On energy exchanges between eddies and the mean flow in quasigeostrophic turbulence

William Barham and Ian Grooms†

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA

(Received xx; revised xx; accepted xx)

We study the term in the eddy energy budget of continuously-stratified quasigeostrophic turbulence that is responsible for energy extraction by eddies from the background mean flow. This term is a quadratic form, and we derive Euler-Lagrange equations describing its eigenfunctions and eigenvalues, the former being orthogonal in the energy inner product and the latter being real. The eigenvalues correspond to the instantaneous energy growth rate of the associated eigenfunction. We find analytical solutions in the Eady problem. We formulate a spectral method for computing eigenfunctions and eigenvalues, and compute solutions in Phillips-type and Charney-type problems. In all problems, instantaneous growth is possible at all horizontal scales in both inviscid problems and in problems with linear Ekman friction. We conjecture that transient growth at small scales is matched by linear transfer to decaying modes with the same horizontal structure, and we provide simulations supporting the plausibility of this hypothesis. In Charney-type problems, where the linear problem has exponentially growing modes at small scales, we expect net energy extraction from the mean flow to be unavoidable, with an associated nonlinear transfer of energy to dissipation.

Key words: Authors should not enter keywords on the manuscript, as these must be chosen by the author during the online submission process and will then be added during the typesetting process (see http://journals.cambridge.org/data/relatedlink/jfm-keywords.pdf for the full list)

1. Introduction

The conversion of large-scale available potential energy to mesoscale eddy energy by baroclinic instability is one of the most important energy pathways in the global ocean (Ferrari & Wunsch 2009). Linear baroclinic instability is a well-studied subject with roots in the work of Charney (1947), Eady (1949), and Phillips (1954), but the connection of linear theory to actual ocean dynamics is tenuous at best: Global studies of linear baroclinic instability in the world ocean (Smith 2007; Tulloch *et al.* 2011) show that the ocean is very far above the onset of linear stability nearly everywhere, and the actual perturbations about the background state are not infinitesimally small, as in linear theory. Nevertheless linear theory is frequently invoked in an effort to at least partially explain the properties of the fully nonlinear turbulent ocean dynamics (e.g. Tulloch *et al.* 2011; Roullet *et al.* 2012; Grooms 2015; Capet *et al.* 2016; Barham *et al.* 2018, inter alia).

A direct connection between energy conversion in fully nonlinear fluid dynamics and non-modal linear stability theory was made by DelSole (2004). The connection is based

† Email address for correspondence: ian.grooms@colorado.edu

on the fact that the mathematical form of the term in the eddy energy budget that corresponds to extraction of energy from a steady mean flow is exactly the same in both the linear stability problem and in the fully nonlinear dynamics (see also the discussion by Barham & Grooms 2019). This term in the eddy (or perturbation) energy budget is a quadratic form in the eddy variables and can therefore be diagonalized by an orthonormal basis of eigenfunctions (Goldstein 1980). The eigenvalues of the quadratic form are all real; they are the instantaneous growth rates of each eigenfunction's energy. The instantaneous optimal is the eigenfunction associated with the maximal eigenvalue. Nonlinear stability analysis (also called global stability and energy stability) examines the eigenvalues of this quadratic form; the background is energy stable when all the eigenvalues are negative (Joseph 1976). When the linear stability problem is normal, the eigenvalues of the quadratic form are simply the real part of the eigenvalues in the linear stability problem. In many problems the linear stability problem is non-normal, and the quadratic form can have positive eigenvalues (i.e. growing perturbations) even when the linear problem has no eigenvalues with positive real part. This non-normal behavior is crucial to subcritical instability and transition to turbulence in shear flow (e.g. Farrell 1984, 1985; Böberg & Brösa 1988; Butler & Farrell 1992; Farrell & Ioannou 1996; Schmid & Henningson 2001; Schmid 2007). In fully nonlinear flow, excitation of eigenfunction perturbations associated with positive eigenvalues is required to sustain turbulence against dissipation (DelSole 2004).

Barham & Grooms (2019) recently found exact analytical expressions for instantaneous optimals and their growth rates in the classical Eady baroclinic instability problem using the hydrostatic approximation without the geostrophic approximation. The results did not have an obvious direct relevance to the phenomenology of geostrophic turbulence because the largest growth rates occurred at small scales and were driven by ageostrophic extraction of kinetic energy from the background mean flow, rather than by geostrophic extraction of large-scale available potential energy as occurs in the world oceans (Ferrari & Wunsch 2009). The type of baroclinic instability in the Eady problem is somewhat different from other kinds more commonly observed in the global oceans (e.g. Charney-type and Phillips-type; see Tulloch *et al.* 2011), but these other types are difficult to analyze without making a geostrophic approximation. This limitation motivates the present study. We analyze instantaneous optimals and their growth rates in the continuously-stratified quasigeostrophic (QG) framework. Previous research in this area includes the work of Farrell (1989) and Farrell & Ioannou (1996), who studied instantaneous optimals for the QG Eady problem. Rivière *et al.* (2001) studied non-normal growth in the two-layer QG problem, i.e. where the vertical coordinate is discretized into two levels. Both Farrell (1989) and Farrell & Ioannou (1996) started by discretizing the vertical coordinate of the linear problem, and then computed optimal perturbations and their growth rates for the discrete problem. More recently Kalashnik & Chkhetiani (2018) studied transient growth in the QG Eady problem analytically, focusing on cases with zero potential vorticity or idealized potential vorticity configurations.

We directly analyze the continuously-stratified QG equations without discretizing the vertical coordinate; we allow arbitrary shear and stratification profiles, including Charney-type and Phillips-type shear; and unlike Kalashnik & Chkhetiani (2018) we allow the perturbations to have arbitrary potential vorticity configurations. We begin by deriving a linear differential eigenvalue problem describing the optimal perturbations and their instantaneous growth rates, and then give closed-form analytical expressions for the instantaneous optimals and their growth rates in the QG Eady problem. We next find that a simple finite-difference approximation of the linear eigenvalue problem describing the optimal perturbations gives unreliable results, and develop a robust spectral method

that we use to compute approximate numerical solutions. We find that applying a discretization to the linear QG problem and then computing instantaneous optimals of the discrete problem (as in the work of Farrell (1989) and Farrell & Ioannou (1996)) gives reliable results comparable to the results obtained by directly analyzing the continuous problem, but with slower convergence.

The behavior in all three problems (Eady, Charney-type, and Phillips-type) is remarkably similar: instantaneous energy growth is possible at all scales, even when linear Ekman friction is included, with growth rates that are nearly independent of the horizontal scale of the perturbation. This is a marked contrast to the non-QG results of Barham & Grooms (2019), where the growth rate increases as the horizontal scale decreases. There is still a disconnect with the theory and simulations of quasigeostrophic turbulence, where energy extraction from the steady mean flow is not observed at small horizontal scales (Smith & Vallis 2002; Roullet *et al.* 2012; Capet *et al.* 2016). We conjecture that net, or sustained energy growth is not observed at small scales because of linear energy transfer from growing perturbations to decaying ones within a fixed horizontal wavenumber. This is only possible when the linear stability problem has no exponentially growing modes at small scales, and wavenumbers where the linear problem has exponentially growing modes must experience a net energy injection. This conjecture makes a connection back to the linear stability problem without needing to make any of the usual assumptions associated with the linear stability problem, and is consistent with the behavior observed in simulations (Smith & Vallis 2002; Roullet *et al.* 2012; Capet *et al.* 2016). It also predicts that Charney-type shear profiles with exponential growth at small scales should see net extraction of energy from the background mean flow by the eddies at small scales along with an associated nonlinear transfer to other scales for eventual dissipation.

The eigenvalue problem describing the instantaneous optimals is derived in §2, along with exact solutions in the case of the Eady problem. Numerical methods and their results are presented in §3. Our conjecture reconciling linear and nonlinear theory is presented in §4, and conclusions are given in §5.

2. Euler-Lagrange Equations and Exact Solutions

The QG equations with general shear and stratification in z coordinates can be written

$$\partial_t q + \mathbf{u}\cdot\boldsymbol{\nabla} q + \bar{u}(z)\partial_x q + v\left(\beta + \partial_y \bar{q}(z)\right) = 0 \tag{2.1a}$$

$$q = \nabla_h^2 \psi + \frac{\mathrm{d}}{\mathrm{d}z}\left(S(z)\frac{\mathrm{d}\psi}{\mathrm{d}z}\right) \tag{2.1b}$$

$$\partial_t \vartheta^+ + \mathbf{u}\cdot\boldsymbol{\nabla}\vartheta^+ + \bar{u}\partial_x\vartheta^+ - v^+(\partial_z\bar{u}) = 0 \text{ and } \vartheta^+ = S(z)\partial_z\psi \text{ at } z = H \tag{2.1c}$$

$$\partial_t \vartheta^- + \mathbf{u}\cdot\boldsymbol{\nabla}\vartheta^- + \bar{u}\partial_x\vartheta^- - v(\partial_z\bar{u}) = -r\nabla_h^2\psi \text{ and } \vartheta^- = S(z)\partial_z\psi \text{ at } z = 0 \tag{2.1d}$$

where q is the relative potential vorticity, ψ is the streamfunction, $\mathbf{u} = (u, v)^T = (-\partial_y\psi, \partial_x\psi)^T$ is the velocity, $\bar{u}(z)$ is the background velocity, $\bar{q}$ is the background potential vorticity, and $\vartheta = (f_0/N)^2\mathrm{d}\psi/\mathrm{d}z$. The background planetary vorticity is $f_0 + \beta y$. Superscripts $^+$ and $^-$ denote evaluation at $z = 0$ and $z = H$, respectively, and $S(z) = f_0^2/N^2(z)$ is the ratio of the Coriolis and Brunt-Vaisala frequencies, squared. The Ekman drag coefficient is $r = -f_0 d/2$, where d is the depth of the Ekman layer.

Multiplying (2.1) by $-\psi$ and integrating over the volume yields an equation for the

evolution of energy

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\iiint \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\partial_x\psi\partial_z\psi \mathrm{d}x\mathrm{d}y\mathrm{d}z - r\iint |\boldsymbol{\nabla}\psi^-|^2 \mathrm{d}x\mathrm{d}y \tag{2.2}$$

$$\text{where } E = \frac{1}{2}\iiint |\boldsymbol{\nabla}\psi|^2 + S(z)\left(\partial_z\psi\right)^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{2.3}$$

(Technically the energy is $\rho_0 E$ where ρ_0 is the density of the fluid, but following standard convention we refer to E as the energy.)

Both the energy and the energy tendency are quadratic forms that are incompletely diagonalized with a Fourier basis. Expanding ψ in a Fourier series with terms $\hat{\psi}_{\mathbf{k}}(z,t)e^{\mathrm{i}(k_x x + k_y y)}$ leads to the following formulas for energy and energy tendency

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \sum_{k_x}\sum_{k_y}\left[k_x\int_0^H \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\hat{\psi}^*_{\mathbf{k}}\partial_z\hat{\psi}_{\mathbf{k}}\right\}\mathrm{d}z - rk^2|\hat{\psi}^-_{\mathbf{k}}|^2\right] \tag{2.4}$$

$$\text{where } E = \frac{1}{2}\sum_{k_x}\sum_{k_y}\int_0^H k^2\left|\hat{\psi}_{\mathbf{k}}\right|^2 + S(z)\left|\partial_z\hat{\psi}_{\mathbf{k}}\right|^2 \mathrm{d}z. \tag{2.5}$$

We assume that the Fourier coefficient is zero at wavenumber $(k_x, k_y) = (0,0)$, and assume (for convenience and without loss of generality) that the horizontal domain is a square of width 1. The superscript * denotes a complex conjugate and $k^2 = k_x^2 + k_y^2$.

We will focus attention on perturbations with a single Fourier mode as the horizontal structure, and will seek perturbations $\hat{\psi}_{\mathbf{k}}(z)$ that optimize the energy growth rate over the set of perturbations with unit energy. To that end, we define the following Lagrangian

$$I_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}, \lambda] = \hat{G}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] - \lambda\hat{E}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] \tag{2.6}$$

where

$$\hat{E}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] = \frac{1}{2}\int_0^H k^2\left|\hat{\psi}_{\mathbf{k}}\right|^2 + S(z)\left|\partial_z\hat{\psi}_{\mathbf{k}}\right|^2 \mathrm{d}z \tag{2.7}$$

and

$$\hat{G}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] = k_x\int_0^H \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\hat{\psi}^*_{\mathbf{k}}\partial_z\hat{\psi}_{\mathbf{k}}\right\}\mathrm{d}z - rk^2|\hat{\psi}^-_{\mathbf{k}}|^2. \tag{2.8}$$

To optimize the energy growth rate over the set of perturbations with unit energy we derive Euler-Lagrange equations for the stationary points of the Lagrangian. The methods used in the derivation are standard, and the derivation is presented in Appendix A. The Euler Lagrange equations are

$$-\mathrm{i}k_x\left(\partial_y\bar{q}\right)\hat{\psi}_{\mathbf{k}} + 2\mathrm{i}k_x(S(z)\partial_z\bar{u})\frac{\mathrm{d}}{\mathrm{d}z}\hat{\psi}_{\mathbf{k}} + \lambda\left(k^2\hat{\psi}_{\mathbf{k}} - \frac{\mathrm{d}}{\mathrm{d}z}\left(S(z)\frac{\mathrm{d}\hat{\psi}_{\mathbf{k}}}{\mathrm{d}z}\right)\right) = 0, \tag{2.9}$$

$$\mathrm{i}k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\hat{\psi}_{\mathbf{k}} - \lambda S(z)\frac{\mathrm{d}\hat{\psi}_{\mathbf{k}}}{\mathrm{d}z} - rk^2\hat{\psi}_{\mathbf{k}} = 0, \text{ at } z = 0, \tag{2.10}$$

$$\mathrm{i}k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\hat{\psi}_{\mathbf{k}} - \lambda S(z)\frac{\mathrm{d}\hat{\psi}_{\mathbf{k}}}{\mathrm{d}z} = 0, \text{ at } z = H. \tag{2.11}$$

It is worth noting that the mean potential vorticity gradient that appears in these equations, $\partial_y\bar{q}$, does not include the planetary potential vorticity gradient β. This should not be surprising, since β does not appear either in the energy or in the energy tendency. It is also worth noting that the eigenvalue λ is the instantaneous energy growth rate

associated with its corresponding eigenfunction, which follows from the fact that the Lagrangian is a quadratic. To be precise, one can write the Lagrangian as

$$I_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}, \lambda] = \langle \hat{\psi}_{\mathbf{k}}, \mathcal{G}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}} \rangle - \lambda \langle \hat{\psi}_{\mathbf{k}}, \mathcal{E}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}} \rangle \tag{2.12}$$

where $\mathcal{G}_{\mathbf{k}}$ and $\mathcal{E}_{\mathbf{k}}$ are linear operators and $\langle \cdot, \cdot \rangle$ is the standard L^2 inner product. The Euler-Lagrange equations are formally

$$\mathcal{G}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}} = \lambda \mathcal{E}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}}. \tag{2.13}$$

Taking the inner product of this equation with an eigenfunction $\hat{\psi}_{\mathbf{k}}$ and assuming this eigenfunction has been normalized to unit energy yields

$$\hat{G}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] = \langle \hat{\psi}_{\mathbf{k}}, \mathcal{G}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}} \rangle = \lambda \langle \hat{\psi}_{\mathbf{k}}, \mathcal{E}_{\mathbf{k}} \hat{\psi}_{\mathbf{k}} \rangle = \lambda. \tag{2.14}$$

The inviscid problem, $r = 0$, has three illuminating symmetries. First, the growth rate scales linearly with the mean shear, since the transformation $\bar{u} \mapsto \alpha \bar{u}$, $\lambda \mapsto \alpha \lambda$ leaves the Euler-Lagrange equations unchanged. Second, the transformation $\hat{\psi}_{\mathbf{k}} \mapsto \hat{\psi}^*_{\mathbf{k}}$, $\lambda \mapsto -\lambda$ leaves the Euler-Lagrange equations unchanged, so eigenvalues λ come in positive and negative pairs. Third, symmetry under the transformation $k_x \mapsto -k_x$, $\lambda \mapsto -\lambda$ together with the second symmetry implies that growth rates are symmetric about the $k_y = 0$ axis. In addition, it is clear that if $k_x = 0$, then λ must also be zero.

2.1. *Bounding the growth rate*

This section provides an analytical upper bound on the absolute value of the eigenvalues in the inviscid problem, showing that the QG problem studied here is qualitatively different from the non-geostrophic problem studied by Barham & Grooms (2019) in that the growth rates in the QG problem remain bounded as $k \to \infty$.

$$\left| \hat{G}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] \right| = |k_x| \left| \int_0^H \left(S(z) \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right) \mathrm{Im} \left\{ \hat{\psi}^*_{\mathbf{k}} \partial_z \hat{\psi}_{\mathbf{k}} \right\} \mathrm{d}z \right| \tag{2.15}$$

$$\leqslant |k_x| \int_0^H S(z) \left| \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right| \left| \hat{\psi}^*_{\mathbf{k}} \partial_z \hat{\psi}_{\mathbf{k}} \right| \mathrm{d}z \tag{2.16}$$

$$\leqslant \frac{|k_x|}{2} \int_0^H S(z) \left| \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right| \left(\epsilon^2 |\hat{\psi}|^2 + \epsilon^{-2} |\partial_z \hat{\psi}|^2 \right) \mathrm{d}z \tag{2.17}$$

$$= \frac{|k_x|}{2} \int_0^H S(z) \left| \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right| \left(\frac{k}{\sqrt{S(z)}} |\hat{\psi}|^2 + \frac{\sqrt{S(z)}}{k} |\partial_z \hat{\psi}|^2 \right) \mathrm{d}z \tag{2.18}$$

$$\leqslant \frac{|k_x| \left\| \sqrt{S} \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right\|_\infty}{2k} \int_0^H \left(k^2 |\hat{\psi}|^2 + S(z) |\partial_z \hat{\psi}|^2 \right) \mathrm{d}z \tag{2.19}$$

$$= \frac{|k_x| \left\| \sqrt{S} \frac{\mathrm{d}\bar{u}}{\mathrm{d}z} \right\|_\infty}{k} \tag{2.20}$$

We make use of Young's inequality

$$|ab| \leqslant \frac{1}{2} \left(\epsilon^2 |a|^2 + \epsilon^{-2} |b|^2 \right) \tag{2.21}$$

with $\epsilon^2 = k/\sqrt{S(z)}$. The final equality assumes that $\hat{\psi}_{\mathbf{k}}$ has unit energy. This shows that $|\hat{G}_{\mathbf{k}}|$ is bounded above by $|k_x| \| \sqrt{S} \mathrm{d}\bar{u}/\mathrm{d}z \|_\infty / k$ for $\hat{\psi}_{\mathbf{k}}$ with $E = 1$. This clearly rules out eigenfunctions $\hat{\psi}_{\mathbf{k}}$ with unit energy and eigenvalues larger in magnitude than

$|k_x|\|\sqrt{S}\mathrm{d}\bar{u}/\mathrm{d}z\|_\infty/k$. It follows that the growth rates λ must remain bounded in the limit $k \to \infty$. For the Eady problem with $S = 1$, $\bar{u} = z$ the bound is also sharp, as shown in the next section.

2.2. *Eady Exact Solution*

The Eady problem is defined by constant stratification $N^2(z)$ and constant shear $\mathrm{d}\bar{u}/\mathrm{d}z$. For simplicity of exposition, we will analyze the Euler-Lagrange equations in the Eady problem by setting $\bar{u} = z$, $H = 1$, and $S(z) = 1$, and will consider inviscid dynamics ($r = 0$). In this case the Euler-Lagrange equations become

$$2\mathrm{i}k_x\partial_z\hat{\psi}_{\mathbf{k}} + \lambda\left(k^2\hat{\psi}_{\mathbf{k}} - \partial_z^2\hat{\psi}_{\mathbf{k}}\right) = 0, \tag{2.22}$$

$$\mathrm{i}k_x\hat{\psi}_{\mathbf{k}} - \lambda\partial_z\hat{\psi}_{\mathbf{k}} = 0, \text{ at } z = 0 \ \& \ 1. \tag{2.23}$$

It is straighforward to obtain an analytical solution, which has the following form:

$$\lambda = \pm\frac{k_x}{\sqrt{k^2 + \pi^2 n^2}} \Rightarrow \hat{\psi}_{\mathbf{k}}(z) = e^{\pm\mathrm{i}z\sqrt{k^2+\pi^2n^2}}\cos(n\pi z) \tag{2.24}$$

where n is an integer, and the eigenfunctions have not been normalized to unit energy. The growth rate is maximized by taking $n = 0$. Modes with $k_y = 0$ have higher growth than modes with $k_y \neq 0$, and modes with $k_x = 0$ have no growth. When $k_y = n = 0$ the maximal growth rate is 1, independent of k_x. Note that the vertical length scale of the eigenfunctions reduces linearly with increasing k, i.e. eigenfunctions have a fixed aspect ratio of horizontal to vertical length scales.

In the problem without the geostrophic approximation studied by Barham & Grooms (2019), one set of eigenfunctions with $k_y = 0$ had a very similar structure with growth rate $\lambda = 1$ independent of k_x. In that study, another type of eigenfunction was also present with λ growing linearly in k_x, and with energy growth driven entirely by ageostrophic shear production. Restriction to the QG approximation has clearly eliminated these ageostrophic perturbations.

3. Numerical Methods and Solutions

3.1. *Discretizing the continuous optimization problem*

Exact analytical solutions to the Euler-Lagrange equations (2.9)–(2.11) are not available in general, so, to study optimal perturbations and their growth rates, we discretize the problem. A simple second-order centred finite-difference approximation of (2.9)–(2.11) results in a generalized eigenvalue problem of the form

$$\boldsymbol{A}\mathbf{x} = \lambda\boldsymbol{B}\mathbf{x} \tag{3.1}$$

where the entries of $\mathbf{x}$ are values of $\hat{\psi}_{\mathbf{k}}$ on a grid from $z = 0$ to $z = H$. The matrix $\boldsymbol{B}$ is a discretization of the operator $\left[k^2(\cdot) - \partial_z\left(S(z)\partial_z(\cdot)\right)\right]$ with Neumann boundary conditions. Application of this method with $S(z) = 1$ and $\bar{u}(z) = (z/H)^2$ yields a full set of complex eigenvalues, where the complex parts are non-trivial and do not diminish with increasing resolution. As an example, we computed the eigenvalues for $S(z) = 1$, $H = 1$, $\bar{u}(z) = z^2/2$ at $k_x = 1$, $k_y = 0$ with 128 grid points; these are plotted in the left panel of Fig. 1. The eigenvalues all have a non-trivial complex component. The complex part does decrease as the resolution is increased, but slowly: doubling the resolution to 256 points reduces the size of the spurious complex component by a factor of two. Eigenvalues of the continuous problem must all be real, so it is evident that the finite difference method is unreliable here.

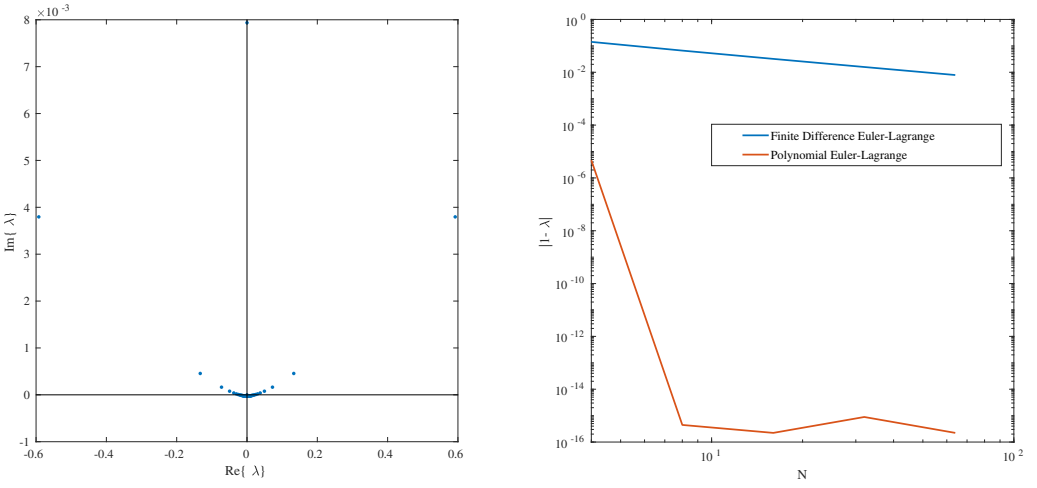

FIGURE 1. Left: Eigenvalues λ computed from a finite-difference discretization of the Euler-Lagrange equations (2.9)–(2.11); the number of grid points is 128 and eigenvalues are plotted in the complex plane. Right: Absolute error in the largest instantaneous growth rate λ in the Eady problem for $k_x = 1$, $k_y = 0$, and $r = 0$ computed by (i) discretizing the continuous QG equations (2.1) and then finding optimals (blue) and (ii) optimizing the Lagrangian of the continuous problem over a subspace of polynomials. In the abscissa for method (i) N refers to the number of grid points, while for method (ii) $N - 1$ is the degree of the polynomials.

Rather than develop an alternate discretization of the Euler-Lagrange equations, we return to the Lagrangian itself (2.6). The Euler-Lagrange equations describe stationary points of the Lagrangian over an infinite-dimensional space (e.g. $\hat{\psi}_{\mathbf{k}} \in C^2([0, H])$), and we discretize the problem by restricting attention to a finite-dimensional subspace, specifically polynomials of degree at most $N - 1$. Finding stationary points of the Lagrangian over this finite-dimensional subspace is straightforward and is guaranteed to produce real eigenvalues λ (which are also Lagrange multipliers); convergence to true values can be monitored by increasing N.

We use the Legendre polynomial basis, $\{P_n(z)\}_{n=0}^{\infty}$, for the space of polynomials on $z \in [0, H]$, so that

$$\hat{\psi}_{\mathbf{k}} = \sum_{n=0}^{N-1} c_n P_n(z). \tag{3.2}$$

Inserting this into $\hat{G}_{\mathbf{k}}$ yields

$$\begin{aligned}\hat{G}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] &= \sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\left[k_x \mathrm{Im}\{c_n^* c_m\}\int_0^H \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right) P_n(z)P_m'(z)\mathrm{d}z - rk^2 c_n^* c_m P_n(0)P_m(0)\right] \\ &= \frac{1}{2}\mathbf{c}^* \boldsymbol{G}\mathbf{c}\end{aligned} \tag{3.3}$$

where $P_n'(z) = \mathrm{d}P_n(z)/\mathrm{d}z$, $\mathbf{c}$ is the vector of coefficients c_n, and

$$\boldsymbol{G}_{n,m} = -2rk^2(-1)^{n+m} + \mathrm{i}k_x \int_0^H \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left(P_m(z)P_n'(z) - P_n(z)P_m'(z)\right)\mathrm{d}z. \tag{3.4}$$

The above uses the Legendre polynomial normalization convention that $P_n(0) = (-1)^n$; we also use the convenient but unconventional indexing for the matrix $\boldsymbol{G}$ starting at 0

rather than 1. Inserting the polynomial into $\hat{E}_{\mathbf{k}}$ yields

$$\begin{aligned}\hat{E}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] &= \frac{1}{2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} c_n^* c_m \int_0^H k^2 P_n(z)P_m(z) + S(z)P_n'(z)P_m'(z)\mathrm{d}z \\ &= \frac{1}{2}\mathbf{c}^*\boldsymbol{E}\mathbf{c} \end{aligned} \tag{3.5}$$

where

$$\boldsymbol{E}_{n,m} = \frac{Hk^2}{2n+1}\delta_{nm} + \int_0^H S(z)P_n'(z)P_m'(z)\mathrm{d}z \tag{3.6}$$

where δ_{nm} is the Kronecker delta, and the indexing again starts at 0 instead of 1. The Lagrangian restricted to this finite-dimensional subspace thus takes the following form

$$\hat{I}_{\mathbf{k}}[\hat{\psi}_{\mathbf{k}}] = \frac{1}{2}\mathbf{c}^*\boldsymbol{G}\mathbf{c} - \lambda\frac{1}{2}\mathbf{c}^*\boldsymbol{E}\mathbf{c}. \tag{3.7}$$

Critical points of the Lagrangian are solutions of the generalized eigenvalue problem

$$\boldsymbol{G}\mathbf{c} = \lambda\boldsymbol{E}\mathbf{c}. \tag{3.8}$$

Both matrices are Hermitian, while the matrix $\boldsymbol{E}$ is positive definite. As a result, a full set of N real eigenvalues is guaranteed to exist.

To validate the spectral discretization described above we compute solutions to the inviscid Eady problem at $k_x = 1$, $k_y = 0$ with $S(z) = 1$, $H = 1$, and $\bar{u}(z) = z$. The absolute error in the maximal eigenvalue is shown in red as a function of N in the right panel of Fig. 1. The error decreases to machine precision by $N = 8$. Results at different wavenumbers are qualitatively similar, though eigenvalues at larger k_x do require higher resolution because they exhibit finer vertical structure as described in the exact analytical solutions.

A natural alternative to the Legendre polynomials is the well-known baroclinic mode basis (Pedlosky 1987, §6.12). Unfortunately these basis functions set $\partial_z\psi = 0$ at the surfaces, which makes them converge much more slowly than polynomials for functions that do not satisfy this condition, including the exact instantaneous optimals of the Eady problem. Furthermore, they are computationally inconvenient, and except for special choices of $N^2(z)$ they are not known analytically.

3.2. *Optimizing the discrete problem*

The foregoing section starts with the continuous equations describing instantaneous optimals and discretizes them with a spectral method. As an alternative one can discretize the linear QG equations (2.1) and then seek instantaneous optimals of the discretized model. The standard second-order finite difference discretization of the continuously-stratified QG system is described in a variety of places including Pedlosky (1987, §6.16 and §6.18) and Vallis (2006, §5.4); the specific details used here are the same as those in Appendix B of Watwood *et al.* (2019). The linear discrete system takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{M}\mathbf{x} = \boldsymbol{K}\mathbf{x} \tag{3.9}$$

where the entries of $\mathbf{x}$ are values of $\hat{\psi}_{\mathbf{k}}$ at the center of N cells from $z = 0$ to $z = H$. The matrix $\boldsymbol{M}$ is real and symmetric (for an equispaced grid). Code to generate the matrices $\boldsymbol{M}$ and $\boldsymbol{K}$ is publicly available (Watwood & Grooms 2019). The energy is $E = H\mathbf{x}^*\boldsymbol{M}\mathbf{x}/(2N)$,

and energy evolution is

$$\frac{\mathrm{d}E}{\mathrm{d}t} = H\mathbf{x}^* \left(\frac{\boldsymbol{K}^* + \boldsymbol{K}}{2N} \right) \mathbf{x}. \tag{3.10}$$

Configurations of $\mathbf{x}$ that optimize the instantaneous growth rate of energy subject to the constraint $E = 1$ are eigenvectors of the generalized eigenvalue problem

$$(\boldsymbol{K}^* + \boldsymbol{K})\,\mathbf{x} = \lambda \boldsymbol{M}\mathbf{x}. \tag{3.11}$$

The discretization is validated by computing solutions to the inviscid Eady problem at $k_x = 1$, $k_y = 0$ with $S(z) = 1$, $H = 1$, and $\bar{u}(z) = z$. The absolute error in the maximal eigenvalue is shown in blue as a function of N in the right panel of Fig. 1. The error decreases algebraically, and far more slowly than the spectral discretization described in the preceeding section: With $N = 64$ the finite difference method is still less accurate than the spectral method with $N = 4$.

3.3. *Numerical Results*

Both of the numerical methods described above were verified to be accurate in the Eady problem. The spectral method is used in all cases with $N = 64$, with exceptions noted. This section describes numerical results for two different configurations of shear. One configuration, the Phillips-type shear, corresponds to two-layer type behavior following on the work of Phillips (1954). The Phillips-type shear consists of a single baroclinic mode with zero shear at the surfaces, and baroclinic instability is related to the fact that the mean potential vorticity gradient $\partial_y \bar{q}$ changes sign once in the interior of the fluid. The second configuration, the Charney-type shear, has a constant mean potential vorticity gradient in the interior of the fluid $\partial_y \bar{q} = -1$ that interacts with shear at the surface to generate instability. The Charney-type setup is associated with surface-intensified nonlinear dynamics and with increased small-scale energy at the upper surface (Capet *et al.* 2016). These two types of shear profile are qualitatively representative of the mean shear profiles that drive baroclinic instability throughout the world oceans (Tulloch *et al.* 2011). In both cases we set $S(z) = 1$ and $H = 1$. Surface-intensified stratification would be more realistic, but the integrals in (3.4) and (3.6) would no longer be analytically tractable. The two shear profiles investigated here are

$$\text{Phillips-type: } \bar{u}(z) = \frac{1}{2}\left(-1 + 6z^2 - 4z^3\right) \tag{3.12}$$

$$\text{Charney-type: } \bar{u}(z) = \frac{z^2}{2}. \tag{3.13}$$

The standard Phillips-type profile (which is a first baroclinic mode) has the form $-\cos(\pi z)/\pi$; a polynomial mean shear was chosen here since it is qualitatively similar to but more convenient numerically than the cosine profile.

The classical Charney problem is posed in a semi-infinite domain, and is identical to the Eady problem except for the inclusion of a planetary vorticity gradient β. Although β is absent from the Lagrangian, a semi-infinite domain constitutes a significant change, and caution should be exercised before extrapolating the current results to the classical Charney problem.

Figure 2 shows numerical results on the instantaneous growth rates λ for all three problems with $r = 0$: Eady (upper row), Phillips-type (middle row), and Charney-type (lower row). The left column shows the most positive eigenvalue λ as a function of wavenumber. The right column shows the five most positive eigenvalues $\lambda_1 > \ldots > \lambda_5$ as

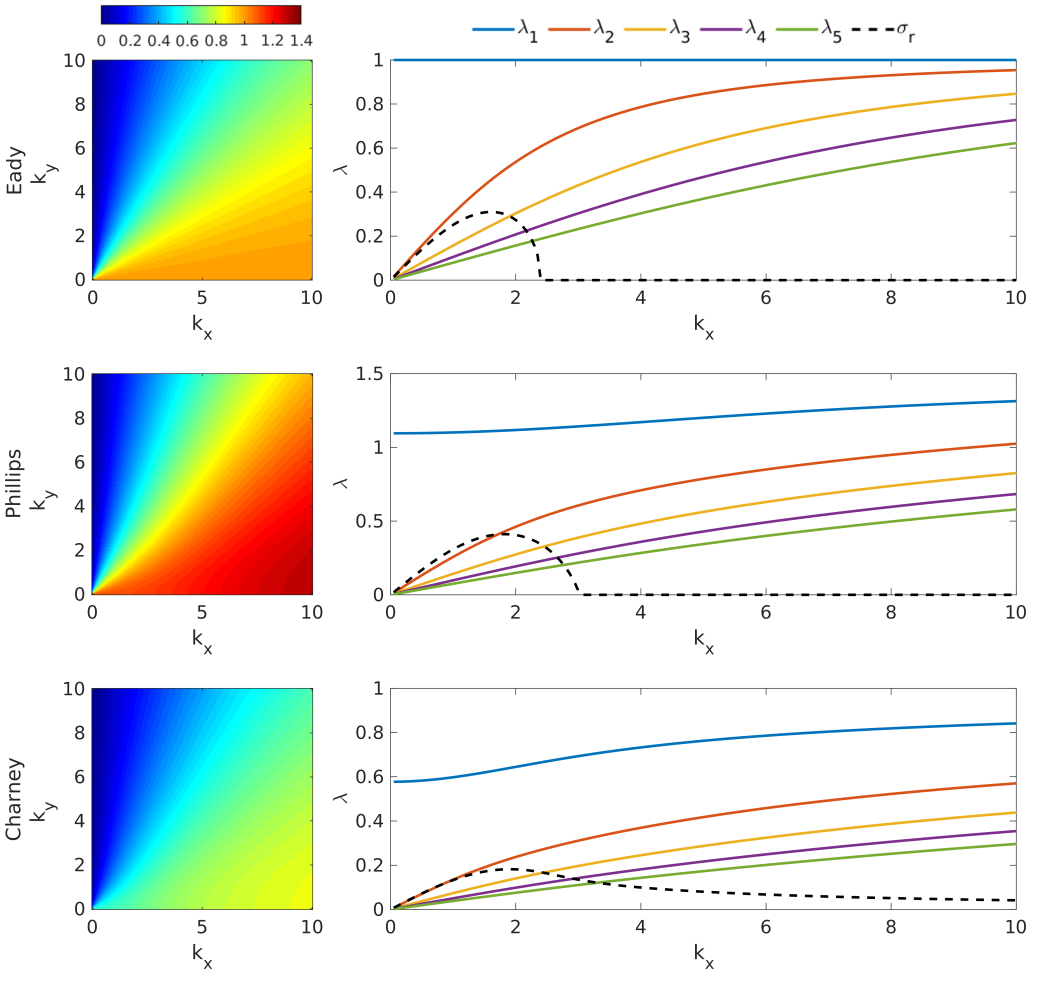

FIGURE 2. Left: Largest eigenvalue λ of the Euler-Lagrange equations as a function of k_x and k_y for the Eady problem (upper), the Phillips-type problem (center), and the Charney-type problem (lower). Right: The five largest eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_5$ (solid; coloured), along with the growth rate σ_r of exponentially-growing solutions for the corresponding linear stability problem (dashed; black). Solutions computed using polynomials of degree $N - 1 = 63$, with $r = 0$ (inviscid).

a function of k_x for $k_y = 0$, together with the growth rate σ_r of the fastest-growing mode of the linear stability problem. The behavior of the three cases is remarkably similar. Growth rates are maximized on the $k_y = 0$ axis, where they are nearly constant. (At larger k_x with $k_y = 0$ the largest growth rates limit towards a constant value; not shown.) At large scales (small k_x) the largest eigenvalue dominates while all the others approach zero.

Figure 3 shows numerical results on the vertical structure of the eigenfunction $\hat{\psi}_{\mathbf{k}}(z)$ associated with the largest eigenvalue at $k_x = 2$, $k_y = 0$ with $r = 0$. The differences between the three problems are more evident here than in their growth rates. The Eady problem has the simple vertical structure proportional to $e^{\mathrm{i}kz}$, as predicted by the analytical solution. In the Phillips-type problem there is no mean shear on the boundary, and the eigenfunctions are required to match this behavior; the energy growth (first term on the right hand side of (2.2)) is maximized in mid-layer, where the mean shear is also

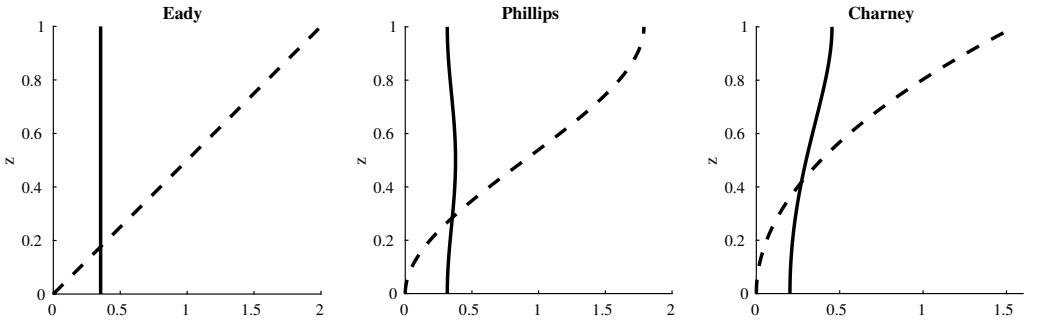

FIGURE 3. Amplitude (solid) and phase (dashed) parts of the optimal perturbations $\hat{\psi}_{\mathbf{k}}$ as a function of z at $k_x = 2$, $k_y = 0$ with $r = 0$ for the Eady (left), Phillips-type (center), and Charney-type (right) problems. Solutions computed with polynomials of degree $N - 1 = 63$.

maximized. In the Charney-type problem there is no mean shear at the lower boundary, and the mean shear is maximized at the upper boundary. The eigenfunctions also have no shear at the lower boundary, and they are surface intensified. In all three problems eigenfunctions corresponding to smaller eigenvalues or to larger wavenumbers k_x have finer vertical structure (not shown).

The effect of linear Ekman friction on the exponentially-growing modes of linear stability theory is markedly different from the simpler intuitive effect of potential vorticity diffusion (Holopainen 1961; Barcilon 1964; Williams & Robinson 1974). To investigate the effect of friction we compute results with $r = 0.5$. Eigenvalues λ as a function of k_x for the Eady problem with $k_y = 0$ are shown in Fig. 4, together with the inviscid results $r = 0$, for comparison. The most positive and most negative eigenvalues for the inviscid Eady problem are $\lambda = 1$ and -1, respectively; these are solid lines in Fig. 4. The most positive and most negative eigenvalues for the Eady problem with $r = 0.5$ are shown as dashed lines. Ekman friction only has a significant impact on the fastest positive growth rate for eigenfunctions with large horizontal scale (small k_x). These growth rates are reduced, though not to zero, presumably because the associated eigenfunctions are nearly barotropic. In contrast, the fastest-decaying eigenvalues are all strongly impacted by Ekman friction. The associated eigenfunctions have a structure that optimizes the decay rate, which increases linearly with k_x. Results for the Phillips-type and Charney-type problems are qualitatively similar (not shown).

4. Discussion

The injection, transfer, and ultimate dissipation of energy is an organizing principle in the study of forced-dissipative turbulence. Energy injection via extraction from the background mean flow in quasigeostrophic turbulence forced by baroclinic shear is associated with a quadratic term in the energy budget. This quadratic term can be diagonalized by a basis of eigenfunctions associated with real eigenvalues. Energy injection in quasigeostrophic turbulence forced by baroclinic shear can therefore be associated with excitation and growth of eigenfunctions associated with positive eigenvalues. These eigenfunctions are related to the instantaneous optimal perturbations of non-normal linear stability theory, and are generically referred to here as instantaneous optimals. But the topic here is strongly nonlinear quasigeostrophic turbulence, which is significantly different from the dynamics of subcritical transition in viscous shear flows where non-normal linear behavior is well-known (Schmid & Henningson 2001).

The partitioning of this energy into kinetic and available potential energy and its

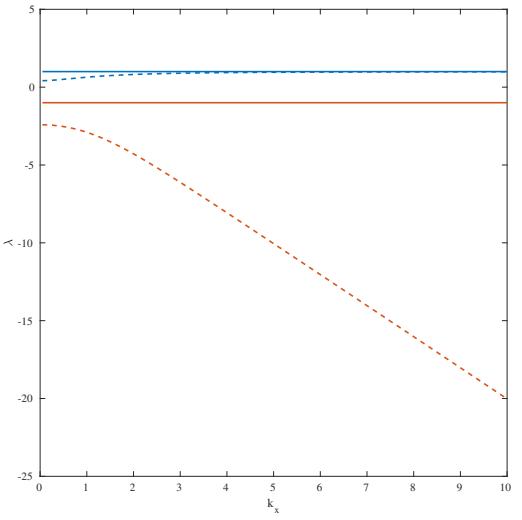

FIGURE 4. Most-positive and most-negative eigenvalues λ for the Eady problem as a function of k_x with $k_y = 0$ for $r = 0$ (solid) and $r = 0.5$ (dashed). The blue lines show the most-positive eigenvalues while the red lines show the most-negative. Solutions computed with polynomials of degree $N - 1 = 63$. Results for the Charney- and Phillips-type problems are similar (not shown).

transfer between forms and between horizontal scales is a complex subject that has been primarily studied in the significantly simpler two-layer model (Salmon 1978; Fu & Flierl 1980; Haidvogel & Held 1980; Scott & Arbic 2007). In both the two-layer model and in models with higher vertical resolution there is a net energy input to the system over a range of scales larger than the deformation radius in the form of eddy available potential energy (see, e.g., Grooms & Majda 2014, Figure 3d), which is also the range of scales where there are exponentially growing modes of the linear stability problem (Smith & Vallis 2002; Roullet *et al.* 2012; Capet *et al.* 2016). The potential energy then cascades downscale; near the deformation radius it is converted to eddy kinetic energy, which then cascades to larger scales where it is dissipated via friction. To calculate the actual rate of change of energy at a given wavenumber due to interaction with the mean flow and friction, one resolves the vertical structure at that wavenumber into a sum of instantaneous optimal eigenfunctions. The total energy growth rate is then obtained by summing over the energy in each eigenfunction multiplied by the growth rate associated with that eigenfunction. Our results are broadly consistent with the simulated phenomenology of baroclinic turbulence in the sense that by multiplying the energy spectrum by nearly flat eigenvalues λ, one obtains an energy injection spectrum concentrated at large scales. This is not completely satisfactory though, because the shape of the energy spectrum presumably depends on the structure of the energy injection and vice versa. Our analysis indicates that energy injection is *possible* at all horizontal scales, so some reconciliation with the observed confinement of energy injection to a limited range of horizontal scales larger than the deformation radius is needed.

To frame the following discussion, write the fluid system in the following generic form

$$\partial_t u = \mathcal{L}[u] + \mathcal{B}[u, u] \tag{4.1}$$

where u generically denotes the flow variables (not the velocity $u = -\partial_x \psi$ of the previous sections), $\mathcal{L}$ is a linear operator including interaction of the flow with the background mean state, and $\mathcal{B}$ is a bilinear nonlinear advection operator. We can further assume the existence of an energy inner product $E = \langle u, u \rangle / 2$, and we can decompose the linear term

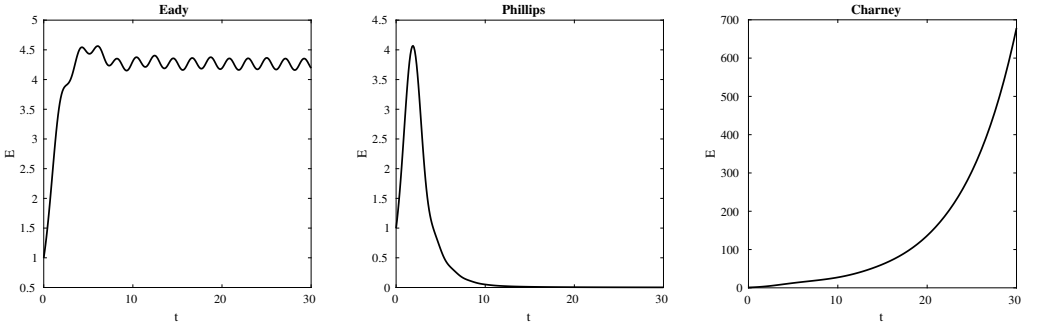

FIGURE 5. Energy as a function of time for simulations at $k_x = 5$, $k_y = 0$, $r = 0$, initialized with the optimal perturbation for the Eady (left), Phillips-type (right), and Charney-type (right) problems.

$\mathcal{L}$ into a piece $\mathcal{L}_S$ that is self-adjoint with respect to the energy inner product, and a piece $\mathcal{L}_A$ that is skew-adjoint with respect to the energy inner product. The eigenvalues and eigenfunctions of the operator $\mathcal{L}_S$ are the growth rates and instantaneous optimals studied in the previous sections.

In the QG problem the flow variable is ψ, the energy inner product is $E = \langle\psi, \psi\rangle/2$ defined by (2.3), and $\langle\psi, \mathcal{L}\psi\rangle$ is $\mathrm{d}E/\mathrm{d}t$ as defined by (2.2). In the QG problem $\mathcal{L}$ has constant coefficients with respect to the horizontal coordinates x and y, and therefore does not transfer energy between horizontal wavenumbers; only the quadratic nonlinearity transfers energy between horizontal wavenumbers via triad interactions. On the other hand, the nonlinearity in the QG equations does not lead to nonlinear self-interaction at a single wavenumber because $\mathbf{u} \cdot \boldsymbol{\nabla} q = 0$ when $\mathbf{u}$ and q are both proportional to $e^{\mathrm{i}(k_x x + k_y y)}$. In a turbulent setting the nonlinear term will guarantee excitation of all modes, so instantaneous energy growth is not only possible at all scales, it is unavoidable. How can this be reconciled with the evident lack of net energy injection at small scales in simulations of quasigeostrophic turbulence? We conjecture that the answer lies in the role of the skew-adjoint linear operator $\mathcal{L}_A$. This operator is not able to transfer energy between wavenumbers, but it can transfer energy from growing eigenfunctions of $\mathcal{L}_S$ to decaying ones. In fact, at wavenumbers where the operator $\mathcal{L}$ has no eigenvalues with positive real part it is possible in principle that there is zero net (or sustained) energy growth because of transfer from growing modes of $\mathcal{L}_S$ to decaying ones by $\mathcal{L}_A$. However, if the operator $\mathcal{L}$ has an eigenvalue with positive real part for some wavenumber then sustained growth at that wavenumber is unavoidable and must be matched by nonlinear transfer to other wavenumbers. This conjecture makes a connection to linear modal stability theory without relying on any assumptions of linearity, and is consistent with the observed lack of energy injection at small horizontal scales in problems like the Eady and Phillips problems where there are no linearly unstable modes at small scales.

To illustrate this mechanism we run simulations of the full QG dynamics using the standard equispaced finite-difference approximation in z, starting from an initial condition that is the optimal perturbation of the finite difference model at wavenumber $k_x = 5$, $k_y = 0$. Simulations are run for all three canonical problems from the preceeding section. We solve the inviscid dynamics ($r = 0$) with $N = 256$ vertical levels, with constant stratification $S(z) = 1$ and with $H = 1$. Since the initial condition consists of a single horizontal wavenumber the nonlinearity is exactly zero for all time, and the dynamics reduce to linear. The time evolution of the energy, starting from unity, is shown in Fig. 5 for all three simulations.

In the Eady problem the energy grows transiently and then sets up a regular oscillation with no further net growth. (With nonzero Ekman friction the energy would eventually decay to zero.) In the Phillips-type problem there is large initial transient growth followed by decay towards zero. In the Charney-type problem the initial condition has a nonzero component in the direction of an linearly unstable mode, and the energy eventually grows exponentially. In the Eady and Phillips-type problems transient growth occurs as the flow extracts energy from the large-scale background, but there is no sustained energy growth. Instead, there is transient decay where the flow returns energy to the background. In the presence of friction some of this transient decay would be associated with friction rather than a return of energy to the background. This behavior, where the flow at small scales transiently extracts energy from the background flow and then either returns it to the background flow or dissipates it frictionally is consistent with the standard theory of QG turbulence that expects no net energy transfer across horizontal scales at scales smaller than the deformation radius (Charney 1971).

Having the small-scale flow transition between extracting energy from the large scales and then returning that energy to the large scales is also connected to the concept of energetic backscatter – transient transfer of energy from small to large scales. Backscatter in idealized ocean models accounts in part for large-scale variability (Kitsios *et al.* 2013), and the development of parameterizations of kinetic energy backscatter for ocean models is a topic of current research (e.g. Kitsios *et al.* 2013; Jansen & Held 2014; Chen *et al.* 2018). The backscatter associated with linear inviscid decay at small scales as seen here is in the form of potential energy while the aforementioned studies address kinetic energy backscatter. The vertical structure of decaying perturbations could perhaps be used to inform parameterizations of potential energy backscatter within the framework of Grooms (2016) and Grooms & Kleiber (2019).

5. Conclusions

The goal of this study is to elucidate the process whereby eddies extract energy from a steady background baroclinic shear in QG turbulence, bridging the gap between linear theory and fully nonlinear dynamics. The term in the eddy energy budget that corresponds to extraction of energy from the background shear is quadratic in the eddy variables and can be diagonalized by an orthonormal basis of eigenfunctions. The associated eigenvalues are the instantaneous energy growth rates associated with each eigenfunction. Any flow configuration can therefore be resolved into a sum of eigenfunctions, each of which is instantaneously extracting energy from the mean or returning energy to the mean. We begin by deriving Euler-Lagrange equations describing the eigenfunctions and eigenvalues in continuously-stratified QG dynamics. An exact analytical solution is found for the inviscid Eady problem showing that instantaneous energy growth is possible at all horizontal wavenumbers, and among perturbations with no variation in the y direction (perpendicular to the background velocity) the fastest-growing ones all have the same growth rate independent of horizontal scale. A spectral numerical method is formulated to compute eigenfunctions and growth rates for vertically-varying shear and stratification profiles. We compute eigenvalues for the Eady problem and for a Phillips-type and a Charney-type problem. The Phillips-type problem has no mean shear at the upper and lower surfaces with a mean potential vorticity gradient that changes sign in the interior, and the Charney-type problem has a constant potential vorticity gradient that interacts with shear at the upper surface. In all experiments the stratification $N^2(z)$ is constant, though the results may extend qualitatively to more realistic surface-intensified stratification. The computed eigenvalues

for Phillips-type and Charney-type background flows are remarkably similar to the Eady results, though the eigenfunctions themselves are different. The amplitude of the growing eigenfunctions is concentrated in regions of high shear; in the Charney-type problem they are surface intensified and in the Phillips-type problem they are intensified in the middle of the layer. Ekman friction reduces growth rates at large scales and also leads to a set of eigenfunctions associated with rapid decay, though it has minimal effect on growth at small scales.

Our analytical results show that instantaneous growth is possible at all scales, unlike modal linear stability theory where growth at small scales only occurs in Charney-type problems where an internal potential vorticity gradient interacts with shear at the boundary. In fully nonlinear turbulence triad interactions guarantee nonzero excitation of all Fourier modes, which means that instantaneous growth is not only possible at all scales, it is unavoidable. This needs to be reconciled with the phenomenology of QG turbulence, where sustained extraction of energy from the background mean flow is not observed at small scales, except in Charney-type problems (Smith & Vallis 2002; Roullet *et al.* 2012; Capet *et al.* 2016). We propose that the transient growth of some eigenfunctions at small scales is matched by transient decay of other eigenfunctions at the same scales, with transfer between them mediated by the skew-adjoint part of the linear operator. The plausibility of this hypothesis is supported by linear simulations where an initial phase of transient growth is either followed by transient decay in the Phillips-type problem, or where it transitions to steady oscillation in the Eady problem. The simulations are inviscid, demonstrating that the mechanism can work without having to match growth by frictional dissipation; the inclusion of frictional dissipation would not change the overall hypothesis though. This hypothesis predicts that Charney-type problems must exhibit sustained energy extraction from the mean flow at all scales, though the rate of extraction at small scales should be weak because of the low energy at small scales. This energy extraction has to be matched by nonlinear transfer to scales where the energy decays, either by frictional processes or by backscatter of energy to the large-scale mean flow.

We are grateful to N. Rodriguez for suggesting to optimize the Lagrangian over a finite-dimensional subspace rather than directly discretize the Euler-Lagrange equations. W. Barham is supported by US National Science Foundation grant DMS 1407340. This work is partially supported by US National Science Foundation grant OCE 1736708.

Declaration of Interests. The authors report no conflict of interest.

Appendix A.

The Euler-Lagrange equations (2.9)–(2.11) are derived in this appendix. Let ϕ represent a perturbation of $\hat{\psi}_{\mathbf{k}}$. Inserting $\hat{\psi}_{\mathbf{k}} + \phi$ into $\hat{I}_{\mathbf{k}}$ and retaining linear terms in ϕ, we obtain:

$$-\frac{\mathrm{i}k_x}{2}\int_0^1 \left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left[\hat{\psi}^*_{\mathbf{k}}\partial_z\phi + \phi^*\partial_z\hat{\psi}_{\mathbf{k}} - \hat{\psi}_{\mathbf{k}}\partial_z\phi^* - \phi\partial_z\hat{\psi}^*_{\mathbf{k}}\right]\mathrm{d}z$$
$$-\frac{\lambda}{2}\int_0^1 k^2\left(\phi\hat{\psi}^*_{\mathbf{k}} + \hat{\psi}_{\mathbf{k}}\phi^*\right) + S(z)\left[\partial_z\phi\partial_z\hat{\psi}^*_{\mathbf{k}} + \partial_z\hat{\psi}_{\mathbf{k}}\partial_z\phi^*\right]\mathrm{d}z - rk^2((\phi^-)^*\hat{\psi}^-_{\mathbf{k}} + \phi^-(\hat{\psi}^-_{\mathbf{k}})^*) \tag{A 1}$$

Simplifying with integration by parts, we obtain:

$$\begin{aligned}\left[-\frac{\mathrm{i}k_x}{2}\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left[\hat{\psi}^*_{\mathbf{k}}\phi-\hat{\psi}_{\mathbf{k}}\phi^*\right]\right]\Big|_{z=0}^{H}&\\ +\frac{\mathrm{i}k_x}{2}\int_0^1\partial_y\bar{q}\left[\hat{\psi}_{\mathbf{k}}\phi^*-\hat{\psi}^*_{\mathbf{k}}\phi\right]+2\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left[\phi^*\partial_z\hat{\psi}_{\mathbf{k}}-\phi\partial_z\hat{\psi}^*_{\mathbf{k}}\right]\mathrm{d}z&\\ -\frac{\lambda}{2}\left[S(z)\left(\phi\partial_z\hat{\psi}^*_{\mathbf{k}}+\phi^*\partial_z\hat{\psi}_{\mathbf{k}}\right)\right]\Big|_{z=0}^{H}&\\ -\frac{\lambda}{2}\int_0^1 k^2\left(\phi\hat{\psi}^*_{\mathbf{k}}+\hat{\psi}_{\mathbf{k}}\phi^*\right)-\phi\partial_z\left(S(z)\partial_z\hat{\psi}^*_{\mathbf{k}}\right)-\phi^*\partial_z\left(S(z)\partial_z\hat{\psi}_{\mathbf{k}}\right)\mathrm{d}z&\\ -rk^2((\phi^-)^*\hat{\psi}^-_{\mathbf{k}}+\phi^-(\hat{\psi}^-_{\mathbf{k}})^*).&\end{aligned}\tag{A 2}$$

The above uses the following identity relating the mean shear to the mean potential vorticity gradient

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)=-\partial_y\bar{q}.\tag{A 3}$$

This simplifies to

$$\begin{aligned}\left[k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\hat{\psi}^*_{\mathbf{k}}\phi\right\}-\lambda S(z)\mathrm{Re}\left\{\phi\partial_z\hat{\psi}^*_{\mathbf{k}}\right\}\right]\Big|_{z=0}^{1}-2rk^2\mathrm{Re}\left\{(\phi^-)^*\hat{\psi}^-_{\mathbf{k}}\right\}&\\ -k_x\int_0^1\partial_y\bar{q}\,\mathrm{Im}\left\{\hat{\psi}_{\mathbf{k}}\phi^*\right\}+2\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\phi^*\partial_z\hat{\psi}_{\mathbf{k}}\right\}\mathrm{d}z&\\ -\lambda\int_0^1\mathrm{Re}\left\{\phi\left(k^2\psi^*_{\mathbf{k}}-(\partial_z S\partial_z\psi^*_{\mathbf{k}})\right)\right\}\mathrm{d}z.&\end{aligned}\tag{A 4}$$

Stationary configurations of $\hat{\psi}_{\mathbf{k}}$ are defined to be those for which the above expression is zero for every ϕ, in particular for ϕ that are zero on the boundaries, which means that the integral and the boundary terms must vanish separately. Thus we have

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\hat{\psi}^*_{\mathbf{k}}\phi\right\}-\lambda S(z)\mathrm{Re}\left\{\phi\partial_z\hat{\psi}^*_{\mathbf{k}}\right\}-2rk^2\mathrm{Re}\left\{(\phi^-)^*\hat{\psi}^-_{\mathbf{k}}\right\}=0\text{ at }z=0,\tag{A 5}$$

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\hat{\psi}^*_{\mathbf{k}}\phi\right\}-\lambda S(z)\mathrm{Re}\left\{\phi\partial_z\hat{\psi}^*_{\mathbf{k}}\right\}=0\text{ at }z=H\tag{A 6}$$

and

$$\begin{aligned}-k_x\partial_y\bar{q}\,\mathrm{Im}\left\{\hat{\psi}_{\mathbf{k}}\phi^*\right\}+2\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\mathrm{Im}\left\{\phi^*\partial_z\hat{\psi}_{\mathbf{k}}\right\}&\\ -\lambda\mathrm{Re}\left\{\phi\left(k^2\psi^*_{\mathbf{k}}-\partial_z\left(S(z)\partial_z\psi^*_{\mathbf{k}}\right)\right)\right\}=0&\end{aligned}\tag{A 7}$$

for $z\in(0,H)$. Let $\hat{\psi}_{\mathbf{k}}=\psi_r+\mathrm{i}\psi_i$ and $\phi=\phi_r+\mathrm{i}\phi_i$ be the complex expression of $\hat{\psi}_{\mathbf{k}}$ and ϕ. Then we may write (A 7) as

$$-k_x\partial_y\bar{q}\left(\phi_r\psi_i-\phi_i\psi_r\right)+2k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left(\phi_r\partial_z\psi_i-\phi_i\partial_z\psi_r\right)\tag{A 8}$$

$$-\lambda\left[\phi_r\left(k^2\psi_r-\partial_z\left(S(z)\partial_z\psi_r\right)\right)\right]-\lambda\left[\phi_i\left(k^2\psi_i-\partial_z\left(S(z)\partial_z\psi_i\right)\right)\right]=0.\tag{A 9}$$

Again, since this has to be true for any ϕ, including cases where $\phi_r=0$ or $\phi_i=0$. From

this follows the system of equations

$$-k_x\partial_y\bar{q}\,\psi_i + 2k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\partial_z\psi_i - \lambda\left(k^2\psi_r - \partial_z\left(S(z)\partial_z\psi_r\right)\right) = 0 \tag{A 10}$$

$$k_x\partial_y\bar{q}\,\psi_r - 2k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\partial_z\psi_r - \lambda\left(k^2\psi_i - \partial_z\left(S(z)\partial_z\psi_i\right)\right) = 0. \tag{A 11}$$

Multiplying the second equation by i and adding, we get

$$-\mathrm{i}k_x\partial_y\bar{q}\,\hat{\psi}_{\mathbf{k}} + 2\mathrm{i}k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\partial_z\hat{\psi}_{\mathbf{k}} + \lambda\left(k^2\hat{\psi}_{\mathbf{k}} - \partial_z\left(S(z)\partial_z\hat{\psi}_{\mathbf{k}}\right)\right) = 0, \text{ for } z\in(0,H). \tag{A 12}$$

This is (2.9). We next return to the boundary conditions. Again splitting into real and imaginary parts, we get

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left[\psi_r\phi_i - \psi_i\phi_r\right] - \lambda S(z)\left[\phi_r\partial_z\psi_r + \phi_i\partial_z\psi_i\right] - rk^2(\phi_r\psi_r + \phi_i\psi_i) = 0, \text{ for } z = 0. \tag{A 13}$$

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\left[\psi_r\phi_i - \psi_i\phi_r\right] - \lambda S(z)\left[\phi_r\partial_z\psi_r + \phi_i\partial_z\psi_i\right] = 0, \text{ for } z = H. \tag{A 14}$$

Alternately setting $\phi_r = 0$ and $\phi_i = 0$ gives rise to the equations

$$-k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\psi_i - \lambda S(z)\partial_z\psi_r = 0 \text{ and} \tag{A 15}$$

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\psi_r - \lambda S(z)\partial_z\psi_i = 0, \text{ for } z = H. \tag{A 16}$$

Multiplying the second equation by i and adding, we get (2.11)

$$\mathrm{i}k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\hat{\psi}_{\mathbf{k}} - \lambda S(z)\partial_z\hat{\psi}_{\mathbf{k}} = 0, \text{ for } z = H. \tag{A 17}$$

The same process at $z = 0$ gives rise to the equations

$$-k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\psi_i - \lambda S(z)\partial_z\psi_r - rk^2\psi_r = 0 \text{ and} \tag{A 18}$$

$$k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\psi_r - \lambda S(z)\partial_z\psi_i - rk^2\psi_i = 0, \text{ for } z = 0. \tag{A 19}$$

Multiplying the second equation by i and adding, we get (2.10)

$$\mathrm{i}k_x\left(S(z)\frac{\mathrm{d}\bar{u}}{\mathrm{d}z}\right)\hat{\psi}_{\mathbf{k}} - \lambda S(z)\partial_z\hat{\psi}_{\mathbf{k}} - 2rk^2\hat{\psi}_{\mathbf{k}} = 0, \text{ for } z = 0. \tag{A 20}$$

REFERENCES

Barcilon, V. 1964 Role of the Ekman layers in the stability of the symmetric regime obtained in a rotating annulus. *J. Atmos. Sci.* **21** (3), 291–299.

Barham, W., Bachman, S. & Grooms, I. 2018 Some effects of horizontal discretization on linear baroclinic and symmetric instabilities. *Ocean Model.* **125**, 106–116.

Barham, W. & Grooms, I. 2019 Exact instantaneous optimals in the non-geostrophic Eady problem and the detrimental effects of discretization. *Theor. Comp. Fluid Dyn.* **33** (2), 125–139.

Böberg, L. & Brösa, U. 1988 Onset of turbulence in a pipe. *Z. Naturforsch.* **43** (8-9), 697–726.

Butler, K. M. & Farrell, B. F. 1992 Three-dimensional optimal perturbations in viscous shear flow. *Phys. Fluids A* **4**, 1637–1650.

Capet, X., Roullet, G., Klein, P. & Maze, G. 2016 Intensification of upper-ocean submesoscale turbulence through charney baroclinic instability. *J. Phys. Ocean.* **46** (11), 3365–3384.

Charney, J. G. 1947 The dynamics of long waves in a baroclinic westerly current. *J. Meteorology* **4** (5), 136–162.

Charney, J. G. 1971 Geostrophic turbulence. *J. Atmos. Sci.* **28** (6), 1087–1095.

Chen, A., Barham, W. & Grooms, I. 2018 Comparing eddy-permitting ocean model parameterizations via Lagrangian particle statistics in a quasigeostrophic setting. *J. Geophys. Res.-Oceans* **123** (8), 5637–5651.

DelSole, T. 2004 The necessity of instantaneous optimals in stationary turbulence. *J. Atmos. Sci.* **61** (9), 1086–1091.

Eady, E. T. 1949 Long waves and cyclone waves. *Tellus* **1** (3), 33–52.

Farrell, B. 1984 Modal and non-modal baroclinic waves. *J. Atmos. Sci.* **41** (4), 668–673.

Farrell, B. 1985 Transient growth of damped baroclinic waves. *J. Atmos. Sci.* **42** (24), 2718–2727.

Farrell, B. F. 1989 Optimal excitation of baroclinic waves. *J. Atmos. Sci.* **46** (9), 1193–1206.

Farrell, B. F. & Ioannou, P. J. 1996 Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.* **53** (14), 2025–2040.

Ferrari, R. & Wunsch, C. 2009 Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annu. Rev. Fluid Mech.* **41**.

Fu, L.-L. & Flierl, G. R. 1980 Nonlinear energy and enstrophy transfers in a realistically stratified ocean. *Dyn. Atmos. Oceans* **4** (4), 219–246.

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

Grooms, I. 2015 Submesoscale baroclinic instability in the Balance Equations. *J. Fluid Mech.* **762**, 256–272.

Grooms, I 2016 A Gaussian-product stochastic Gent-McWilliams parameterization. *Ocean Model.* **106**, 27–43.

Grooms, I. & Kleiber, W. 2019 Diagnosing, modeling, and testing a multiplicative stochastic Gent-McWilliams parameterization. *Ocean Model.* **133**, 1–10.

Grooms, I. & Majda, A. J. 2014 Stochastic superparameterization in quasigeostrophic turbulence. *J. Comput. Phys.* **271**, 78–98.

Haidvogel, D. B. & Held, I. M. 1980 Homogeneous quasi-geostrophic turbulence driven by a uniform temperature gradient. *J. Atmos. Sci.* **37** (12), 2644–2660.

Holopainen, E.O. 1961 On the effect of friction in baroclinic waves. *Tellus* **13**, 363–367.

Jansen, M. F. & Held, I. M. 2014 Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Model.* **80**, 36–48.

Joseph, D.D. 1976 *Stability of Fluid Motions I*. Springer-Verlag.

Kalashnik, M. V. & Chkhetiani, O. 2018 An analytical approach to the determination of optimal perturbations in the Eady model. *J. Atmos. Sci.* **75**, 2741–2761.

Kitsios, V., Frederiksen, J. S. & Zidikheri, M. J. 2013 Scaling laws for parameterisations of subgrid eddy–eddy interactions in simulations of oceanic circulations. *Ocean Model.* **68**, 88–105.

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

Phillips, N. A. 1954 Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus* **6** (3), 274–286.

Rivière, G., Hua, B. L. & Klein, P. 2001 Influence of the β-effect on non-modal baroclinic instability. *Q. J. Roy. Meteor. Soc.* **127** (574), 1375–1388.

Roullet, G., McWilliams, J. C., Capet, X. & Molemaker, M. J. 2012 Properties of steady geostrophic turbulence with isopycnal outcropping. *J. Phys. Ocean.* **42** (1), 18–38.

Salmon, R. 1978 Two-layer quasi-geostrophic turbulence in a simple special case. *Geophys. Astro. Fluid* **10** (1), 25–52.

Schmid, P. J. 2007 Nonmodal stability theory. *Annu. Rev. Fluid Mech.* **39**, 129–162.

Schmid, P. J. & Henningson, D. S. 2001 *Stability and transition in shear flows*. Springer.

Scott, R. B. & Arbic, B. K. 2007 Spectral energy fluxes in geostrophic turbulence: Implications for ocean energetics. *J. Phys. Ocean.* **37** (3), 673–688.

Smith, K. S. 2007 The geography of linear baroclinic instability in Earth's oceans. *J. Marine Res.* **65** (5), 655–683.

Smith, K. S. & Vallis, G. K. 2002 The scales and equilibration of midocean eddies: Forced–dissipative flow. *J. Phys. Ocean.* **32** (6), 1699–1720.

Tulloch, R., Marshall, J., Hill, C. & Smith, K. S. 2011 Scales, growth rates, and spectral fluxes of baroclinic instability in the ocean. *J. Phys. Ocean.* **41** (6), 1057–1076.

Vallis, G. K. 2006 *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press.

Watwood, M. & Grooms, I. 2019 QG-Galerkin. `https://github.com/mwatwood-cu/QG-Galerkin`.

Watwood, M., Grooms, I., Julien, K.A. & Smith, K. S. 2019 Energy-conserving Galerkin approximations for quasigeostrophic dynamics. *J. Comput. Phys.* **388**, 23–40.

Williams, G.P. & Robinson, J.B. 1974 Generalized Eady waves with Ekman pumping. *J. Atmos. Sci.* **31** (7), 1768–1776.