal{D}] \oplus L^2[\mathcal{D}_0] \oplus L^2[\mathcal{D}_0], \tag{14}$$

where $\oplus$ is the direct sum. Equation (14) states that any element of $\mathscr{P}$ is the sum of a function defined on the volume $\mathcal{D}$ and an element of $L^2[\mathcal{D}]$, plus two functions defined on the area $\mathcal{D}_0$ that are elements of $L^2[\mathcal{D}_0]$. By construction, the potential vorticity distribution $\mathfrak{Q}$ is determined by a function Q on the volume $\mathcal{D}$ and two functions R_1 and R_2 on the area $\mathcal{D}_0$. We conclude that $\mathfrak{Q} \in \mathscr{P}$ and that $\mathscr{P}$ is the space of all possible quasigeostrophic states. We thus call $\mathscr{P}$ the quasigeostrophic phase space.

c. The phase space in terms of the streamfunction

The geostrophic streamfunction, ψ, contains the same dynamical information as $\mathfrak{Q}$, and as such it offers an equivalent description of the dynamics. We can thus choose to represent the phase space, $\mathscr{P}$, either in terms of $\mathfrak{Q}$ or in terms of ψ. Consequently, the space of all possible geostrophic streamfunctions is identical[4] to the space of all possible potential vorticity distributions $\mathfrak{Q}$. That is, both spaces are $\mathscr{P}$. Just as $Q \in L^2[\mathcal{D}]$ is independent of $R_i \in L^2[\mathcal{D}_0]$, then $\psi(x, y, z, t)$, for $z \in (z_1, z_2)$, is independent of $\psi|_{z=z_1}$ and $\psi|_{z=z_2}$.

One may have expected the streamfunction, ψ, as a function defined on the volume $\mathcal{D}$, to belong to $L^2[\mathcal{D}]$. But ψ instead belongs to the larger space, $\mathscr{P}$, given in equation (14). To make this point explicit, write ψ as the sum

$$\psi = \psi_{\text{int}} + \psi_{\text{low}} + \psi_{\text{upp}}. \tag{15}$$

ψ_{int} is the flow generated by the volume potential vorticity density, Q, in the fluid interior

$$\begin{aligned} Q - f = \nabla^2 \psi_{\text{int}} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\frac{\partial \psi_{\text{int}}}{\partial z}\right) \quad &\text{for } z \in (z_1, z_2) \\ 0 = \frac{f_0^2}{N^2}\frac{\partial \psi_{\text{int}}}{\partial z} \quad &\text{for } z = z_1, z_2, \end{aligned} \tag{16}$$

whereas ψ_{low} is the flow generated by the surface potential vorticity density, R_1, on the lower boundary

$$\begin{aligned} 0 = \nabla^2 \psi_{\text{low}} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi_{\text{low}}}{\partial z}\right) \quad &\text{for } z \in (z_1, z_2) \\ R_1 - g_1 = \frac{f_0^2}{N^2}\frac{\partial \psi_{\text{low}}}{\partial z} \quad &\text{for } z = z_1 \\ 0 = \frac{f_0^2}{N^2}\frac{\partial \psi_{\text{low}}}{\partial z} \quad &\text{for } z = z_2, \end{aligned} \tag{17}$$

and ψ_{upp} is the flow generated by the surface potential vorticity density, R_2, on the upper boundary

$$\begin{aligned} 0 = \nabla^2 \psi_{\text{upp}} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi_{\text{upp}}}{\partial z}\right) \quad &\text{for } z \in (z_1, z_2) \\ 0 = \frac{f_0^2}{N^2}\frac{\partial \psi_{\text{upp}}}{\partial z} \quad &\text{for } z = z_1 \\ R_2 + g_2 = -\frac{f_0^2}{N^2}\frac{\partial \psi_{\text{upp}}}{\partial z} \quad &\text{for } z = z_2. \end{aligned} \tag{18}$$

We observe that ψ consists of three independent components[5]: ψ_{int} belongs to $L^2[\mathcal{D}]$, $\psi_{\text{low}}|_{z=z_1}$ belongs to $L^2[\mathcal{D}_0]$, and $\psi_{\text{upp}}|_{z=z_2}$ belongs to $L^2[\mathcal{D}_0]$. It follows that ψ is an element of the phase space $\mathscr{P}$.

Equations (16)–(18) motivate the definition of the relative potential vorticity densities, $q = Q - f$ and $r_j = R_j - (-1)^{j+1} g_j$, which are the portions of the potential vorticity providing a source for a streamfunction. Explicitly, the relative potential vorticity densities are

$$q = \nabla^2 \psi + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right) \quad \text{for } z \in (z_1, z_2) \tag{19a}$$

$$r_1 = \frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z} \quad \text{for } z = z_1 \tag{19b}$$

$$r_2 = -\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z} \quad \text{for } z = z_2. \tag{19c}$$

We then define a relative potential vorticity density distribution, $\mathfrak{q}$, analogous to $\mathfrak{Q}$ in equation (11)

$$\mathfrak{q} = q + \sum_{j=1}^{2} r_j \, \delta(z - z_j). \tag{20}$$

[3]The definition of $L^2[\mathcal{D}]$ is more subtle than presented here. Namely, elements of $L^2[\mathcal{D}]$ are not functions, but rather equivalence classes of functions leading to the unintuitive properties seen in this section. See Yassin (in prep.) and citations within for more details.

[4]Streamfunctions in the phase space are defined up to a gauge transformation (see Schneider et al. 2003). We consider two streamfunctions that give rise to the same potential vorticity distribution as equivalent in the phase space $\mathscr{P}$.

[5]We emphasize that $\psi_{\text{low}} = \psi_{\text{low}}(x, y, z, t)$ and $\psi_{\text{upp}} = \psi_{\text{upp}}(x, y, z, t)$, that is, both are functions of *three* spatial coordinates. However, both ψ_{low} and ψ_{upp} are determined by the distribution of surface potential vorticity $R_j - (-1)^{j+1} g_j$ on the jth boundary.

d. The vertical structure phase space

Since the fluid region, $\mathcal{D}$, is separable, we can expand the potential vorticity density distribution $\mathfrak{q}$ and the streamfunction ψ in terms of the eigenfunctions, $e_{\boldsymbol{k}}$, of the horizontal Laplacian. For a horizontal domain $\mathcal{D}_0$, the eigenfunction $e_{\boldsymbol{k}}(\boldsymbol{x})$ satisfies

$$-\nabla^2 e_{\boldsymbol{k}} = k^2\, e_{\boldsymbol{k}}. \tag{21}$$

where $\boldsymbol{x} = (x, y)$ is the horizontal position vector, $\boldsymbol{k} = (k_x, k_y)$ is the horizontal wavevector, and $k = |\boldsymbol{k}|$ is the horizontal wavenumber. For example, in a horizontally periodic domain the eigenfunctions $e_{\boldsymbol{k}}(\boldsymbol{x})$ are proportional to complex exponentials, $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$. The horizontal eigenfunctions $e_{\boldsymbol{k}}$ obey the orthogonality relation

$$\delta_{\boldsymbol{a},\boldsymbol{b}} = \frac{1}{A}\int_{\mathcal{D}_0} e^*_{\boldsymbol{a}}(\boldsymbol{x})\, e_{\boldsymbol{b}}(\boldsymbol{x})\,\mathrm{d}A, \tag{22}$$

where $\delta_{\boldsymbol{a},\boldsymbol{b}}$ is the Kronecker delta, A is the area of the region $\mathcal{D}_0$, and $e^*_{\boldsymbol{a}}$ is the complex conjugate of $e_{\boldsymbol{a}}$.

Projecting the relative potential vorticity density distribution, $\mathfrak{q}$, onto the horizontal eigenfunctions, $e_{\boldsymbol{k}}$, yields

$$\mathfrak{q}(\boldsymbol{x}, z, t) = \sum_{\boldsymbol{k}} \mathfrak{q}_{\boldsymbol{k}}(z, t)\, e_{\boldsymbol{k}}(\boldsymbol{x}). \tag{23}$$

Only the vertical dependence is retained in the vertical structure, $\mathfrak{q}_{\boldsymbol{k}}$, of the potential vorticity density distribution

$$\mathfrak{q}_{\boldsymbol{k}}(z, t) = q_{\boldsymbol{k}}(z, t) + \sum_{j=1}^{2} r_{j\boldsymbol{k}}(t)\,\delta(z - z_j) \tag{24}$$

where $q_{\boldsymbol{k}}$ is a function of z and $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$ are independent of z. Hence, $q_{\boldsymbol{k}}$ is an element of $L^2[(z_1, z_2)]$ whereas $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$ are elements of the space of complex numbers[6] , $\mathbb{C}$.

We conclude that the vertical structure of the potential vorticity, $\mathfrak{q}_{\boldsymbol{k}}$, is an element of

$$\widehat{\mathscr{P}} = L^2[(z_1, z_2)] \oplus \mathbb{C} \oplus \mathbb{C}, \tag{25}$$

so that the vertical structure, $\mathfrak{q}_{\boldsymbol{k}}$, of the potential vorticity distribution is determined by a function, $q_{\boldsymbol{k}}$, in $L^2[(z_1, z_2)]$ and two z-independent elements, $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$, of $\mathbb{C}$. Similarly, the streamfunction can be represented as

$$\psi(\boldsymbol{x}, z, t) = \sum_{\boldsymbol{k}} \psi_{\boldsymbol{k}}(z, t)\, e_{\boldsymbol{k}}(\boldsymbol{x}), \tag{26}$$

where $\psi_{\boldsymbol{k}}$ and $\mathfrak{q}_{\boldsymbol{k}}$ are related by

$$q_{\boldsymbol{k}} = -k^2\psi_{\boldsymbol{k}} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi_{\boldsymbol{k}}}{\partial z}\right) \tag{27a}$$

$$r_{j\boldsymbol{k}} = (-1)^{j+1}\left(\frac{f_0^2}{N^2}\frac{\partial \psi_{\boldsymbol{k}}}{\partial z}\right)\bigg|_{z=z_j}. \tag{27b}$$

As before, knowledge of the vertical structure of the streamfunction, $\psi_{\boldsymbol{k}}(z)$, is equivalent to knowing the vertical structure of the potential vorticity distribution, $\mathfrak{q}_{\boldsymbol{k}}(z)$. Thus, $\psi_{\boldsymbol{k}}$ must also belong to $\widehat{\mathscr{P}}$. As a consequence of the surface potential vorticity, $r_{j\boldsymbol{k}}$, on the boundaries, the properties of $\psi_{\boldsymbol{k}}(z_j)$ on the lower and upper boundaries are independent of the interior values of the streamfunction.

That $\psi_{\boldsymbol{k}}$ belongs to $\widehat{\mathscr{P}}$ and not $L^2[(z_1, z_2)]$ underlies much of the confusion over baroclinic modes. Assertions of completeness, based on Sturm-Liouville theory, assume that ψ is an element of $L^2[(z_1, z_2)]$. However, as we have shown, that is an incorrect assumption. A collection of functions $\{\phi_n(z)\}_{n=0}^{\infty}$ is complete (in quasigeostrophic theory) only if this collection forms a basis of the function space $\widehat{\mathscr{P}}$. In the context of quasigeostrophic theory, the space $\widehat{\mathscr{P}}$ first appeared in Smith and Vanneste (2012). More generally, $\widehat{\mathscr{P}}$ appears in the presence of non-trivial boundary dynamics (Yassin in prep.).

We call $\widehat{\mathscr{P}}$ the vertical structure phase space, and for convenience we denote $L^2[(z_1, z_2)]$ by L^2 for the remainder of the article. The vertical structure phase space $\widehat{\mathscr{P}}$ is then written as the direct sum

$$\widehat{\mathscr{P}} = L^2 \oplus \mathbb{C}^2. \tag{28}$$

e. Representing the energy and potential enstrophy

We find it convenient to represent several quadratic quantities in terms of the eigenfunctions of the horizontal Laplacian, $e_{\boldsymbol{k}}(\boldsymbol{x})$. The energy per unit mass in the volume $\mathcal{D}$ is given by

$$E = \frac{1}{V}\int_{\mathcal{D}}\left[|\boldsymbol{\nabla}\psi|^2 + \frac{f_0^2}{N^2}\left|\frac{\partial\psi}{\partial z}\right|^2\right]\mathrm{d}A\,\mathrm{d}z = \sum_{\boldsymbol{k}} E_{\boldsymbol{k}}, \tag{29}$$

where the horizontal energy mode is given by the vertical integral

$$E_{\boldsymbol{k}} = \frac{1}{H}\int_{z_1}^{z_2}\left[k^2\,|\psi_{\boldsymbol{k}}|^2 + \frac{f_0^2}{N^2}\left|\frac{\partial\psi_{\boldsymbol{k}}}{\partial z}\right|^2\right]\mathrm{d}z, \tag{30}$$

with $V = A\,H$ the domain volume and $H = z_2 - z_1$ the domain depth.

Similarly, for the relative volume potential enstrophy density, Z, we have

$$Z = \frac{1}{V}\int_{\mathcal{D}} |q|^2\mathrm{d}A\,\mathrm{d}z = \sum_{\boldsymbol{k}} Z_{\boldsymbol{k}} \tag{31}$$

[6]Since all physical fields must be real, only a single degree of freedom is gained from $\mathbb{C}$. Furthermore, when complex notation is used (e.g., complex exponentials for the horizontal eigenfunctions $e_{\boldsymbol{k}}$) it is only the real part of the fields that is physical.

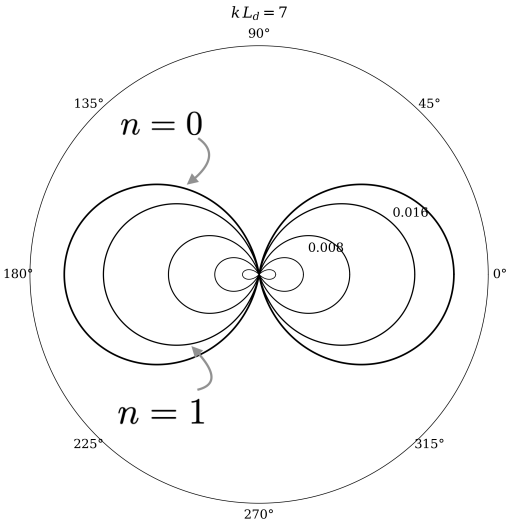

FIG. 2. Polar plots of the absolute value of the non-dimensional angular frequency $|\omega_n|/(\beta L_d)$ of the first five modes of the traditional eigenvalue problem (section 3a) as a function of the wave propagation direction, $\boldsymbol{k}/|\boldsymbol{k}|$, for constant stratification. The outer most ellipse, with the largest absolute angular frequency, represents the angular frequency of the barotropic ($n = 0$) mode. The higher modes have smaller absolute frequencies and are thus concentric and within the barotropic angular frequency curve. Since the absolute value of the angular frequency of the barotropic mode becomes infinitely large at small horizontal wavenumbers k, we have chosen a large wavenumber k, given by $k\,L_d = 7$, so that the angular frequency of the first five modes can be plotted in the same figure. We have chosen $f_0 = 10^{-4}\ \mathrm{s}^{-1}, \beta = 10^{-11}\ \mathrm{m}^{-1}\,\mathrm{s}^{-1}, N_0 = 10^{-2}\ \mathrm{s}^{-1}$ and $H = 1$ km leading to a deformation radius $L_d = N_0\,H/f_0 = 100$ km. Numerical solutions to all eigenvalue problems in this article are obtained using `Dedalus` (Burns et al. 2020).

where

$$Z_{\boldsymbol{k}} = \frac{1}{H}\int_{z_1}^{z_2} |q_{\boldsymbol{k}}|^2\,\mathrm{d}z. \tag{32}$$

Finally, analogous to Z, we have the relative surface potential enstrophy densities, Y_j, on the area $\mathcal{D}_0$

$$Y_j = \frac{1}{A}\int_{\mathcal{D}_0} |r_j|^2\,\mathrm{d}A = \sum_{\boldsymbol{k}} Y_{j\boldsymbol{k}} \tag{33}$$

where

$$Y_{j\boldsymbol{k}} = |r_{j\boldsymbol{k}}|^2. \tag{34}$$

3. Rossby waves in a linear ocean

In this section, we study Rossby waves in an otherwise quiescent ocean; in other words, we examine the *physical* normal modes of a quiescent ocean. Linearizing the equations of motion (13) about a resting ocean yields

$$\frac{\partial q}{\partial t} + \beta v = 0 \quad \text{for } z \in (z_1, z_2) \tag{35a}$$

$$\frac{\partial r_j}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\left[(-1)^{j+1} g_j\right] = 0 \quad \text{for } z = z_j. \tag{35b}$$

We assume that the prescribed surface potential vorticity densities at the lower and upper boundaries, g_1 and g_2, are linear, which ensures the resulting eigenvalue problem is separable.

The linear quasigeostrophic system (35) with no background mean-flow is artificial since the ocean has a circulation while the above equations do not. A consequence is that the linear system (35) has infinitely many Rossby wave modes and no advective continuum modes. However, once a mean-flow is present, then it is possible that only a finite number of Rossby waves exist (Killworth and Anderson 1977) and a large number of degrees of freedom will generally be associated with advective continuum modes which the above model lacks (the situation is analogous to Drazin et al. 1982; Balmforth and Morrison 1994, 1995).

The importance of the linear problem (35) is that it provides all possible discrete Rossby wave normal modes in a quasigeostrophic fluid. Substituting a wave ansatz of the form [compare with equation (1) for *physical* normal modes]

$$\psi(\boldsymbol{x}, z, t) = \hat{\psi}(z)\, e_{\boldsymbol{k}}(\boldsymbol{x})\, \mathrm{e}^{-\mathrm{i}\omega t} \tag{36}$$

into the linear problem (35) renders

$$(-\mathrm{i}\,\omega)\left[-k^2\hat{\psi} + \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\hat{\psi}}{\mathrm{d}z}\right)\right] + \mathrm{i}\,k_x\beta\hat{\psi} = 0 \tag{37}$$

for $z \in (z_1, z_2)$, and

$$(-\mathrm{i}\,\omega)\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\hat{\psi}}{\mathrm{d}z}\right) + \mathrm{i}\,\hat{\boldsymbol{z}}\cdot\left(\boldsymbol{k}\times\boldsymbol{\nabla}g_j\right)\hat{\psi} = 0 \tag{38}$$

for $z = z_1, z_2$.

a. Traditional Rossby wave problem

We first study the traditional case of linear fluctuations to a quiescent ocean with isentropic lower and upper boundaries. Setting $\boldsymbol{\nabla}g_1 = \boldsymbol{\nabla}g_2 = 0$ in the eigenvalue problem (37)–(38) gives

$$\omega\left[-k^2\phi + \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\phi}{\mathrm{d}z}\right)\right] - \beta\,k_x\,\phi = 0 \tag{39a}$$

$$\omega\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\phi}{\mathrm{d}z}\right)\bigg|_{z=z_j} = 0, \tag{39b}$$

where $\hat{\psi}(z) = \hat{\psi}_0\,\phi(z)$ and ϕ is a non-dimensional function. There are two cases to consider depending on whether ω vanishes.

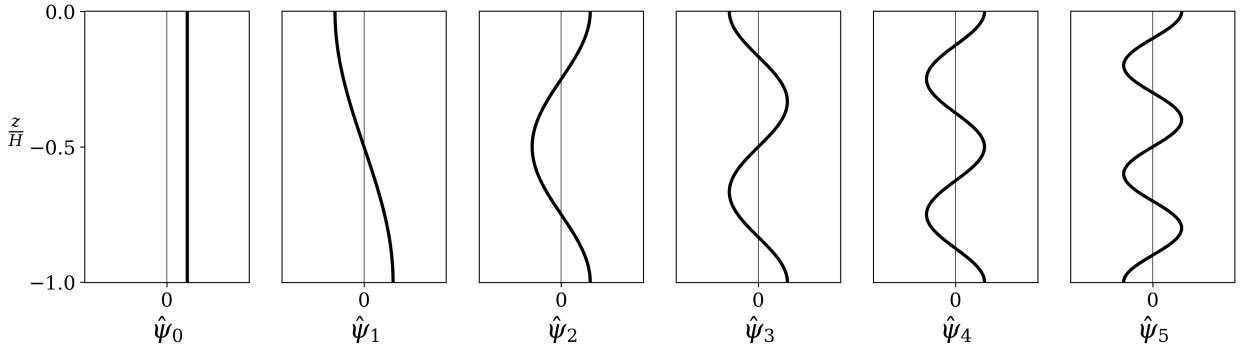

FIG. 3. The first six baroclinic modes in constant stratification for a propagation direction $\boldsymbol{k}/|\boldsymbol{k}|$ with $k_x \neq 0$. Other parameters are as in figure 2. See section 3a for details.

1) Traditional baroclinic modes

Assuming $\omega \neq 0$ in the eigenvalue problem (39) renders a Sturm-Liouville eigenvalue problem in L^2

$$-\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\phi}{\mathrm{d}z}\right) = \lambda\,\phi \quad \text{for } z \in (z_1, z_2) \tag{40a}$$

$$\frac{f_0^2}{N^2}\frac{\mathrm{d}\phi}{\mathrm{d}z} = 0 \quad \text{for } z = z_1, z_2, \tag{40b}$$

where the eigenvalue λ is the square vertical wavenumber, given by

$$\omega = -\frac{\beta\, k_x}{k^2 + \lambda}. \tag{41}$$

See figure 2 for an illustration of the dependence of $|\omega|$ on the wavevector $\boldsymbol{k}$.

From Sturm-Liouville theory, the eigenvalue problem (40) has infinitely many eigenfunctions $\phi_0, \phi_1, \phi_2, \ldots$ with distinct and ordered eigenvalues, λ_n, satisfying

$$0 = \lambda_0 < \lambda_1 < \cdots \to \infty. \tag{42}$$

The nth mode, ϕ_n, has n internal zeros in the interval (z_1, z_2) (see figure 3). The eigenfunctions are orthonormal with respect to the inner product, $\langle\cdot,\cdot\rangle_0$, given by

$$\langle \varphi, \phi \rangle_0 = \frac{1}{H}\int_{z_1}^{z_2} \varphi\,\phi\,\mathrm{d}z, \tag{43}$$

with orthonormality meaning that

$$\delta_{mn} = \langle \phi_m, \phi_n \rangle_0 \tag{44}$$

where δ_{mn} is the Kronecker delta. A powerful and commonly used result of Sturm-Liouville theory is that the set $\{\phi_n\}_{n=0}^{\infty}$ forms an orthonormal basis of L^2.

2) Stationary step-modes

There are two additional solutions to the Rossby wave eigenvalue problem (39) not previously noted in the literature. If $\omega = 0$ then the eigenvalue problem (39) becomes

$$\beta\, k_x\, \phi = 0 \quad \text{for } z \in (z_1, z_2) \tag{45a}$$

$$0 = 0 \quad \text{for } z = z_1, z_2. \tag{45b}$$

Consequently, if $k_x \neq 0$, then $\phi(z) = 0$ for $z \in (z_1, z_2)$. That is, ϕ must vanish in the interior of the interval. However, since $\omega = 0$ in (39b), we obtain tautological boundary conditions (45b). As a result, $\hat{\psi}$ can take arbitrary values at the lower and upper boundaries. Thus two solutions are

$$\phi_j^{\mathrm{step}}(z) = \begin{cases} 1 & \text{for } z = z_j \\ 0 & \text{otherwise.} \end{cases} \tag{46}$$

The two step-mode solutions (46) are independent of the traditional baroclinic modes. An expansion of the step-mode ϕ_j^{step} in terms of the baroclinic modes will fail and produce a series that is identically zero.

The two stationary step-modes, ϕ_1^{step} and ϕ_2^{step}, correspond to the two inert degrees of freedom in eigenvalue problem (39). These two solutions are neglected in the traditional eigenvalue problem (40) through the assumption that $\omega \neq 0$. We will see that these two step-waves are obtained as limits of boundary-trapped modes as the boundary buoyancy gradients $N^2\boldsymbol{\nabla} g_j/f_0$ become small.

3) The general solution

For a wavevector $\boldsymbol{k}$ with $k_x \neq 0$, the vertical structure of the streamfunction must be of the form

$$\psi_{\boldsymbol{k}}(z, t=0) = \Psi(z) + \sum_{j=1}^{2} \Psi_j\, \phi_j^{\mathrm{step}}(z) \tag{47}$$

where $\Psi(z)$ is a twice differentiable function satisfying $\mathrm{d}\Psi(z_j)/\mathrm{d}z = 0$ for $j = 1,2$ and Ψ_1, Ψ_2 are arbitrary constants. We can represent Ψ according to the expansion

$$\Psi = \sum_{n=0}^{\infty} \langle \Psi, \phi_n \rangle_0 \, \phi_n \tag{48}$$

and so the time-evolution is

$$\psi_{\boldsymbol{k}}(z,t) = \sum_{n=0}^{\infty} \langle \Psi, \phi_n \rangle_0 \, \phi_n \, \mathrm{e}^{-\mathrm{i}\omega_n t} + \sum_{j=1}^{2} \Psi_j \, \phi_j^{\text{step}}. \tag{49}$$

It is this time-evolution expression, which is valid only in linear theory for a quiescent ocean, that gives the baroclinic modes a clear physical meaning. More precisely, equation (49) states that the vertical structure $\Psi(z)$ disperses into its constituent Rossby waves with vertical structures ϕ_n. Outside the linear theory of this section, baroclinic modes do not have a physical interpretation, although they remain a mathematical basis for L^2.

b. The generalized Rhines problem

We now study the case with constant buoyancy gradients $\boldsymbol{\nabla} g_1$ and $\boldsymbol{\nabla} g_2$ at the lower and upper boundaries. The special case of a "flat" upper boundary (i.e., isentropic upper boundary $\boldsymbol{\nabla} g_2 = 0$), a meridional "bottom slope", and constant stratification was first investigated by Rhines (1970). Subsequently, Charney and Flierl (1981) extended the analysis to realistic stratification and Straub (1994) examined the dependence of the waves on the propagation direction. In Yassin (in prep.), the author applies the mathematical theory of eigenvalue problems with λ-dependent boundary conditions and obtains various completeness and expansions results as well as a qualitative theory for the modes. Below, we generalize these results, study the two limiting boundary conditions, and examine the phase speed of the resulting waves.

1) The eigenvalue problem

Let $\hat{\psi}(z) = \hat{\psi}_0 \, \varphi(z)$ where φ is a non-dimensional function. We then manipulate the eigenvalue problem (37)–(38) to obtain

$$-\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\varphi}{\mathrm{d}z}\right) = \lambda\,\varphi \quad \text{for } z \in (z_1, z_2) \tag{50a}$$

$$-k^2\varphi + (-1)^j \gamma_j^{-1}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\varphi}{\mathrm{d}z}\right) = \lambda\,\varphi \quad \text{for } z = z_j, \tag{50b}$$

where the length-scale γ_j is given by

$$\begin{aligned}\gamma_j &= (-1)^{j+1}\frac{\hat{\boldsymbol{z}}\cdot(\boldsymbol{k}\times\boldsymbol{\nabla} g_j)}{\hat{\boldsymbol{z}}\cdot(\boldsymbol{k}\times\boldsymbol{\nabla} f)} \\ &= (-1)^{j+1}\left(\frac{\alpha_j\, k}{\beta\, k_x}\right)\sin(\Delta\theta_j)\end{aligned} \tag{51}$$

where $\alpha_j = |\boldsymbol{\nabla} g_j|$ and $\Delta\theta_j$ is the angle between the wavevector $\boldsymbol{k}$ and $\boldsymbol{\nabla} g_j$ measured counterclockwise from $\boldsymbol{k}$. The parameter γ_j depends only on the direction of the wavevector $\boldsymbol{k}$ and not its magnitude $k = |\boldsymbol{k}|$. If $\gamma_j = 0$, then the jth boundary condition can be written as a λ-independent boundary condition. For now, we consider the case $\gamma_j \neq 0$.

Since the eigenvalue λ appears in both boundary conditions (50b), the eigenvalue problem (50) occurs in the function space $L^2 \oplus \mathbb{C}^2$ and hence is not a Sturm-Liouville problem. Instead, we must use the mathematical theory developed in Yassin (in prep.) for eigenvalue problems with λ-dependent boundary conditions [7].

2) Characterizing the eigen-solutions

The eigenvalue problem (50) has a countable infinity of eigenfunctions $\varphi_0, \varphi_1, \varphi_2, \ldots$ with ordered and distinct non-zero eigenvalues λ_n satisfying

$$\lambda_0 < \lambda_1 < \lambda_2 < \cdots \to \infty. \tag{52}$$

The inner product $\langle \cdot, \cdot \rangle$ induced by the eigenvalue problem (50) is

$$\langle \varphi, \phi \rangle = \frac{1}{H}\left(\int_{z_1}^{z_2} \varphi\,\phi \,\mathrm{d}z + \sum_{j=1}^{2} \gamma_j \, \varphi(z_j)\,\phi(z_j)\right), \tag{53}$$

which depends on the direction of the horizontal wavevector $\boldsymbol{k}$ through γ_1 and γ_2. Furthermore, γ_1 and γ_2 are not necessarily positive, with one consequence being that some functions φ may have a negative square $\langle \varphi, \varphi \rangle < 0$. Orthonormality of the modes φ_n then takes the form

$$\pm\delta_{mn} = \langle \varphi_m, \varphi_n \rangle, \tag{54}$$

where at most two modes φ_n satisfy $\langle \varphi_n, \varphi_n \rangle = -1$.

The theory in Yassin (in prep.) provides the following inequality

$$\left(k^2 + \lambda_n\right)\langle \varphi_n, \varphi_n \rangle > 0 \tag{55}$$

which, using the dispersion relation (41), implies that modes φ_n with $\langle \varphi_n, \varphi_n \rangle > 0$ correspond to waves with a westward phase speed while modes φ_n with $\langle \varphi_n, \varphi_n \rangle < 0$ correspond to waves with an eastward phase speed (assuming $\beta > 0$).

We distinguish the following cases depending on the signs of γ_1 and γ_2 and as depicted in figures 4, 5, and 6. In the following, we assume $k \neq 0$. [These results are a consequence of the mathematical theory in Yassin (in prep.)]

i. $\gamma_1 > 0$ and $\gamma_2 > 0$. All eigenvalues satisfy $\lambda_n > -k^2$, all modes satisfy $\langle \varphi_n, \varphi_n \rangle > 0$, and all waves propagate westward. The nth mode, φ_n, has n internal

[7] To apply the theory in Yassin (in prep.), let $\tilde{\lambda} = \lambda - k^2$ be the eigenvalue in place of λ; the resulting eigenvalue problem for $\tilde{\lambda}$ will then satisfy the left-definiteness conditions in Yassin (in prep.).

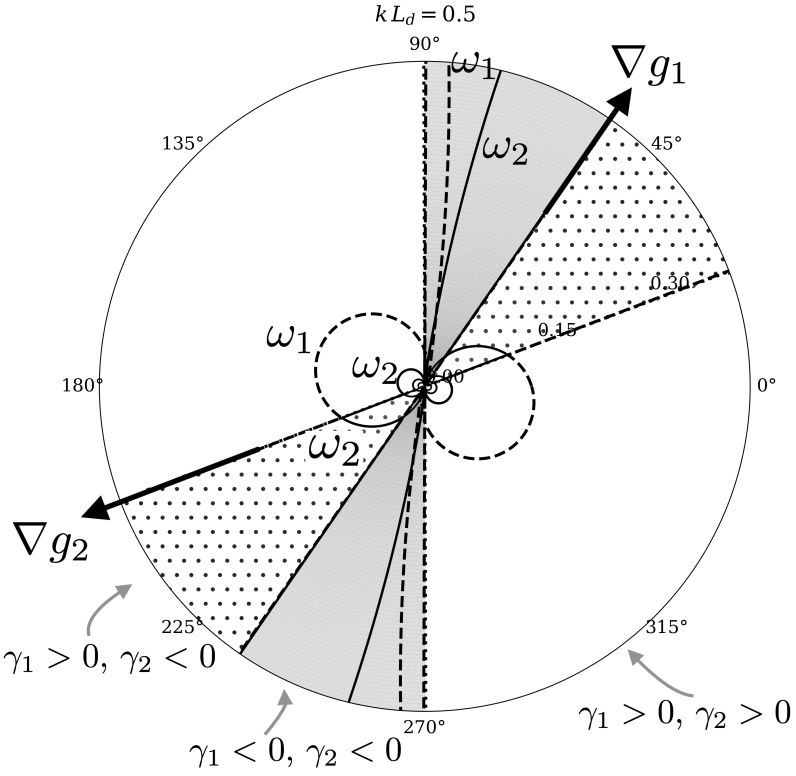

FIG. 4. Polar plots of the absolute value of the non-dimensional angular frequency $|\omega_n|/(\beta L_d)$ of the first five modes from section 3b as a function of the wave propagation direction $\boldsymbol{k}/|\boldsymbol{k}|$ for a horizontal wavenumber given by $k\,L_d = 0.5$ in constant stratification. The dotted line corresponds to ω_0 (not visible in this figure, but see figure 5), the dashed line to ω_1, with these two modes becoming boundary trapped at large wavenumbers $k = |\boldsymbol{k}|$. The remaining modes, ω_n for $n = 2, 3, 4$, are shown with solid lines. White regions are angles where $\gamma_1 > 0$ and $\gamma_2 > 0$. All Rossby waves with a propagation direction lying in the white region have negative angular frequencies ω_n and so have a westward phase speed. Gray regions are angles where $\gamma_1 < 0$ and $\gamma_2 < 0$. The two gravest angular frequencies ω_0 and ω_1 are both positive while the remaining angular frequencies ω_n for $n > 1$ are negative. Consequently, in the gray regions, ω_0 and ω_1 each correspond to a Rossby waves with an eastward phase speed whereas the remaining Rossby waves have westward phase speeds . Stippled regions are angles where $\gamma_1 > 0$ and $\gamma_2 < 0$. In the stippled region, ω_0 is positive and has an eastward phase speed. The remaining Rossby waves in the stippled region have negative angular frequencies and have westward phase speeds. The lower boundary buoyancy gradient, proportional to $\boldsymbol{\nabla} g_1$, points towards 55° while the upper boundary buoyancy gradient, proportional to $\boldsymbol{\nabla} g_2$, points towards 200°. These buoyancy gradients correspond to topographic gradients at the lower and upper boundaries through $g_j = f_0\, h_j$. The g_j are chosen so that $|\boldsymbol{\nabla} h_1| = 1.5 \times 10^{-5}$ and $|\boldsymbol{\nabla} h_2| = 10^{-5}$, thus leading to $\gamma_1/H = 0.15$ and $\gamma_2/H = 0.1$. The remaining parameters are as in figure 2. Plots of the first four modes, φ_n, for angles $\theta = 180°, 225°, 265°$ are shown in figure 6.

zeros. See the regions in white in figures 4 and 5 and plots (a) and (b) in figure 6.

ii. $\gamma_1\,\gamma_2 < 0$. There is one mode, φ_0, with a negative square, $\langle \varphi_0, \varphi_0 \rangle < 0$, corresponding to an eastward propagating wave. The associated eigenvalue, λ_0, satisfies $\lambda_0 < -k^2$. The remaining modes, φ_n for $n > 1$, have positive squares, $\langle \varphi_n, \varphi_n \rangle > 0$, corresponding to westward propagating waves and have eigenvalues λ_n satisfying $\lambda_n > -k^2$. Both φ_0 and φ_1 have no internal zeros whereas the remaining modes, φ_n, have $n-1$ internal zeros for $n > 1$. See the stippled regions in figures 4 and 5 and plots (c) and (d) in figure 6.

iii. $\gamma_1 < 0$ and $\gamma_2 < 0$. There are two modes φ_0 and φ_1 with negative squares, $\langle \varphi_n, \varphi_n \rangle < 0$, that propagate eastward and have eigenvalues, λ_n, satisfying $\lambda_n < -k^2$ for $n = 1, 2$. The remaining modes, φ_n, for $n > 1$ have positive squares, $\langle \varphi_n, \varphi_n \rangle > 0$, propagate westward, and have eigenvalues, λ_n, satisfying $\lambda_n > -k^2$. The zeroth mode, φ_0, has one internal zero, the first and second modes, φ_1 and φ_2, have no internal zeros, and the remaining modes, φ_n, have $n-2$ internal zeros for $n > 2$. See the shaded regions in figures 4 and 5 and plots (e) and (f) in figure 6.

Figures 4 and 5 illustrate that the mode number, n, no longer refers to a unique physical mode. Instead, n is

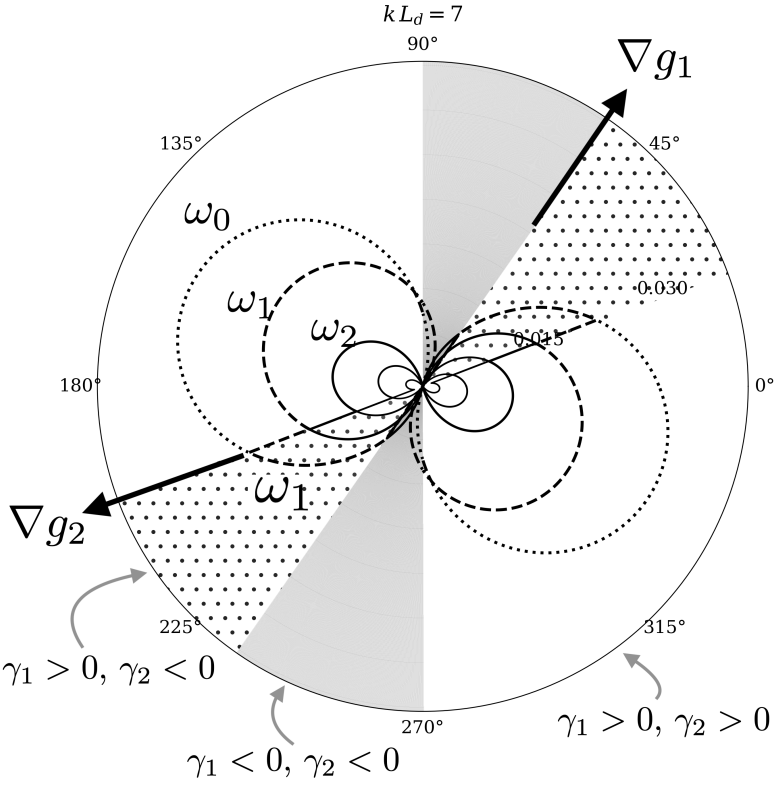

FIG. 5. As in figure 4 but with a larger horizontal wavenumber k given by $k\,L_d = 7$. The smaller horizontal length-scale means that ω_0 is small enough to be visible in this figure, unlike in figure 4 where ω_0 is large and thus sits outside of the figure. Plots of the first four modes, φ_n, for angles $\theta = 180°, 225°, 265°$ are shown in figure 6.

simply a mathematical label. For instance, in the white regions in figure 5, the dotted line ($n = 0$) corresponds to the bottom-trapped mode. Following the dotted line into the stippled region reveals the dotted line becomes a dashed line indicating that the bottom-trapped mode now has $n = 1$. A similar sort of mode switching was described by Hallberg (1997) in a two-layer quasigeostrophic model with bottom topography. As noted by Straub (1994), symmetry in figures 4 and 5 is related to the $k_y \to -k_y, \partial_x g_j \to -\partial_x g_j$ symmetry in the eigenvalue problem (50).

To elucidate the meaning of $\lambda_n < -k^2$ in cases (i)–(iii) above, note that a pure surface quasigeostrophic mode[8] has a vertical structure $e^{k(z-z_2)}$ decaying from the upper boundary at $z = z_2$; this exponential vertical structure implies a square vertical wavenumber $\lambda = -k^2$ (e.g., an e-folding scale of $|\lambda|^{-1/2} = k^{-1}$). Thus $\lambda_n < -k^2$ means that at least one of the boundary-trapped modes decays away from the boundary more rapidly than a pure surface quasigeostrophic wave. Indeed, the limit of $\lambda_0 \to -\infty$ yields one of the step-modes (46) of the previous subsection.

The step-mode limit is obtained as $\gamma_j \to 0^-$. This limit is found as either $|\boldsymbol{\nabla} g_j| \to 0$ for propagation directions in which $\gamma_j < 0$ or as $\boldsymbol{k}$ becomes parallel or anti-parallel to $\boldsymbol{\nabla} g_j$ (whichever limit satisfies $\gamma_j \to 0^-$). In this limit, we obtain a step-mode exactly confined at the boundary (the e-folding scale is now $|\lambda|^{-1/2} = 0$) with zero phase speed [see figure 7(a)]. The remaining modes then satisfy the isentropic boundary condition

$$\left(\frac{f_0^2}{N^2} \frac{\mathrm{d}\varphi_n}{\mathrm{d}z} \right) \bigg|_{z=z_j} = 0. \tag{56}$$

The other limit is that of $|\gamma_j| \to \infty$ which is obtained as the buoyancy gradient becomes large, $|\boldsymbol{\nabla} g_j| \to \infty$. In this limit, the eigenvalue $\lambda_m \to -k^2$ where λ_m is the eigenvalue of the mode φ_m associated with the jth boundary [see figure 7(b)]. Furthermore, the phase speed of the wave trapped to the jth boundary becomes infinite, an indication that the quasigeostrophic approximation breaks down. Indeed, the large buoyancy gradient

[8] A pure surface quasigeostrophic mode is the mode found after setting $\beta = 0$ with a lower boundary at $z_1 = -\infty$.

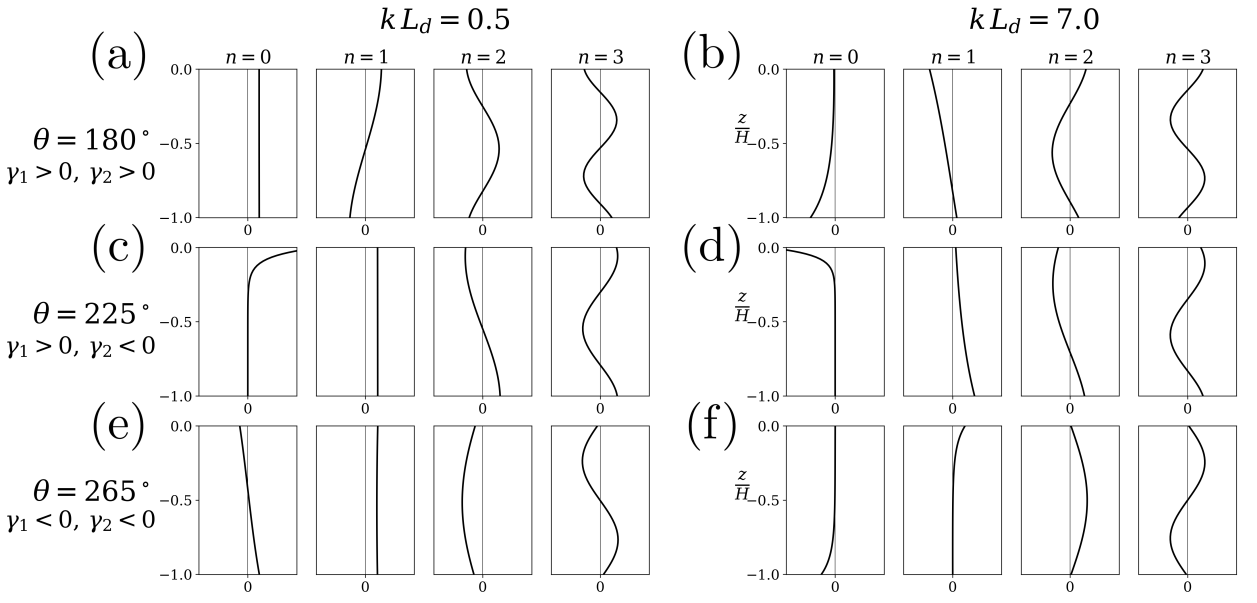

FIG. 6. This figure illustrates the dependence of the vertical structure φ_n of the streamfunction to the horizontal wavevector $\boldsymbol{k}$ as discussed in section 3b . Three propagation directions are shown $\theta = 180°, 225°, 265°$ and correspond to the rows in the figure [e.g., the row containing (a) and (b) are the vertical structures of waves at $\theta = 180°$]; two wavenumbers $k\,L_d = 0.5, 7$ are shown (where $k = |\boldsymbol{k}|$) and they correspond to the columns in the above figure [e.g., (b), (d) and (f) are the vertical structure of waves with $k\,L_d = 7$]. The parameters for the above figure are identical to figure 4. We emphasize two features in this figure. First, note how the boundary modes ($n = 0, 1$) are generally only boundary-trapped at small horizontal scales. At larger horizontal scales, we typically obtain a depth-independent mode along with another mode with large-scale features in the vertical. Second, note that for $\gamma_1, \gamma_2 > 0$, as in panels (a) and (b), the nth mode has n internal zeros, as in Sturm-Liouville theory; for $\gamma_1 > 0, \gamma_2 < 0$, as in panels (c) and (d), the first two modes ($n = 0, 1$) have no internal zeros; and for $\gamma_1, \gamma_2 < 0$, the zeroth mode φ_0 has one internal zero, the first and second modes, φ_1 and φ_2 have no internal zeros, and the third mode φ_2 has one internal zero. Note that the zero-crossing for the $n = 0$ mode in panel (f) is difficult to observe since the amplitude of φ_0 is small near the zero-crossing.

limit corresponds to steep topographic slopes and so we obtain the topographically-trapped internal gravity wave of Rhines (1970), which has an infinite phase speed in quasigeostrophic theory. The remaining modes then satisfy the vanishing pressure boundary condition

$$\varphi(z_j) = 0 \tag{57}$$

as in the surface modes of de La Lama et al. (2016) and LaCasce (2017).

A notable difference between the traditional baroclinic modes and the generalized Rhines modes is that the generalized Rhines modes depend on the horizontal wavevector, $\boldsymbol{k}$. However, as Straub (1994) observed, only the first few modes have a strong dependence on $\boldsymbol{k}$ whereas higher modes are generally insensitive to $\boldsymbol{k}$. All modes, φ_n with $n \gg 1$, satisfy the $\boldsymbol{k}$-independent boundary condition (57) if $\gamma_j \neq 0$. Alternatively, if $\gamma_j = 0$, the $n \gg 1$ modes instead satisfy boundary condition (56).

Boundary buoyancy gradients can modify the speed of Rossby waves by modifying their vertical structure. From the dispersion relation (41), the zonal phase speed $c_x = \omega/k_x$ is

$$c_x = -\frac{\beta}{k^2 + \lambda}, \tag{58}$$

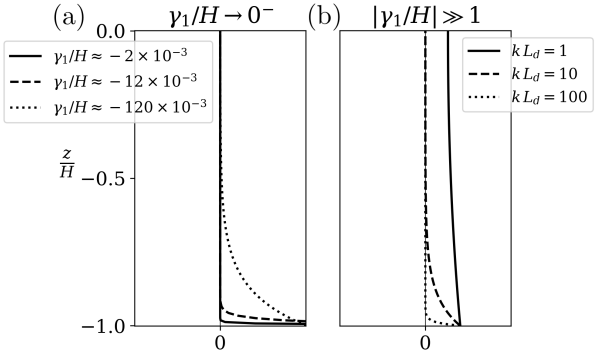

FIG. 7. The two limits of the boundary-trapped surface quasi-geostrophic waves, as discussed in section 3b. (a) Convergence to the step mode given in equation (46) with $j = 1$ as $\gamma_1 \to 0^-$ for three values of γ_1 at a wavenumber $k = |\boldsymbol{k}|$ given by $k\,L_d = 1$. The phase speed approaches zero in the limit $\gamma_1 \to 0^-$. (b) Here, $\gamma_1/H \approx 10$ for the three vertical structures φ_n shown. Consequently, the bottom trapped wave has a square vertical wavenumber $\lambda \approx -k^2$ and the phase speeds are large. The vertical structure, φ, for three values of $k\,L_d$ are shown, illustrating the dependence on k of this mode, which behaves as a boundary-trapped exponential mode with an e-folding scale of $|\lambda|^{-1/2} = k^{-1}$. In both (a) and (b), the wave propagation direction $\theta = 260°$. All other parameters are identical to figure 4.

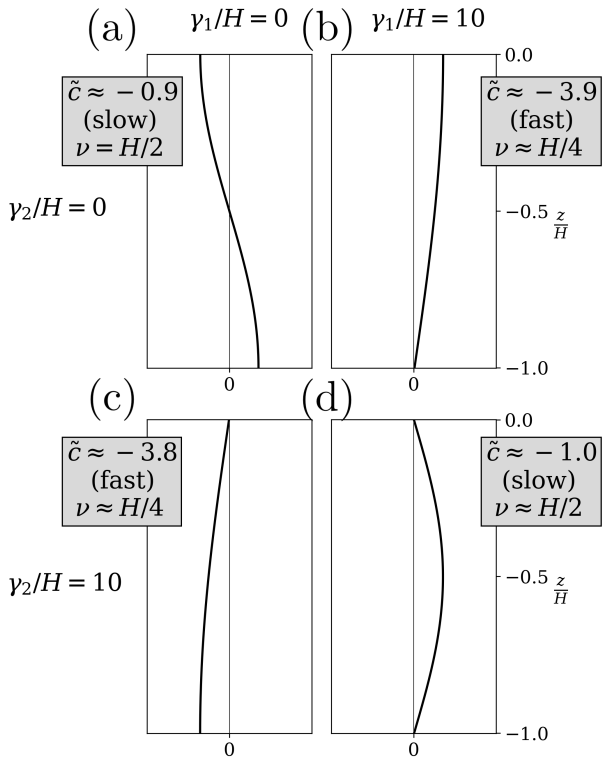

FIG. 8. The dependence of the Rossby wave phase speed on the vertical structure φ_n, as discussed in section 3b. The first internal mode is plotted. Two columns with $\gamma_1/H = 0, 10$ and two rows $\gamma_2/H = 0, 10$ are shown; e.g., (a) has $\gamma_1 = 0$ and $\gamma_2 = 0$ while (b) has $\gamma_1 = 10$ and $\gamma_2 = 0$. The non-dimensional phase speed $\tilde{c}$ is given by $\tilde{c} = c/(\beta L_d^2)$ and the vertical wavelength is given by $\nu = 2\pi/\sqrt{\lambda}$. The horizontal wavevector $\boldsymbol{k}$ points in the $\theta = 260°$ direction and has a magnitude given by $k L_d = 0.1$. Other parameters are as in figure 4.

which shows that the phase speed is inversely proportional to the square vertical wavenumber, λ. As shown by Rhines (1970), bottom topography increases the speed of Rossby waves by modifying the bottom boundary condition and hence decreasing the square vertical wavenumber, λ. Figure 8 shows that boundary buoyancy gradients at the upper surface can also modify the Rossby phase speed. Notably, when buoyancy gradients are present at the upper surface, bottom topography now decreases the phase speed. For example, the phase speed decreases as we increase γ_1 from $\gamma_1 = 0$ in figure 8(c) to $\gamma_1 = 10$ in figure 8(d). Of course, surface buoyancy gradients in the ocean arise in the presence of background mean flows and so equation (58) for c_x no longer holds. This analysis suggests that it would be interesting to investigate how the Rossby wave phase speed depends on surface buoyancy gradients.

In the above analysis, we have assumed $\gamma_j \neq 0$ for $j = 1, 2$. Suppose now that $\gamma_2 = 0$ so that the upper boundary is an isentropic surface. Then we must replace the upper boundary condition by the isentropic boundary condition (56) with $j = 2$; the resulting modes form a basis for $L^2 \oplus \mathbb{C}$ (i.e., there is now one "missing" degree of freedom). Condition (i) above holds if $\gamma_1 > 0$ and condition (ii) holds if $\gamma_1 < 0$ [see Yassin (in prep.) for more details]. The case when $\gamma_1 = 0$ but $\gamma_2 \neq 0$ is similar. If both γ_1 and γ_2 vanish then we recover the results of section 3a.

3) The general time-dependent solution

At some wavevector $\boldsymbol{k}$, the observed vertical structure now has the form

$$\psi_{\boldsymbol{k}}(z, t = 0) = \Psi(z) \tag{59}$$

where Ψ is a twice continuously differentiable function. For such functions we can write

$$\Psi = \sum_{n=0}^{\infty} \frac{\langle \Psi, \varphi_n \rangle}{\langle \varphi_n, \varphi_n \rangle} \varphi_n. \tag{60}$$

so that the time-evolution is

$$\psi_{\boldsymbol{k}}(z, t) = \sum_{n=0}^{\infty} \frac{\langle \Psi, \varphi_n \rangle}{\langle \varphi_n, \varphi_n \rangle} \varphi_n(z)\, \mathrm{e}^{-\mathrm{i}\omega_n t}. \tag{61}$$

Again, it is the above expression, which is valid only in linear theory with a quiescent background state, that gives the generalized Rhines modes φ_n physical meaning. Outside the linear theory of this section, the generalized Rhines modes do not have any physical interpretation and instead merely serve as a mathematical basis for $L^2 \oplus \mathbb{C}^2$.

c. The vertical velocity eigenvalue problem

Let $\hat{w}(z) = \hat{w}_0 \chi(z)$ where $\chi(z)$ is a non-dimensional function. For the Rossby waves with isentropic boundaries of section 3a (the traditional baroclinic modes), the corresponding vertical velocity modes satisfy

$$-\frac{\mathrm{d}^2 \chi}{\mathrm{d}z^2} = \lambda \left(\frac{N^2}{f_0^2} \right) \chi \tag{62}$$

with vanishing vertical velocity boundary conditions

$$\chi(z_j) = 0 \tag{63}$$

(see the appendix for details). The resulting modes $\{\chi_n\}_{n=0}^{\infty}$ form an orthonormal basis of L^2 with orthonormality given by

$$\delta_{mn} = \frac{1}{H} \int_{z_1}^{z_2} \chi_m \chi_n \left(\frac{N^2}{f_0^2} \right) \mathrm{d}z. \tag{64}$$

One can obtain the eigenfunctions, χ_n, by solving the eigenvalue problem (62)–(63) or by differentiating the streamfunction modes ϕ_n according to equation (A12). The quasigeostrophic vertical velocity eigenvalue problem (62)–(63) is often used as an alternative to the streamfunction eigenvalue problem (40) (e.g., Chelton et al. 1998; Ferrari et al. 2010).

1) Relating to the gravity wave problem

The quasigeostrophic vertical velocity eigenvalue problem (62)–(63) is isomorphic to the vertical velocity eigenvalue problem for internal gravity waves with rigid boundaries (e.g. Gill 1982, section 6.11). In the Rossby wave problem, the eigenvalue λ is given by the dispersion relation (41), while in the internal gravity wave problem, the eigenvalue λ is given by

$$c^2_{\text{gravity}} = \frac{\lambda}{f_0^2} \tag{65}$$

where c_{gravity} is the speed of internal gravity waves.

However, the isomorphism is lost once boundary dynamics are activated. In the case of gravity waves, boundary dynamics are activated in the presence of a free-surface [see Kelly (2016) and Yassin (in prep.)]. In contrast, one can never apply a free-surface boundary condition to the *continuously stratified* quasigeostrophic problem as it is asymptotically inconsistent (Pedlosky 1982, section 6.9, p. 345). Instead, there is a genuine rigid-lid at the level of the continuously stratified quasigeostrophic approximation[9].

2) Quasigeostrophic boundary dynamics

As seen earlier, boundary buoyancy gradients activate boundary dynamics in the quasigeostrophic problem. In this case, boundary conditions for the quasigeostrophic vertical velocity problem (62) become

$$-(-1)^j \gamma_j k^2 \frac{\mathrm{d}\chi}{\mathrm{d}z}\bigg|_{z_j} = \lambda \left[\chi|_{z_j} + (-1)^j \gamma_j \frac{\mathrm{d}\chi}{\mathrm{d}z}\bigg|_{z_j} \right] \tag{66}$$

(see the appendix). The resulting modes $\{\chi_n\}_{n=0}^{\infty}$ form an orthonormal basis of $L^2 \oplus \mathbb{C}^2$ and satisfy a peculiar orthogonality relation given by equation (A18).

In the presence of boundary buoyancy gradients, $\hat{\psi}$ and $\hat{w}$ become proportional at the boundaries [see equation (A15)]. Indeed, figure 9 shows that, while for the $n > 1$ internal modes $\hat{\psi}_n$ and $\hat{w}_n$ are related by differentiation, for the two boundary-trapped modes ($n = 0, 1$), the vertical structures of $\hat{\psi}$ and $\hat{w}$ are nearly identical. It follows that, in the limit $\gamma_j \to 0^-$, we obtain a stationary vertical velocity step-mode

$$\chi_j^{\text{step}} = \phi_j^{\text{step}}. \tag{67}$$

Similarly, in the limit of $|\gamma_j| \to \infty$, we obtain an infinitely fast boundary-trapped vertical velocity mode with a vertical e-folding of k^{-1}.

4. Eigenfunction expansions

Motivated by the Rossby waves of the previous section, we now investigate various sets of normal modes for quasigeostrophic theory. Let $\{\phi_n\}_{n=0}^{\infty}$ be a collection of L^2 normal modes, and assume $\psi_{\boldsymbol{k}}(z,t)$ is twice continuously differentiable in z. Completeness in L^2 implies that the eigenfunction expansion $\psi_{\boldsymbol{k}}^{\text{exp}}$, defined by

$$\psi_{\boldsymbol{k}}^{\text{exp}}(z,t) = \sum_{n=0}^{\infty} \psi_{\boldsymbol{k}n}(t)\, \phi_n(z) \tag{68}$$

where

$$\psi_{\boldsymbol{k}n} = \langle \psi_{\boldsymbol{k}}, \phi_n \rangle_0 \tag{69}$$

satisfies

$$\int_{z_1}^{z_2} |\psi_{\boldsymbol{k}}(z) - \psi_{\boldsymbol{k}}^{\text{exp}}(z)|^2 \mathrm{d}z = 0. \tag{70}$$

Significantly, the vanishing of the integral (70) does not imply $\psi_{\boldsymbol{k}} = \psi_{\boldsymbol{k}}^{\text{exp}}$ since the two functions may differ at some points $z \in [z_1, z_2]$.

We now list physically desirable properties for a collection of normal modes. In general $\psi_{\boldsymbol{k}} \neq \psi_{\boldsymbol{k}}^{\text{exp}}$, but this lack of equality is not desirable physically. Hence, our first condition is that the expansion $\psi_{\boldsymbol{k}}^{\text{exp}}$ be pointwise equal to $\psi_{\boldsymbol{k}}$ on the closed interval $[z_1, z_2]$,

$$\psi_{\boldsymbol{k}}(z) = \psi_{\boldsymbol{k}}^{\text{exp}}(z) \quad \text{for } z \in [z_1, z_2]. \tag{71}$$

If such pointwise equality is our only requirement for the normal modes, then it is not necessary to restrict ourselves to quasigeostrophic normal modes. Indeed, any basis of L^2 or $L^2 \oplus \mathbb{C}^2$ satisfying equation (71) will suffice. However, we also require that the resulting series must be differentiable (we will see this requirement excludes L^2 bases) and that the normal modes partition the energy, volume potential vorticity, and surface potential vorticity amongst the various modes. That is, we insist that the normal modes diagonalize the energy and potential enstrophy integrals as in Section 2e. This diagonalization requirement restricts us to quasigeostrophic normal modes, and we will show that only the $L^2 \oplus \mathbb{C}^2$ modes satisfy all our requirements.[10]

a. *The four possible L^2 modes*

There are only four L^2 bases in quasigeostrophic theory that diagonalize the energy. All four sets of corresponding normal modes satisfy the differential equation

$$-\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\phi}{\mathrm{d}z}\right) = \lambda\,\phi \quad z \in (z_1, z_2), \tag{72}$$

but differ in boundary conditions according to the following.

- *Baroclinic modes*: Vanishing vertical velocity at both boundaries (Neumann),

$$\frac{\mathrm{d}\phi(z_1)}{\mathrm{d}z} = 0, \quad \frac{\mathrm{d}\phi(z_2)}{\mathrm{d}z} = 0. \tag{73}$$

[9] In contrast, one can include a free surface in stacked shallow-water quasigeostrophic models. See chapter 5 in Vallis (2017) for more details.

[10] The reader should consult Yassin (in prep.) for appropriate references regarding the mathematical facts used in this section.

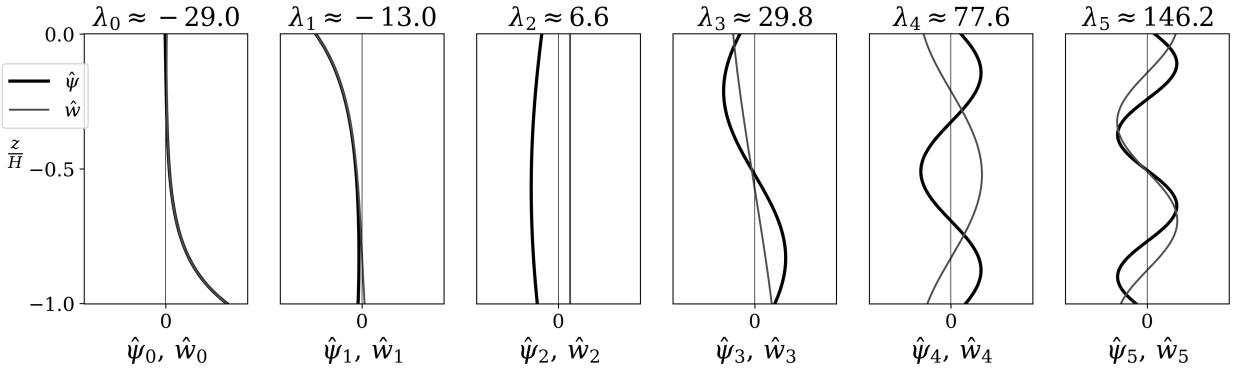

FIG. 9. The first six vertical velocity normal modes χ_n (thin grey lines) and streamfunction normal modes φ_n (black lines) (see section 3c). The propagation direction is $\theta = 75°$ with a wavenumber of $k\,L_d = 2$. The remaining parameter are as in figure 4. Note that χ_n and φ_n are nearly indistinguishable for the boundary-trapped modes $n = 0, 1$ while they are related by a vertical derivative for the internal modes $n > 1$. The eigenvalue in the figure is non-dimensionalized by the deformation radius L_d.

- *Anti-baroclinic modes*: Vanishing pressure[11] at both boundaries (Dirichlet),

$$\phi(z_1) = 0, \quad \phi(z_2) = 0. \tag{74}$$

- *Surface modes*: (mixed Neumann/Dirichlet)

$$\phi(z_1) = 0, \quad \frac{\mathrm{d}\phi(z_2)}{\mathrm{d}z} = 0. \tag{75}$$

- *Anti-surface modes*: (mixed Neumann/Dirichlet)

$$\frac{\mathrm{d}\phi(z_1)}{\mathrm{d}z} = 0, \quad \phi(z_2) = 0. \tag{76}$$

All four sets of modes are missing two modes. Each boundary condition of the form

$$\frac{\mathrm{d}\phi(z_j)}{\mathrm{d}z} = 0, \tag{77}$$

implies a missing step-mode while a boundary condition of the form

$$\phi(z_j) = 0, \tag{78}$$

implies a missing boundary-trapped exponential mode.

b. Expansions with L^2 modes

We here examine the utility of eigenfunction expansions in terms of L^2 modes, considering both pointwise equality and differentiability.

[11] Recall that the geostrophic streamfunction ψ is proportional to pressure (e.g., Vallis 2017, section 5.4).

1) POINTWISE EQUALITY ON $[z_1, z_2]$

We now consider the convergence behaviour of the series expansion $\psi_{\boldsymbol{k}}^{\mathrm{exp}}$ of $\psi_{\boldsymbol{k}}$ in terms of L^2 modes. For all four sets of L^2 modes, if $\psi_{\boldsymbol{k}}$ is twice continuously differentiable in z, we obtain equality in the interior

$$\psi_{\boldsymbol{k}}(z) = \psi_{\boldsymbol{k}}^{\mathrm{exp}}(z) \quad \text{for } z \in (z_1, z_z). \tag{79}$$

The behaviour at the boundaries depends on the boundary conditions the modes ϕ_n satisfy. If the ϕ_n satisfy the vanishing pressure boundary condition at the jth boundary

$$\phi_n(z_j) = 0 \tag{80}$$

then

$$\psi_{\boldsymbol{k}}^{\mathrm{exp}}(z_j) = 0 \tag{81}$$

regardless of the values of $\psi_{\boldsymbol{k}}(z_j)$. Hence, $\psi_{\boldsymbol{k}}^{\mathrm{exp}}$ generally has a jump discontinuity at the jth boundary and is not equal to $\psi_{\boldsymbol{k}}$ on the closed interval $[z_1, z_2]$. In contrast, if the ϕ_n satisfy a zero vertical velocity boundary condition at the jth boundary

$$\frac{\mathrm{d}\phi_n(z_j)}{\mathrm{d}z} = 0 \tag{82}$$

then

$$\psi_{\boldsymbol{k}}(z_j) = \psi_{\boldsymbol{k}}^{\mathrm{exp}}(z_j). \tag{83}$$

Consequently, of the four sets of L^2 modes, only with the baroclinic modes do we obtain the pointwise equality $\psi_{\boldsymbol{k}}(z) = \psi_{\boldsymbol{k}}^{\mathrm{exp}}(z)$ on the closed interval $[z_1, z_2]$. Indeed, with the baroclinic modes, the convergence of the series Ψ_{exp} to Ψ on $[z_1, z_2]$ is uniform and so no Gibbs phenomenon is present.

However, even though we can represent $\psi_{\boldsymbol{k}}$ in terms of the baroclinic modes, we are unable to represent the corresponding velocity $w_{\boldsymbol{k}}$ in terms of the vertical velocity

baroclinic modes since the modes vanish at both boundaries. Analogous considerations show that only the anti-baroclinic vertical velocity modes (see the appendix) can represent arbitrary vertical velocities.

2) Differentiability of the series expansion

Although we obtain pointwise equality on the whole interval $[z_1, z_2]$, we have lost two degrees of freedom in the expansion process. Recall that the degrees of freedom in the quasigeostrophic phase space are determined by the potential vorticity. The volume potential vorticity, $q_{\boldsymbol{k}}$, is associated with the L^2 degrees of freedom while the surface potential vorticities, $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$, are associated with the $\mathbb{C}^2$ degrees of freedom.

The series expansion $\psi^{\text{exp}}_{\boldsymbol{k}}$ of $\psi_{\boldsymbol{k}}$ in terms of the baroclinic modes is differentiable in the interior (z_1, z_2). Consequently, we can differentiate the $\psi^{\text{exp}}_{\boldsymbol{k}}$ series for $z \in (z_1, z_2)$ to recover $q_{\boldsymbol{k}}$, that is,

$$q_{\boldsymbol{k}} = \sum_{n=0}^{\infty} q_{\boldsymbol{k}n}\, \phi_n \tag{84}$$

where

$$q_{\boldsymbol{k}n} = -(k^2 + \lambda_n)\, \psi_{\boldsymbol{k}n}. \tag{85}$$

However, $\psi^{\text{exp}}_{\boldsymbol{k}}$ is not differentiable at the boundaries, $z = z_1, z_2$, so we are unable to recover the surface potential vorticities, $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$. Two degrees of freedom are lost by projecting onto the baroclinic modes.[12]

The energy at wavevector $\boldsymbol{k}$ is indeed partitioned between the modes

$$E_{\boldsymbol{k}} = \sum_{n=0}^{\infty} (k^2 + \lambda_n)\, \psi_{\boldsymbol{k}n} \tag{86}$$

and similarly for the potential enstrophy

$$Z_{\boldsymbol{k}} = \sum_{n=0}^{\infty} (k^2 + \lambda_n)^2 \psi_{\boldsymbol{k}n}. \tag{87}$$

However, as we have lost $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$ in the projection process, the surface potential enstrophies $Y_{1\boldsymbol{k}}$ and $Y_{2\boldsymbol{k}}$, defined in equation (34), are not partitioned.

[12]To see that $\psi^{\text{exp}}_{\boldsymbol{k}}$ is non-differentiable at $z = z_1, z_2$, suppose that the series $\psi^{\text{exp}}_{\boldsymbol{k}}$ is differentiable and that $\mathrm{d}\psi_{\boldsymbol{k}}(z_j)/\mathrm{d}z \neq 0$ for $j = 1, 2$. But then

$$0 \neq \frac{\mathrm{d}\psi_{\boldsymbol{k}}(z_j)}{\mathrm{d}z} = \sum_{n=0}^{\infty} \psi_{\boldsymbol{k}n} \frac{\mathrm{d}\phi_n(z_j)}{\mathrm{d}z} = 0,$$

which is a contradiction.

c. *Quasigeostrophic $L^2 \oplus \mathbb{C}^2$ modes*

Consider the eigenvalue problem

$$-\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\varphi}{\mathrm{d}z}\right) = \lambda\varphi \quad \text{for } z \in (z_1, z_2) \tag{88a}$$

$$-k^2\varphi + (-1)^j D_j^{-1}\left(\frac{f_0^2}{N^2}\frac{\mathrm{d}\varphi}{\mathrm{d}z}\right) = \lambda\varphi \quad \text{for } z = z_j \tag{88b}$$

where D_1 and D_2 are non-zero real constants. This eigenvalue problem differs from the generalized Rhines eigenvalue problem (50) in that D_j are generally not equal to the γ_j defined in equation (51). The inner product $\langle\cdot,\cdot\rangle$ induced by the eigenvalue problem (88) is given by equation (53) with the γ_j replaced by the D_j.

Smith and Vanneste (2012) investigate an equivalent eigenvalue problem to (88) and conclude that, when D_1 and D_2 are positive, the resulting eigenfunctions form a basis of $L^2 \oplus \mathbb{C}^2$. However, such a completeness result is insufficient for the Rossby wave problem of section 3b, in which case $D_j = \gamma_j$ and γ_j can be negative. Subsequently, Yassin (in prep.) examines the general theory of eigenvalue problems with λ-dependent boundary conditions. A consequence of the theory is that the eigenfunctions of (88) form a complete basis of $L^2 \oplus \mathbb{C}^2$ when $D_1, D_2 \neq 0$.

The eigenvalue problem (88) contains as a special case every collection of discrete normal modes for quasigeostrophic theory. The generalized Rhines modes of section 3b are obtained by choosing $D_j = \gamma_j$ where γ_j is given in equation (51). The surface-aware basis of Smith and Vanneste (2012) correspond to the case $D_1, D_2 > 0$. All four L^2 bases of section 4a, such as the traditional baroclinic modes and the surface modes, are singular limits when $|D_j| \to 0$ or $|D_j| \to \infty$.

d. *Expansion with $L^2 \oplus \mathbb{C}^2$ modes*

When D_1, D_2 in the eigenvalue problem (88) are finite and non-zero, the resulting eigenmodes $\{\varphi_n\}_{n=0}^{\infty}$ form a basis for the vertical structure phase space $L^2 \oplus \mathbb{C}^2$. Thus, the projection

$$\psi^{\text{exp}}_{\boldsymbol{k}}(z) = \sum_{n=0}^{\infty} \psi_{\boldsymbol{k}n}\, \varphi_n(z) \tag{89}$$

where

$$\psi_{\boldsymbol{k}n} = \frac{\langle \psi_{\boldsymbol{k}}, \varphi_n \rangle}{\langle \varphi_n, \varphi_n \rangle} \tag{90}$$

is an *equivalent* representation of $\psi_{\boldsymbol{k}}$. Not only do we have pointwise equality

$$\psi_{\boldsymbol{k}}(z) = \psi^{\text{exp}}_{\boldsymbol{k}}(z) \quad \text{for } z \in [z_1, z_2], \tag{91}$$

but the series $\psi^{\text{exp}}_{\boldsymbol{k}}$ is also differentiable on the *closed* interval $[z_1, z_2]$. Thus given $\psi^{\text{exp}}_{\boldsymbol{k}}$, we can differentiate to obtain

both $q_{\boldsymbol{k}}$ and $r_{j\boldsymbol{k}}$ and thereby recover all quasigeostrophic degrees of freedom. Indeed, we have

$$q_{\boldsymbol{k}}(z,t) = \sum_{n=0}^{\infty} q_{\boldsymbol{k}n}(t)\,\varphi_n(z) \tag{92}$$

$$r_{j\boldsymbol{k}}(z,t) = \sum_{n=0}^{\infty} r_{j\boldsymbol{k}n}(t)\,\varphi_n(z_j) \tag{93}$$

where

$$q_{\boldsymbol{k}n} = -(k^2+\lambda_n)\frac{\langle \Psi, \varphi_n\rangle}{\langle \varphi_n, \varphi_n\rangle} \tag{94}$$

$$r_{j\boldsymbol{k}n} = D_j\, q_{\boldsymbol{k}n} \tag{95}$$

for $j = 1, 2$. In fact, we have just obtained an expansion of the relative potential vorticity distribution of Bretherton (1966) given by equation (24),

$$\mathfrak{q}_{\boldsymbol{k}} = \sum_{n=0}^{\infty} q_{\boldsymbol{k}n}\left[1 + \sum_{j=1}^{2} D_j\,\delta(z-z_j)\right]\varphi_n. \tag{96}$$

In addition, the energy, $E_{\boldsymbol{k}}$, volume potential enstrophy, $Z_{\boldsymbol{k}}$, and surface potential enstrophies, $Y_{1\boldsymbol{k}}$ and $Y_{2\boldsymbol{k}}$, are partitioned (diagonalized) between the modes

$$E_{\boldsymbol{k}} = \sum_{n=0}^{\infty} (k^2+\lambda_n)\psi_{\boldsymbol{k}n}, \tag{97}$$

$$Z_{\boldsymbol{k}} + \frac{1}{H}\sum_{j=1}^{2}\frac{1}{D_j}Y_{j\boldsymbol{k}} = \sum_{n=0}^{\infty}(k^2+\lambda_n)^2\psi_{\boldsymbol{k}n}. \tag{98}$$

5. Discussion

a. Projecting model/observational data

If one seeks modes to project model or observational data, which modes should be chosen? If our only concern is reproducing the shape of the vertical structure, then we can use any L^2 basis (not just those listed in this article) where the basis functions do not vanish at the boundary. Alternatively, any $L^2 \oplus \mathbb{C}^2$ basis can be used and these modes will ensure the differentiability of the series expansion as well.

However, it is typically desirable that the projection also diagonally partitions the energy and potential vorticity amongst the eigenmodes. Of the four quasigeostrophic L^2 modes discussed in section 4a, only the baroclinic modes are capable of exactly representing a function and partitioning the energy. The baroclinic modes, however, can only partition the volume potential vorticity, $q_{\boldsymbol{k}}$, amongst the modes whereas the surface potential vorticities, $r_{j\boldsymbol{k}}$, are lost in the projection process. In addition a series expansion in terms of the baroclinic modes will generally not be differentiable at the boundaries.

For the $L^2 \oplus \mathbb{C}^2$ modes, any set of modes obtained from the eigenvalue problem (88) with $D_1, D_2 \neq 0$ will diagonally partition the energy, the volume potential vorticity, $q_{\boldsymbol{k}}$, and the surface potential vorticities, $r_{j\boldsymbol{k}}$, amongst the modes. Therefore, in a non-linear ocean with non-isentropic boundaries, there are infinitely many ways (parameterized by D_1 and D_2) of decomposing the vertical structure into normal modes. We have no physical argument prefering one decomposition over another.

Indeed, we emphasize that modal projection is a mathematical procedure that merely provides an alternate representation of the vertical structure as a sum of normal modes. For instance, the presence of the barotropic mode in some decompositions does not imply that there is any coherent barotropic motion. Rather, it is only in the appropriate linear setting with quiescent background, as discussed in section 3, that such an interpretation holds.

We also note that the projection process in terms of the baroclinic modes can be highly misleading. For instance, consider a realistic ocean with strong surface quasigeostrophic dynamics at the upper surface and weak interior dynamics. Most of the energy in this ocean is associated with motion having large upper surface potential vorticity, $r_{2\boldsymbol{k}}$, and small volume potential vorticity, $q_{\boldsymbol{k}}$. However, a projection onto baroclinic modes will partition the energy amongst the traditional baroclinic modes and may give the incorrect impression that the motion is all due to the interior potential vorticity $q_{\boldsymbol{k}}$.

b. What vertical mode does the altimeter reflect?

It is the misleading nature of the baroclinic modes that is presently the basis of the debate on what mode a surface altimeter signal reflects. Returning to the controversy from the introduction between Wunsch (1997) and Lapeyre (2009), we only state that the question of what "mode" the surface signal represents is itself ill-posed. Outside of the linear theory of section 3, the concept of a "mode" is somewhat arbitrary and certainly does not correspond to any coherent motion.

An alternate and more physical question is the following: Is the altimeter signal due to interior (induced by Q) or surface (induced by R_2) dynamics? However, it is not possible to answer this question using modal decompositions alone.

c. Galerkin approximations with L^2 modes

Both the L^2 baroclinic modes and the $L^2 \oplus \mathbb{C}^2$ modes have infinitely many degrees of freedom. In contrast, numerical simulations only contain a finite number of degrees of freedom. Consequently, it should be possible to use baroclinic modes to produce a Galerkin approximation to quasigeostrophic theory with non-trivial boundary dynamics. Such an approach has already been proposed by Rocha et al. (2015).

Projecting $\psi_{\boldsymbol{k}}$ onto the baroclinic modes produces a series expansion, $\psi^{\text{exp}}_{\boldsymbol{k}}$, that is differentiable in the interior but not at the boundaries. By differentiating the series *in the interior* we obtain equation (85) for $q_{\boldsymbol{k}n}$. If instead we integrate by parts twice and avoid differentiating $\psi^{\text{exp}}_{\boldsymbol{k}}$, we obtain

$$q_{\boldsymbol{k}n} = -(k^2+\lambda_n)\psi_{\boldsymbol{k}n} - \frac{1}{H}\sum_{j=1}^{2} r_{j\boldsymbol{k}}\,\phi_n(z_j). \tag{99}$$

The two expressions (85) and (99) are only equivalent when $r_{1\boldsymbol{k}} = r_{2\boldsymbol{k}} = 0$. For non-zero $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$, the singular nature of the expansion means we have a choice between equations (85) and (99).

By choosing equation (99) and avoiding the differentiation of $\psi^{\text{exp}}_{\boldsymbol{k}}$, Rocha et al. (2015) produced a least-squares approximation to quasigeostrophic dynamics that conserves the surface potential enstrophy integrals (33). This is a conservation property underlying the approximation's success.

6. Conclusion

In this article, we have studied all possible non-continuum collections of streamfunction normal modes that diagonalize the energy and potential enstrophy. There are four possible L^2 modes: the baroclinic modes, the anti-baroclinic modes, the surface modes, and the anti-surface modes. Additionally, we developed a family of $L^2\oplus\mathbb{C}^2$ bases parameterized by the parameters D_1, D_2. If $D_j = \gamma_j$, where γ_j is given by equation (51) for $j = 1, 2$, the resulting modes are the vertical structure of Rossby waves in a quiescent ocean with prescribed boundary buoyancy gradients. We have also examined the associated L^2 and $L^2\oplus\mathbb{C}^2$ vertical velocity modes.

For the streamfunction L^2 modes, only the baroclinic modes are capable of exactly representing any quasigeostrophic state on the interval $[z_1, z_2]$, whereas for the vertical velocity L^2 modes, only the anti-baroclinic modes are capable. However, in both cases, the resulting eigenfunction expansion is not differentiable at the boundaries, $z = z_1, z_2$. Consequently, while we can recover the volume potential vorticity density, $q_{\boldsymbol{k}}$, we cannot recover the surface potential vorticity densities, $r_{1\boldsymbol{k}}$ and $r_{2\boldsymbol{k}}$. Thus, we lose two degrees of freedom when projecting onto the baroclinic modes. In contrast, $L^2\oplus\mathbb{C}^2$ modes provide an equivalent representation of the function in question. Namely, the eigenfunction expansion is differentiable on the closed interval $[z_1, z_2]$ so that we can recover $q_{\boldsymbol{k}}$, $r_{1\boldsymbol{k}}$, $r_{2\boldsymbol{k}}$ from the series expansion.

A natural application of the proposed $L^2\oplus\mathbb{C}^2$ normal modes is to the study of weakly non-linear wave-interaction theories of geostrophic turbulence found in Fu and Flierl (1980) and Smith and Vallis (2001), extending their work to include prescribed surface buoyancy gradients. Another application is to the extension of equilibrium statistical mechanical calculations (e.g., Venaille et al. 2012) to include surface quasigeostrophic degrees of freedom. Both applications will be reported in subsequent contributions.

Acknowledgments. We offer sincere thanks to Stephen Garner, Robert Hallberg, Isaac Held, and Sonya Legg for comments and suggestions that greatly helped our presentation. This report was prepared by Houssam Yassin under award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

Data availability statement.

APPENDIX

Polarization relations and the vertical velocity eigenvalue problem

a. Polarization relations

The linear quasigeostrophic vorticity and buoyancy equations, computed about a resting background state, are

$$\frac{\partial\zeta}{\partial t} + \beta\frac{\partial\psi}{\partial x} = f_0\frac{\partial w}{\partial z}, \tag{A1}$$

$$\frac{\partial b}{\partial t} = -N^2 w \tag{A2}$$

in the interior $z \in (z_1, z_2)$. The vorticity, ζ, and buoyancy, b, are given in terms of the geostrophic streamfunction via

$$\zeta = \nabla^2\psi \tag{A3}$$

$$b = f_0\frac{\partial\psi}{\partial z}, \tag{A4}$$

The no-normal flow at the lower and upper boundaries implies

$$f_0 w = \boldsymbol{u}\cdot\boldsymbol{\nabla} g_j \tag{A5}$$

for $j = 1, 2$. Substituting equation (A5) into the linear buoyancy equation (A2), yields the boundary conditions

$$\partial_t b + \boldsymbol{u}\cdot\boldsymbol{\nabla}\left(\frac{N^2}{f_0}g_j\right) = 0 \quad \text{for } z = z_j. \tag{A6}$$

We now assume solutions of the form

$$\psi = \hat{\psi}(z)\, e_{\boldsymbol{k}}(\boldsymbol{x})\, \mathrm{e}^{-\mathrm{i}\omega t} \tag{A7}$$

and similarly for the variables ζ, b, w, u, v. Substituting such solutions into equations (A1)–(A2) and using $\boldsymbol{u} = \hat{\boldsymbol{z}}\times\boldsymbol{\nabla}\psi$ gives

$$\hat{\zeta} = -k^2\hat{\psi} \tag{A8}$$

$$\hat{u} = -\mathrm{i}\,k_y\hat{\psi} \tag{A9}$$

$$\hat{v} = \mathrm{i}\,k_x\hat{\psi} \tag{A10}$$

for $z \in [z_1, z_2]$, and

$$\hat{b} = -\mathrm{i}\,\frac{N^2}{\omega}\,\hat{w} \tag{A11}$$

$$\frac{\mathrm{d}\hat{\psi}}{\mathrm{d}z} = -\mathrm{i}\,\frac{N^2}{f_0\,\omega}\hat{w} \tag{A12}$$

$$\frac{\mathrm{d}\hat{w}}{\mathrm{d}z} = \mathrm{i}\,\frac{\omega}{f_0}\left[k^2 + \frac{\beta\,k_x}{\omega}\right]\hat{\psi}. \tag{A13}$$

for $z \in (z_1, z_2)$. At the boundaries $z = z_1, z_2$, we use equations (A5) and (A6) to obtain

$$\hat{b} = -\frac{N^2}{f_0\,\omega}\,\hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla} g_j \tag{A14}$$

$$\hat{w} = \mathrm{i}\,\frac{1}{f_0}\,\hat{\boldsymbol{u}}\cdot\boldsymbol{\nabla} g_j. \tag{A15}$$

b. The vertical velocity eigenvalue problem

Taking the vertical derivative of (A13) and using (A12) yields

$$-\frac{\mathrm{d}^2\chi}{\mathrm{d}z^2} = \lambda\left(\frac{N^2}{f_0^2}\right)\chi \tag{A16}$$

where $\hat{w} = w_0\,\chi(z)$ and χ is non-dimensional. The boundary conditions at $z = z_j$ are

$$-(-1)^j\,\gamma_j\,k^2\,\frac{\mathrm{d}\chi}{\mathrm{d}z} = \lambda\left[\chi + (-1)^j\,\gamma_j\,\frac{\mathrm{d}\chi}{\mathrm{d}z}\right] \tag{A17}$$

as obtained by using equations (A13) and (A12) in boundary conditions (50b).

As shown in Yassin (in prep.), the eigenfunctions $\{\chi_n\}_{n=0}^{\infty}$ of the above problem form an orthonormal basis of $L^2 \oplus \mathbb{C}^2$ when $\gamma_j \neq 0$ for $j = 1, 2$. In this case, the orthonormality condition is

$$\pm\delta_{mn} = \frac{1}{H}\left[\int_{z_1}^{z_2}\chi_m\,\chi_n\left(\frac{N^2}{f_0^2}\right)\mathrm{d}z - \frac{1}{k^2}\sum_{j=1}^{2}\frac{1}{\gamma_j}\left(\mathcal{B}_j\chi_m\right)\left(\mathcal{B}_j\chi_n\right)\right] \tag{A18}$$

where

$$\mathcal{B}_j\chi = \chi(z_j) + (-1)^j\,\gamma_j\,\frac{\mathrm{d}\chi(z_j)}{\mathrm{d}z}. \tag{A19}$$

c. The vertical velocity L^2 modes

Analogously with the streamfunction L^2 modes, we have the following sets of vertical velocity L^2 modes.

- *Baroclinic modes*: Vanishing vertical velocity at both boundaries,

$$\chi(z_1) = 0, \quad \chi(z_2) = 0. \tag{A20}$$

- *Anti-baroclinic modes*: Vanishing pressure at both boundaries,

$$\frac{\mathrm{d}\chi(z_1)}{\mathrm{d}z} = 0, \quad \frac{\mathrm{d}\chi(z_2)}{\mathrm{d}z} = 0. \tag{A21}$$

- *Surface modes*:

$$\frac{\mathrm{d}\chi(z_1)}{\mathrm{d}z} = 0, \quad \chi(z_2) = 0. \tag{A22}$$

- *Anti-surface modes*:

$$\chi(z_1) = 0, \quad \frac{\mathrm{d}\chi(z_2)}{\mathrm{d}z} = 0. \tag{A23}$$

References

Balmforth, N. J., and P. J. Morrison, 1994: Normal modes and continuous spectra. Tech. Rep. DOE/ET/53088–686, Texas University, 26 pp. https://inis.iaea.org/Search/search.aspx?orig_q=RN:26051560.

Balmforth, N. J., and P. J. Morrison, 1995: Singular eigenfunctions for shearing fluids I. Tech. Rep. DOE/ET/53088–692, Texas University, 79 pp. http://inis.iaea.org/Search/search.aspx?orig_q=RN:26061992.

Bretherton, F. P., 1966: Critical layer instability in baroclinic flows. *Quart. J. Roy. Meteor. Soc.*, **92**, 325–334, https://doi.org/10.1002/qj.49709239302.

Brink, K. H., and J. Pedlosky, 2019: The Structure of Baroclinic Modes in the Presence of Baroclinic Mean Flow. *J. Phys. Oceanogr.*, **50**, 239–253, https://doi.org/10.1175/JPO-D-19-0123.1.

Burns, K. J., G. M. Vasil, J. S. Oishi, D. Lecoanet, and B. P. Brown, 2020: Dedalus: A flexible framework for numerical simulations with spectral methods. *Phys. Rev. Res.*, **2**, 023068, https://doi.org/10.1103/PhysRevResearch.2.023068.

Charney, J. G., 1971: Geostrophic Turbulence. *J. Atmos. Sci.*, **28**, 1087–1095, https://doi.org/10.1175/1520-0469(1971)028<1087:GT>2.0.CO;2.

Charney, J. G., and G. R. Flierl, 1981: Oceanic Analogues of Large-scale Atmospheric Motions. *Evolution of Physical Oceanography*, B. A. Warren, and C. Wunsch, Eds., MIT press, 448–504.

Chelton, D. B., R. A. deSzoeke, M. G. Schlax, K. El Naggar, and N. Siwertz, 1998: Geographical Variability of the First Baroclinic Rossby Radius of Deformation. *J. Phys. Oceanogr.*, **28**, 433–460, https://doi.org/10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2.

de La Lama, M. S., J. H. LaCasce, and H. K. Fuhr, 2016: The vertical structure of ocean eddies. *Dynamics and Statistics of the Climate System*, **1**, dzw001, https://doi.org/10.1093/climsys/dzw001.

Drazin, P. G., D. N. Beaumont, and S. A. Coaker, 1982: On Rossby waves modified by basic shear, and barotropic instability. *J. Fluid Mech.*, **124**, 439–456, https://doi.org/10.1017/S0022112082002572.

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

Ferrari, R., and C. Wunsch, 2010: The distribution of eddy kinetic and potential energies in the global ocean. *Tellus*, **62**, 92–108, https://doi.org/10.1111/j.1600-0870.2009.00432.x.

Flierl, G. R., 1978: Models of vertical structure and the calibration of two-layer models. *Dyn. Atmos. Oceans*, **2**, 341–381, https://doi.org/10.1016/0377-0265(78)90002-7.

Fu, L.-L., and G. R. Flierl, 1980: Nonlinear energy and enstrophy transfers in a realistically stratified ocean. *Dyn. Atmos. Oceans*, **4**, 219–246, https://doi.org/10.1016/0377-0265(80)90029-9.

Gill, A. E., 1982: *Atmosphere-ocean dynamics*. International geophysics series, Acad. Press, San Diego, 662 pp.

Hallberg, R., 1997: Localized Coupling between Surface and Bottom-Intensified Flow over Topography. *J. Phys. Oceanogr.*, **27**, 977–998, https://doi.org/10.1175/1520-0485(1997)027<0977:LCBSAB>2.0.CO;2.

Held, I. M., R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson, 1995: Surface quasi-geostrophic dynamics. *J. Fluid Mech.*, **282**, 1–20, https://doi.org/10.1017/S0022112095000012.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946, https://doi.org/10.1002/qj.49711147002.

Kelly, S. M., 2016: The Vertical Mode Decomposition of Surface and Internal Tides in the Presence of a Free Surface and Arbitrary Topography. *J. Phys. Oceanogr.*, **46**, 3777–3788, https://doi.org/10.1175/JPO-D-16-0131.1.

Killworth, P. D., and D. L. T. Anderson, 1977: Meaningless modes? *Mode Hot-Line News*, **(72)**, unpublished manuscript.

LaCasce, J. H., 2017: The Prevalence of Oceanic Surface Modes. *Geophys. Res. Lett.*, **44**, 11,097–11,105, https://doi.org/10.1002/2017GL075430.

Lapeyre, G., 2009: What Vertical Mode Does the Altimeter Reflect? On the Decomposition in Baroclinic Modes and on a Surface-Trapped Mode. *J. Phys. Oceanogr.*, **39**, 2857–2874, https://doi.org/10.1175/2009JPO3968.1.

Lapeyre, G., 2017: Surface Quasi-Geostrophy. *Fluids*, **2**, 7, https://doi.org/10.3390/fluids2010007.

Pedlosky, J., 1982: *Geophysical Fluid Dynamics*. 2nd ed., Springer, 624 pp.

Rhines, P. B., 1970: Edge-, bottom-, and Rossby waves in a rotating stratified fluid. *Geophys. Astrophys. Fluid Dyn.*, **1**, 273–302, https://doi.org/10.1080/03091927009365776.

Rocha, C. B., W. R. Young, and I. Grooms, 2015: On Galerkin Approximations of the Surface Active Quasigeostrophic Equations. *J. Phys. Oceanogr.*, **46**, 125–139, https://doi.org/10.1175/JPO-D-15-0073.1.

Roullet, G., J. C. McWilliams, X. Capet, and M. J. Molemaker, 2012: Properties of Steady Geostrophic Turbulence with Isopycnal Outcropping. *J. Phys. Oceanogr.*, **42**, 18–38, https://doi.org/10.1175/JPO-D-11-09.1.

Schneider, T., I. M. Held, and S. T. Garner, 2003: Boundary Effects in Potential Vorticity Dynamics. *J. Atmos. Sci.*, **60**, 1024–1040, https://doi.org/10.1175/1520-0469(2003)60<1024:BEIPVD>2.0.CO;2.

Scott, R. B., and D. G. Furnival, 2012: Assessment of Traditional and New Eigenfunction Bases Applied to Extrapolation of Surface Geostrophic Current Time Series to Below the Surface in an Idealized Primitive Equation Simulation. *J. Phys. Oceanogr.*, **42**, 165–178, https://doi.org/10.1175/2011JPO4523.1.

Smith, K. S., and G. K. Vallis, 2001: The Scales and Equilibration of Midocean Eddies: Freely Evolving Flow. *J. Phys. Oceanogr.*, **31**, 554–571, https://doi.org/10.1175/1520-0485(2001)031<0554:TSAEOM>2.0.CO;2.

Smith, K. S., and J. Vanneste, 2012: A Surface-Aware Projection Basis for Quasigeostrophic Flow. *J. Phys. Oceanogr.*, **43**, 548–562, https://doi.org/10.1175/JPO-D-12-0107.1.

Straub, D. N., 1994: Dispersive effects of zonally varying topography on quasigeostrophic Rossby waves. *Geophys. Astrophys. Fluid Dyn.*, **75**, 107–130, https://doi.org/10.1080/03091929408203650.

Tulloch, R., and K. S. Smith, 2009: Quasigeostrophic Turbulence with Explicit Surface Dynamics: Application to the Atmospheric Energy Spectrum. *J. Atmos. Sci.*, **66**, 450–467, https://doi.org/10.1175/2008JAS2653.1.

Vallis, G. K., 2017: *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*. 2nd ed., Cambridge University Press, 946 pp.

Venaille, A., G. K. Vallis, and S. M. Griffies, 2012: The catalytic role of the beta effect in barotropization processes. *J. Fluid Mech.*, **709**, 490–515, https://doi.org/10.1017/jfm.2012.344.

Wunsch, C., 1997: The Vertical Partition of Oceanic Horizontal Kinetic Energy. *J. Phys. Oceanogr.*, **27**, 1770–1794, https://doi.org/10.1175/1520-0485(1997)027<1770:TVPOOH>2.0.CO;2.

Yassin, H., in prep.: Normal Modes with Boundary Dynamics. *J. Fluid Mech.*