The geography of linear baroclinic instability in Earth's oceans

by K. Shafer Smith[1]

ABSTRACT

Satellite observations reveal a mesoscale oceanic circulation dominated by turbulence that is correlated, in most cases, with local baroclinicity. Linear baroclinic instability theory has proved useful in understanding the time and space scales of atmospheric eddies. The question addressed here is, to what degree can the observed oceanic eddy activity be understood through a local, linear stability analysis? This question is addressed as follows. A local quasigeostrophic linear stability calculation is performed on a grid of wavenumbers, ranging in magnitude about the local deformation wavenumber, for each vertical profile in a dataset of neutral density for the world's oceans. The initial results show that nearly the entire ocean is unstable, but in many places, particularly in low latitudes, the instability is dominated by surface intensified modes, resulting in very small-scale, quickly growing waves. At higher latitudes, the primary instabilities are due to thermocline depth shears and have a broader vertical structure. For each unstable wave, at each location, the mean-to-eddy energy conversion rate is also calculated and used to select the growing waves that are both fast and have significant energetic conversion potential. This procedure removes most of the surface-instabilities, which cannot lead to significant energy conversion, and reveals the slower but more powerful thermocline-level instabilities where they exist.

The time and space scales of these growing waves are compared to estimates of the Eady growth rate and deformation scale, respectively. It is found that while the timescale is well-approximated by the Eady-estimate, the spatial scales are uniformly smaller than the deformation scale, typically by a factor of 4. The zonally averaged spatial scales are then compared to observed eddy scales. The spatial scales of maximum growth are everywhere significantly smaller than the observed eddy scales. In the Antarctic Circumpolar Current, for example, the scale of maximum growth is about 5 km, much smaller than the observed eddy scales, estimates of which range from 30–100 km. A possible, and unsurprising conclusion is that the observed eddy scales are the result of an inverse cascade, and cannot be understood by linear theory alone.

1. Introduction

Satellite altimetric observations of the ocean surface reveal a circulation dominated by turbulent flow on scales of 50 to 250 km, the oceanic mesoscale. The apparent eddy dominance at the oceanic mesoscale was first noted, however, from ship-going observations in the early 1970s. Using available data, Gill, Green, and Simmons (1974, hereafter GGS)

---

1. Courant Institute of Mathematical Sciences and Center for Atmosphere Ocean Science, New York University, New York, New York, 10012, U.S.A. *email: shafer@cims.nyu.edu*

noted the concurrence of eddy activity with steep isopycnal gradients, and pointed out that the oceanic mean available potential energy resulting from these isopycnal gradients is about 1000 times larger than the kinetic energy of the gyre-scale circulation. These two facts, they argued, are consistent with the hypothesis that eddies are generated by baroclinic instability. Baroclinic instability develops in flow where both rotation and stratification are important, as they are at the oceanic mesoscale, and acts to convert mean available potential energy to eddy kinetic energy. GGS also perform some idealized linear instability calculations and show that the resulting growth rates are roughly consistent with observed eddy timescales.

Recent analyses of the global statistics of eddies using satellite altimetry (Stammer, 1997; Chelton *et al.*, 2007) are consistent with most aspects of this picture, demonstrating that eddy frequencies, for example, are well correlated with the approximate Eady growth rate $f \langle \mathrm{Ri}^{-1/2} \rangle$ from hydrographic data. Stammer (1997) also analyzes the horizontal *scale* of the observed eddies, and argues that it is linearly correlated with the local first deformation scale. In the simplest examples of linear baroclinic instability, such as the Eady and Phillips models (see, e.g., Pedlosky, 1987), the scale of fastest linear growth is also near the deformation scale, and so this apparent correlation seems to be reasonable. Therefore, one might argue that the observed eddy structure can be largely predicted and understood through linear theory. The implication is that the ocean is in a near-linear state, characterized by weakly growing waves and mild turbulence. Is this a reasonable conclusion?

One can go beyond estimates of the Eady timescale and deformation scale by systematically investigating the local linear instability of the oceanic mean state, as envisioned by GGS. There appears to be only the calculation by Killworth and Blundell (2007) of the full instability problem using observed hydrography (and their purposes and presentation were quite different than that undertaken here), likely because only recently have global hydrographic atlases approached completeness. The goals of this paper are

1. To present a systematic analysis of the linear stability characteristics of the oceanic mean state—a reference calculation—, and
2. To compare the scales of fastest, energetically significant growth to observed eddy statistics.

The dataset of Gouretski and Koltermann (2004, denoted GK04 hereafter) provides monthly, seasonal and annual averages of temperature and salinity gridded at 1/2 degree in the horizontal, on 45 fixed depth levels, from 72S to 90N, at all longitudes. In the present study we use annually averaged data, though comparisons were made to seasonally averaged results, showing negligible differences. Notably, the GK04 dataset interpolation is performed in such a way that neutral stability is ensured (this is not the case for the World Ocean Atlas 2001 dataset). See GK04 for details and an explicit comparison to the World Ocean Atlas 2001 dataset. The intrinsic buoyant stability of the dataset allows for the straightforward computation of "neutral" density (Jackett and McDougall, 1997) for the

entire world ocean. Topographic slopes are included in the calculation via estimates from the Smith and Sandwell (1997) topographic atlas.

The complete wave and instability characteristics are presented for a few prototypical locations, and these demonstrate a number of important features. First and foremost, it is found that in most locations, fastest growth occurs at a scale *smaller* than the internal deformation scale. Unlike the Eady instability, the primary instabilities are typically due to mean PV gradient zero crossings of the internal shear at thermocline depths. While Eady instabilities have fastest growth at scales larger than the first deformation radius ($R_{eady} = NH/(1.6f) = 1.96R_1$ for constant N), these internal instabilities have scales of order $R_1/4$ and smaller. Moreover, in many locations there is significant growth at much smaller scale (down to scales of 100s of meters) due to the strong shears and weak stratification near the surface. However, (1) at these scales, quasigeostrophic scaling is not valid, in particular because the stratification is weak and the shears strong; (2) the mean hydrography is not accurate on such small space and time scales; and (3) it is shown here that these surface instabilities are less energetically potent (in terms of their potential energetic transfer rate for a given perturbation amplitude) than the thermocline shear-driven instabilities. This is not to say that these instabilities are unimportant—they are likely a crucial part of the submesoscale dynamics of the surface mixed layer, as demonstrated by the more complete analysis of Boccaletti *et al.* (2007). The focus of the present analysis is the larger scale and energetically potent instabilities of the mean hydrography.

The global picture that emerges indicates that while the time-scales of observed eddies are close to that predicted by linear theory, the spatial scales are uniformly larger than the fastest growing modes. The discrepancy in scale is most extreme in the Antarctic Circumpolar Current, where eddy scales are of order 30–100 km (Stammer 1997; Chelton *et al.* 2007, and R. Scott, pers. comm.), while the largest scales with significant growth are on order 5 km, and the fastest growing modes are at much smaller scale still. Possible explanations are that other, larger-scale instabilities missed by the present analysis lead to the observed eddy scales, or that errors in the mean state lead to spurious results. Scott and Wang (2005), however, present striking evidence of a strong inverse cascade throughout the ocean, consistent with the finding here that linear analysis is insufficient to predict eddy scales.

The paper is organized as follows. Basic quantities of relevance to mesoscale eddy dynamics, such as the local neutral modes and deformation scales, and the Eady timescale, are calculated from hydrography in Section 2. The details of the linear normal mode calculation are presented in Section 3, and basic results of linear analysis are reviewed. As is well known, sufficient conditions for instability are based on the mean potential vorticity and buoyancy gradients—these quantities, as well as an estimate of the Charney depth, are computed for the global dataset in Section 4. Linear instability calculations for four profiles at different locations are performed in Section 5, and from these generic features are identified. The local results guide a complete calculation of linear instability at every horizontal location in the ocean in Section 6. Energetic conversion rates are also computed

and used to discriminate the growth rates by energetic relevance. The paper concludes with a discussion in Section 7, and includes three appendices.

2. Neutral modes, deformation scales and the Eady growth rate

The baroclinic instabilities of the oceanic mean state are typically assumed to be characterized by a length-scale near the first Rossby radius of deformation, and a timescale given by an estimate of the Eady growth rate. Therefore, before calculating the linear wave solutions, we explain and compute these fundamental characteristics of the mean hydrography. These computations appear elsewhere in the literature, but are repeated here to provide a self-contained baseline to which the full instability calculation can be compared.

First, neutral density is computed from the GK04 dataset using the Ocean Data View software package. As done earlier by Chelton *et al.* (1998), we then compute the neutral modes and deformation scales of the neutral density field. The neutral modes are the vertical structure of solutions to the quasigeostrophic equations linearized about a state of rest (see, e.g., ref. Pedlosky, 1987). Specifically, the neutral modes are solutions $\phi^m = \phi^m(z)$ with eigenvalues K_m to the Sturm-Liouville eigenvalue equation

$$\Gamma\phi^m \equiv \frac{d}{dz}\left(\frac{f^2}{N^2}\frac{d\phi^m}{dz}\right) = -K_m^2\phi^m, \quad \text{with} \quad \left.\frac{d\phi^m}{dz}\right|_0 = \left.\frac{d\phi^m}{dz}\right|_{-H} = 0, \qquad (2.1)$$

where $N^2(z) = -(g/\rho_0)d\bar\rho/dz$, for each vertical profile of neutral density $\bar\rho = \bar\rho(z)$. In the above, $f = 2\Omega\sin(\theta)$ is the Coriolis parameter at latitude θ, with rotation rate Ω, g is the gravitational acceleration and ρ_0 is the mean oceanic density, and here we assume a flat bottom at depth $z = -H$ (though topography is included in the instability calculation later) and a rigid lid at $z = 0$. The resulting eigenvalues K_m are the local internal deformation wavenumbers, or the reciprocals of the Rossby deformation radii $R_m = 1/K_m$, and the eigenfunctions ϕ^m form a complete set.

Solutions to (2.1) are computed numerically using centered differences. Note that there are a number of GK04 profiles with levels at which the centered estimate of the vertical derivative of the density between two prescribed depth levels is zero. At these points we simply skip down to the next level at which there is an increase in density, and so always retrieve a non-zero vertical profile of the density derivative. The resulting contours of the first deformation radii are plotted in Figure 1a. More detailed analysis and discussion of the deformation radii can be found in Chelton *et al.* (1998).

An intrinsic eddy timescale can be estimated by consideration of the energy conversion generated by baroclinic instability. Assuming horizontally constant flow in thermal wind balance with the density field, the mean available potential energy in a square area of length L is

$$\text{APE} = \frac{\rho_0}{2}\int_{-H}^{0}\frac{f^2}{N^2}\Psi_z^2 dxdydz = \frac{\rho_0 L^4}{2}\int_{-H}^{0}\frac{f^2}{N^2}\left(\frac{U_z^2 + V_z^2}{3} - \frac{U_z V_z}{4}\right)dz$$

First Rossby Radius of Deformation (Km)
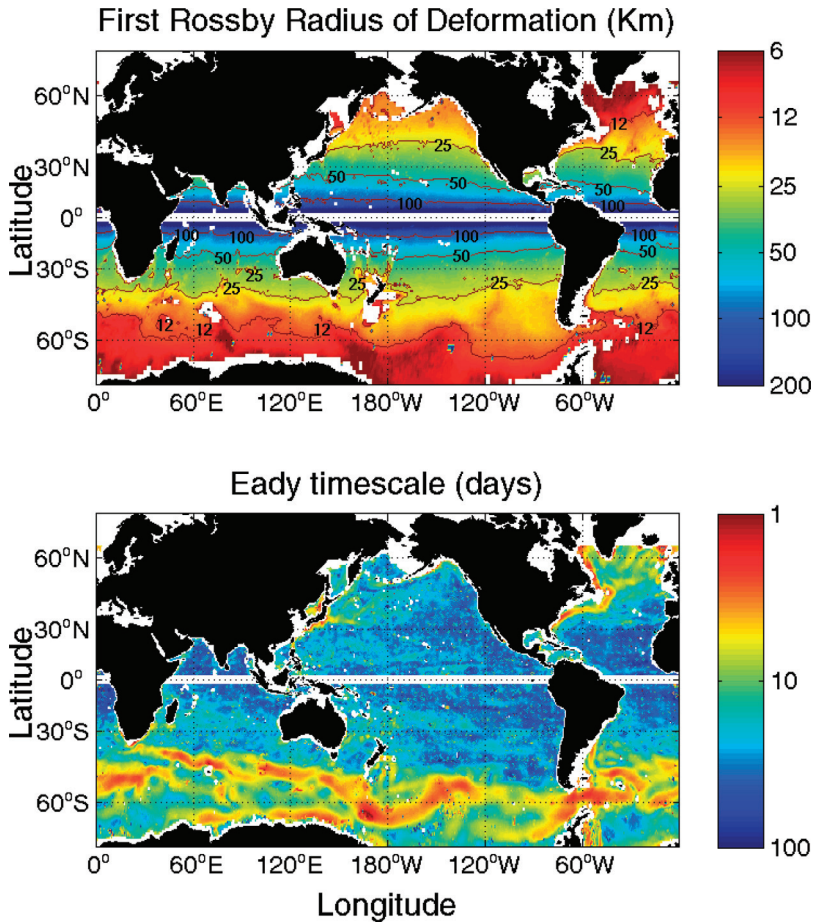

Eady timescale (days)

Figure 1. (a) The first internal deformation radii, R_1, as calculated from the GK04 dataset, plotted with logarithmic color scale ranging from 6 km in red to 200 km in blue. Contours at 12, 25, 50 and 100 km are overlayed. (b) The Eady timescale σ^{-1}, from (2.3), in days.

where $\Psi = -Uy + Vx$ is the streamfunction for the mean flow, quasigeostrophic scaling has been assumed and the subscript z denotes a partial derivative. Neglecting the cross term, the APE can be expressed as

$$\text{APE} \simeq \frac{\rho_0 L^4}{6} \int_{-H}^{0} \frac{f^2}{\text{Ri}(z)} dz,$$

where

$$\text{Ri} = \frac{N^2}{U_z^2 + V_z^2} \tag{2.2}$$

is the Richardson number. We can compare this quantity to the growth rate in the Eady problem for baroclinic instability between two rigid surfaces with a uniform temperature gradient, $0.31 f/\sqrt{\mathrm{Ri}}$—since Ri is constant in the Eady problem, no integral is necessary. Accordingly, a natural estimate of baroclinic growth rates is

$$\sigma = f \sqrt{\frac{1}{H} \int_{-H}^{0} \frac{dz}{\mathrm{Ri}(z)}} \propto \sqrt{\mathrm{APE}}, \tag{2.3}$$

and we call the inverse of this quantity the "Eady timescale". The Eady timescale σ^{-1} for the global ocean is calculated using the thermal wind shear (see Appendix A for details of the mean shear computation), and the result is shown in Figure 1b. Stammer (1997) and Treguier *et al.* (1997) both compute similar quantities, but use only hydrography from the upper ocean. As was shown by Stammer (1997), the locations of high σ are correlated with locations of high observed eddy kinetic energy—this is the primary reason to assign eddy generation to baroclinic instability. In Section 6 we show that the inverse Eady timescale calculated here is, in fact, a good estimate of the actual growth rates of the energetically significant instabilities.

3. Linear normal-mode instability

The inviscid quasigeostrophic equation, linearized about the local mean state characterized by $\boldsymbol{U} = U(z)\hat{\boldsymbol{i}} + V(z)\hat{\boldsymbol{j}}$, N^2 and ∇Q, is

$$q_t + \boldsymbol{U} \cdot \nabla q + \boldsymbol{u} \cdot \nabla Q = 0, \quad -H < z < 0, \tag{3.1a}$$

$$\psi_{zt} + \boldsymbol{U} \cdot \nabla \psi_z + \boldsymbol{u} \cdot \nabla(\Psi_z + f^{-1}N^2\eta) = 0, \quad z = -H, 0, \tag{3.1b}$$

where $q = \nabla^2\psi + \Gamma\psi$ is the eddy quasigeostrophic potential vorticity (QGPV), $\boldsymbol{u} = -\psi_y\hat{\boldsymbol{i}} + \psi_x\hat{\boldsymbol{j}}$ is the eddy velocity field expressed in terms of the horizontal streamfunction, $\psi = \psi(x, y, z, t)$, Γ is defined in (2.1), $\nabla\Psi_z = V_z\hat{\boldsymbol{i}} - U_z\hat{\boldsymbol{j}}$ is proportional to the mean buoyancy gradient and η is the height of the topography at $z = -H$ (set $\eta = 0$ at $z = 0$ of course). The mean quasigeostrophic potential vorticity gradient ∇Q is

$$\nabla Q = \Gamma V\hat{\boldsymbol{i}} + (\beta - \Gamma U)\hat{\boldsymbol{j}} = \beta\hat{\boldsymbol{j}} - f\boldsymbol{s}_z,$$

where $\beta = 2\Omega R_e^{-1} \cos(\theta)$, with R_e the radius of the Earth, Ω its rotation rate and θ the latitude. In the second expression, thermal wind balance is used to express the mean vortex stretching in terms of $\boldsymbol{s} = -\nabla\bar{\rho}/\bar{\rho}_z$, the isopycnal slope vector.[2]

---

2. Here we have used thermal wind balance to write

$$f\frac{\partial}{\partial z}(-V, U) = \frac{g}{\rho_0}\nabla\bar{\rho}$$

and that $N^2 = -(g/\rho_0)\partial\bar{\rho}/\partial z$ to write the mean vortex stretching in terms of the mean isopycnal slope. Formally, we have violated QG scaling by allowing horizontal variations in N^2, but this approximation is necessary to compute the mean shears from hydrography.

The mean velocity and stratification are taken to be horizontally local and slowly varying, hence at each horizontal location depend only on z (see Pedlosky, 1984). Consistent with the assumption of a slowly-varying background state that is locally horizontally homogeneous, the domain is taken to be horizontally-periodic. Substitution of a plane-wave solution of the form $\psi = \mathfrak{Re}\{\hat{\psi}(z) \exp[i(kx + \ell y - \omega t)]\}$, where $\hat{\psi}$ is the complex amplitude, into (3.1) gives

$$(\boldsymbol{K} \cdot \boldsymbol{U} - \omega)(\Gamma - K^2)\hat{\psi} = -\Pi\hat{\psi}, \quad -H < z < 0, \tag{3.2a}$$

$$(\boldsymbol{K} \cdot \boldsymbol{U} - \omega)\hat{\psi}_z = \Lambda\hat{\psi}, \qquad z = -H, 0, \tag{3.2b}$$

where $\boldsymbol{K} = (k, \ell)$, $K = \sqrt{k^2 + \ell^2}$, $\Pi = \Pi(z) = kQ_y - \ell Q_x$ and $\Lambda = \Lambda(z) = k(U_z - N^2\alpha^y/f) + \ell(V_z + N^2\alpha^x/f)$. In the expression for Λ, $\alpha^{x,y} = 0$ are the mean topographic slopes at the bottom (and are 0 for the upper boundary condition). Equations (3.2) form an eigenvalue problem for the normal modes $\hat{\psi}$ (the eigenfunctions) and the frequencies ω (the eigenvalues). If the frequency has a nonzero imaginary part, the wave solution grows exponentially. The wavenumber of fastest growth K_{max} is defined such that $\omega_i(K_{max}) = \max(\omega_i)$, where $\omega_i = \mathfrak{Im}\{\omega\}$.

The vertically discrete problem is derived in Appendix B. There it is shown that the boundary conditions (3.2b) are included in the construction of the discrete stretching operator, Γ_{nm}. Before proceeding to the direct calculations, though, it is useful to consider the structure of the mean gradients and their influence on the nature of the instabilities.

4. Mean hydrographic gradients

Much can be understood about the instability characteristics of the mean state by considering the mean potential and relative vorticity gradients. In the special case of a zonal mean flow ($V = 0$), a necessary condition for an imaginary frequency is given by the Charney-Stern-Pedlosky criterion (Charney and Stern, 1962; Pedlosky, 1964), which states that one of the following must hold: (a) Q_y changes sign in the interior, (b) U_z has the same sign at the upper and lower boundaries (as in the Eady problem Eady, 1949), (c) U_z at the upper boundary has the opposite sign of Q_y in the interior, or (d) U_z at the lower boundary has the same sign as Q_y in the interior (as in the Charney problem Charney, 1947). See Pedlosky (1987) or Vallis (2006) for a review. Since we have explicitly removed horizontal gradients from the mean velocities, sign changes in Q_y must be due solely to sign changes of mean vertical vortex stretching in the vertical. In the more general case considered here ($V \neq 0$), one can derive a similar necessary condition, but there are many more possibilities for instability (see Appendix C). Nonetheless, one can generalize the instabilities as resulting from the interaction of the shears at the upper and lower surfaces, the interaction of the interior PV gradient with boundary shears, or from sign changes of the interior PV gradient alone. These three types of interactions will be termed Eady-type, Charney-type and Phillips-type (after Phillips, 1954), respectively.

a. Quasigeostrophic potential vorticity gradients

The mean meridional QGPV gradient includes the environmental vorticity gradient β. Where flows are non-zonal, instability is present with any shear, whereas for zonal flow, the planetary vorticity gradient β can suppress instability if it is large enough compared to the mean stretching term. The ratio $\nabla Q/\beta$ is a rough measure of the supercriticality of the flow (an exact measure in the two-layer, zonal flow case), and for energetically significant instabilities, its magnitude is an indicator of the strength of eddy generation (Held and Larichev, 1996).

Figure 2 displays the nondimensional zonal and meridional mean QGPV gradients along the transect at 31W. The figure demonstrates the overwhelming dominance of the mean stretching term relative to the background planetary vorticity gradient for latitudes above about 30 degrees, and especially in the southern ocean, where the ratio can exceed 100. The large increase in the magnitude of the mean QGPV gradients with latitude is due to the increase in the mean vortex stretching term, $\Gamma U = (f^2/N^2 U_z)_z$. The increase in mean vortex stretching is due both to the increase in the Coriolis term f, as well as to the decrease in static stability N. In the ACC, the effect is more pronounced because the shear itself (and its curvature) are so large.
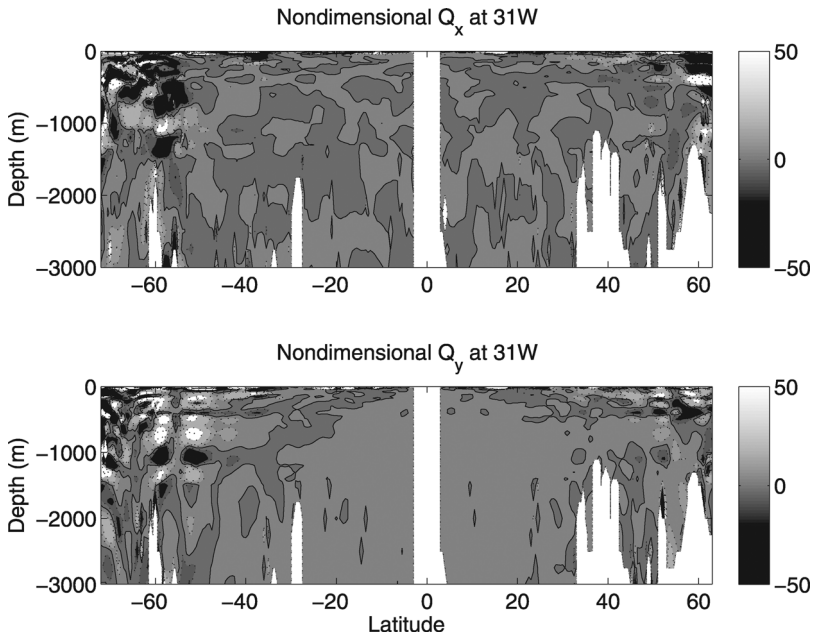

Figure 2. Zonal (a) and meridional (b) QGPV gradient in the upper 3000 m of the Atlantic, along 31W. Values are nondimensionalized by local values of β. The zero contour is drawn with a solid line. The color-scale is saturated at ±50, but the maxima and minima are in the 100s for each panel.

Figure 2 shows a clear pattern of gradient reversals with depth, most prominently in the meridional gradient (panel b). Results are presented here for a transect rather than a zonal average, since it is necessary to preserve the sign of the gradient. The patterns vary somewhat along other transects, but generically, one finds strong gradient reversals, exceeding β by factors of 10–100, at all longitudes and in both northern and southern hemispheres. Moreover, the strongest gradient reversals are in the upper 600 meters in the northern hemisphere, but extend to well below 1000 meters (and with larger gradients) in the Southern Ocean. The structure of the mean QGPV gradient demonstrates the likelihood of active baroclinic instability in much of the extratropical ocean, concentrated at depths between the surface and 600 meters in the northern oceans, and between the surface and about 1500 m in the Southern Ocean.

b. The Charney depth

The structure of the mean QGPV gradients indicates that the mean state is highly susceptible to baroclinic instabilities of the Phillips type (generated by interior QGPV gradient sign changes). However, there are also strong shears at the surface. To what extent are Charney-type instabilities present, and what can we expect of their structure? The original Charney instability problem considers a mean vertical shear of horizontal velocity that is constant with height over a rigid lower boundary, with decay conditions above. The interior QGPV gradient is $Q_y = \beta$, and so the instability is of type (d) above. The calculation is mathematically challenging (see Pedlosky, 1987), but the important result is that growing waves from this type of instability have a self-selecting scale height $h_C = U_z f^2 / \beta N^2$, where $N = N(z = 0)$, independent of the depth of the fluid (which in this case is infinite). The horizontal length of fastest growth scales like $N h_C / f$ and the growth rate scales like $f U_z / N$.

For an interior QGPV gradient set by the interior shear instead of β (as is apparently the case in most of the extratropical ocean), we can generalize the Charney height as $h_C = U_z f^2 / Q_y N^2 = s^{\text{surf}} / s_z^{\text{int}}$, where $s = -|\nabla \bar{\rho}| / \bar{\rho}_z$ is the isopycnal slope. Thus we expect horizontal growth at a scale $N s^{\text{surf}} / f s_z^{\text{int}}$ and a growth rate that scales the same as in the original problem. Figure 3 shows the estimated Charney depth scale h_C, where s^{surf} is averaged over the top 50 m and s_z^{int} is averaged over depths 75–500 m. Apart from some isolated areas, the typical effective Charney depth is on order 100–300 m. Given that, as found above, $s_z \gg \beta$, it is not surprising that h_C is rather small. In the full instability calculations presented in the next two sections, surface-intensified Charney-like instabilities are common.

c. Relative vorticity gradients

The likelihood of barotropic instability can be assessed by considering the mean meridional vorticity gradient at the upper surface, where velocities are strongest. Consider the nondimensional estimate of the mean relative vorticity gradient $(\beta - U_{yy})/\beta$. If no vortex
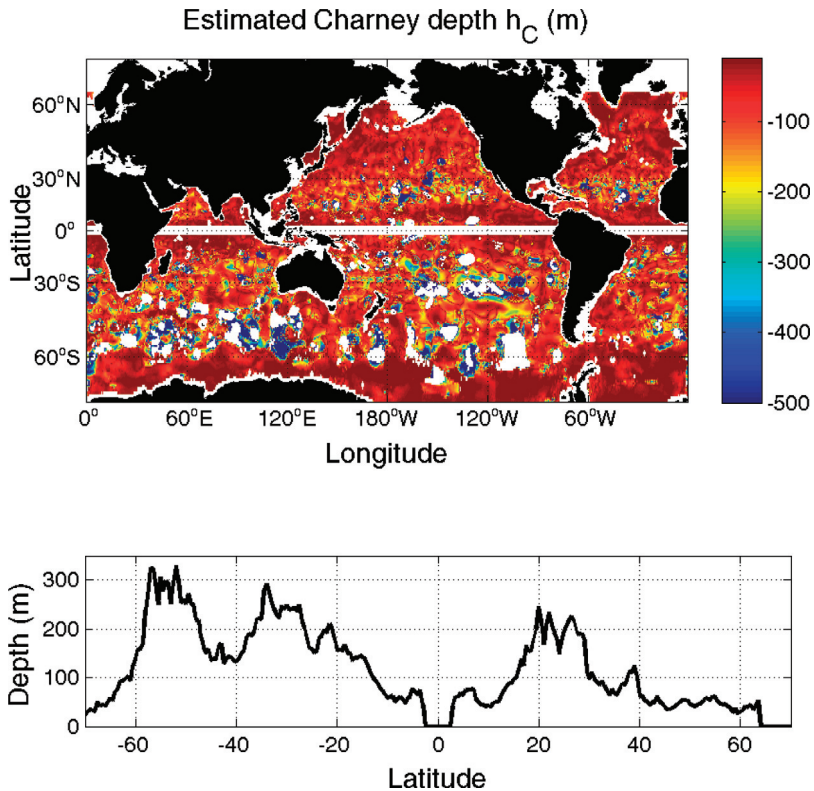
Estimated Charney depth h_C (m)

Figure 3. (a) The estimated depth h_C of eddies generated by Charney-type baroclinic instabilities, in meters (see Section 3b). (b) The zonal average of h_C.

stretching were present, negative values of this quantity would indicate sufficient conditions for barotropic instability. In fact, this quantity never goes below 1/2 (plot not shown, but is available from the author), and given that the strength of the vortex stretching term outside the tropics is between 20 and 100 times larger than β, we may safely conclude that barotropic instability is not actively produced by the mean circulation. Of course, in regions near coastlines and possibly near the equator (both filtered from this analysis), as well as in intermittently generated sharp currents, barotropic instability may be important—the present results pertain only to the annually averaged hydrography outside the equatorial region.

5. Analysis of a few specific locations

We first consider the solutions at a few individual locations, and identify some generic features. For each location, the linear wave solutions are calculated for a range of k and ℓ

distributed about the local first deformation wavenumber K_1. For each wavenumber, there are as many eigensolutions as there are levels on the local vertical grid. The eigensolution with with the largest imaginary frequency (growth rate) is chosen and the others discarded (typically only a few other modes have imaginary parts—most are stable).

a. The Gulf-Stream Jet: 60W, 40N

The mean neutral density and mean horizontal velocity profiles are shown in Figure 4a–b. The location chosen is in the region of the Gulf Stream jet extension in the western North Atlantic, dominated by westward flow in and above the thermocline. The density structure is characterized by two regions of rapid increase: the first in the upper 100 m, the second at
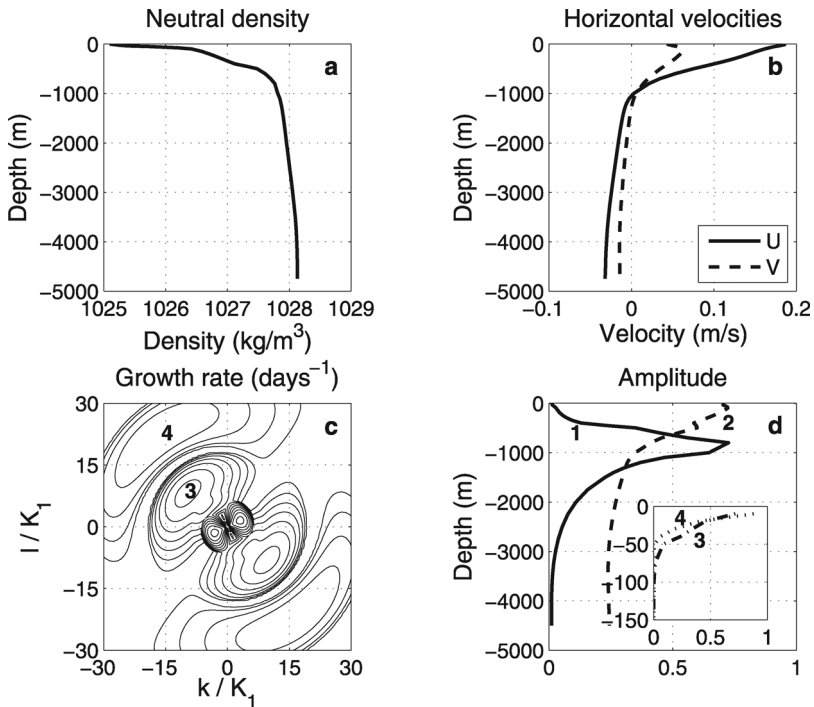

Figure 4. Mean stratification, horizontal velocity and linear instability at 60W, 40N. (a) Neutral density (kgm^{-3}); (b) Zonal (solid) and meridional (dashed) mean velocity (m/s). (c) Growth rate in days^{-1} as a function of horizontal wavenumber nondimensionalized by the local first deformation wavenumber, K_1. (d) Amplitudes $|\hat{\psi}|$ for growing modes at four particular wavenumbers. Amplitude '1' corresponds to the largest growth rate in panel c, and '2' corresponds to a point very near the origin. Amplitudes '3' and '4' correspond to the labeled peaks in panel c. The peaks have growth rates and locations as follows. '1': $\omega_i = 0.11$ days^{-1}, $K/K_1 = (3.1, 1.5)$; '2': $\omega_i = 0.082$ days^{-1}, $K/K_1 = (0.59, 0.59)$; '3': $\omega_i = 0.079$ days^{-1}, $K/K_1 = (-9.8, 8.2)$; and '4': $\omega_i = 0.054$ days^{-1}, $K/K_1 = (-16, 22)$.

about 400 m. The zonal velocity structure is simpler, smoothly decreasing from the surface to about 1000 m depth, then more gently in the abyss (as mentioned before, the absolute mean velocity profiles are chosen to have no barotropic component). The meridional velocity is about 1/2 the magnitude of the zonal velocity, but peaks at about 150 m depth.

The linear instabilities were calculated on a grid of wavenumbers in the upper half plane ranging from $K_1/10$ to $500K_1$ in each horizontal direction. The resulting growth rate is shown in Figure 4c, as a function of k/K_1 and ℓ/K_1, and the nondimensional amplitudes $|\hat{\psi}|(z)$ of the four largest distinct growth rate peaks labeled numerically in panel c are shown in Figure 4d. The growth rates in Figure 4c are shown on the full k-ℓ plane (the lower half plane is given by conjugate symmetry), but only for wavenumbers ranging in absolute value from $K_1/10$ to $30K_1$ in either direction. There are no significant instabilities at smaller scales.

We can summarize the results as follows:

- The largest peak (fastest growth rate—labeled '1') corresponds to an amplitude with its peak at a depth of about 800 m, and has a horizontal wavenumber of $3.4K_1$, or equivalently a horizontal scale $R_1/3.4$, where R_1 is the first Rossby radius of deformation. This length is consistent with a scaling down from R_1 by the ratio of the vertical extent of the instability to the full depth of the ocean, or $R_{\max} \simeq (N/f)h_{\text{thermocline}}$. The vertical structure of this instability indicates that it likely results from interior PV gradient sign changes (Phillips-type).

- Peak '2' is not obviously a peak in this figure, but a detailed closeup of the small-wavenumber region reveals that it is a separate instability. These larger-scale modes, found generically, are Green modes (Green, 1960), and accordingly have much smaller growth rates than the primary peak. The vertical structure, while surface intensified, has amplitude at all depths. Such modes are also found in the idealized computations of GGS.

- Peak '3' occurs at one twelfth the local deformation radius, and '4' is at $R_1/28$. The small scales of these instabilities are consistent with their vertical structures: both are extremely surface intensified—their amplitudes are entirely contained in the top 100 m of the ocean (see inset in Fig. 4d). The orientation of these instabilities in the $k - \ell$ plane is nearly orthogonal to those of the larger-scale instabilities, consistent with the reversal of the meridional shear near the surface. The vertical scale is consistent with the typical vertical scale estimated for Charney-like instabilities, and so it seems likely that these modes result from the interaction between surface shears and subsurface QGPV gradients. Another possibility is that they arise from small kinks in the density or shear profile, resulting in a 'defect' instability (Samelson, 1999).

b. Antarctic Circumpolar Current: 142E, 51S

This location is just south of Tasmania, in the Antarctic Circumpolar Current (ACC), and is the location of a field campaign reported on by Phillips and Rintoul (2000). A more
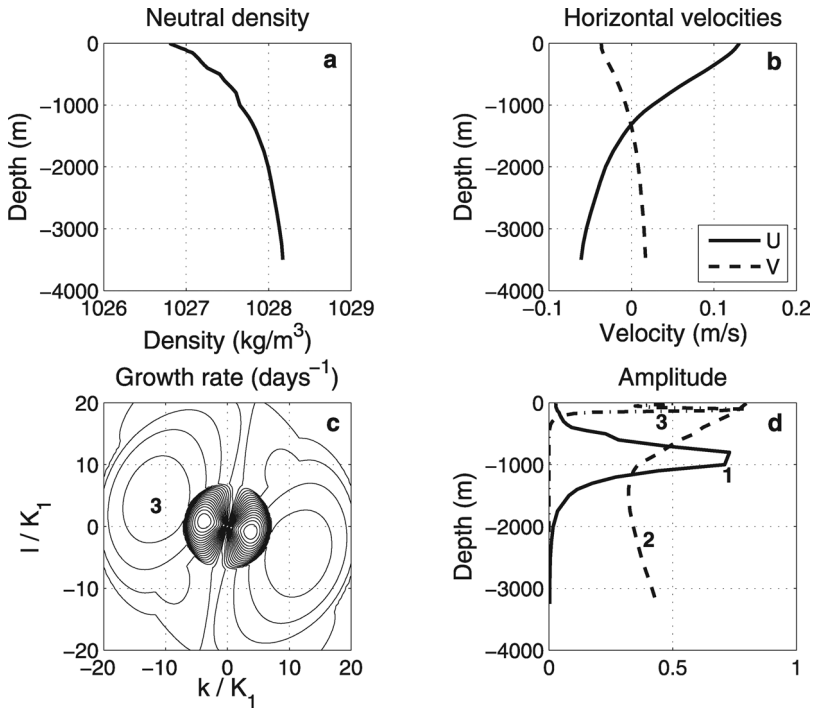
Figure 5. Same as Figure 4, but for 142E, 51S, in the Antarctic Circumpolar Current, south of Tasmania. The peaks have growth rates and locations as follows. '1': $\omega_i = 0.18$ days^{-1}, $K/K_1 = (-3.7, 0.86)$; '2': $\omega_i = 0.11$ days^{-1}, $K/K_1 = (-0.55, 0.15)$; and '3': $\omega_i = 0.048$ days^{-1}, $K/K_1 = (-11, 4.2)$.

detailed investigation of the dynamics of ACC instabilities and the eddies they produce, as well as a comparison to the field data in this particular location, is taken up in Smith and Marshall (2008). The mean state and instability structure are presented in Figures 5a–d in the same format as for the previous location. The location chosen is near the axis of the ACC jet and so the velocity is primarily zonal, with most of the shear located near thermocline depths. The density structure is nearly exponential with depth, apart from a mixed layer near the surface, and a vertical jump near 400 m. Overall the stratification is weaker than in the previous location, and the vertical structure is less surface-intensified.

The apparently simple structure of the mean velocity and density profiles nevertheless leads to mean QGPV gradients that are complex, with multiple zero-crossings in the vertical (refer to Fig. 2). The linear instabilities are calculated as for the previous location, and the results are again displayed as a function of horizontal wavenumber in Figure 5c. Unstable scales range from greater than the local deformation scale down to scales less than $R_1/10$

(at this location, $R_1 \simeq 10$ km, and so growth at $K/K_1 \lesssim 10$ corresponds to less than 1 km). The strongest instability has a growth rate of 0.16 days^{-1}, almost twice the maximum for the Gulf Stream location (0.09 days^{-1}). There are weak instabilities at much smaller scales (down to $R_1/100$), but these are not shown (all are surface-trapped, Charney-like instabilities, and are occurring at scales not resolved by the hydrography, and are not well-represented by quasigeostrophic scaling).

The amplitudes $|\hat{\psi}|(z)$ of the three most prominent, distinct instabilities are shown in Figure 5d. These are similar to those in the Gulf Stream location, and the generalizations made for those instabilities apply here as well.

c. The Mid Pacific: 172E, 44N

We consider this location, in the Kuroshio extension, because the mean current is directed northeastward, not zonally. Baroclinic instability generated by nonzonal mean flow is not inhibited by β, and so even weak shears can lead to nonlinear growth. However, at latitudes outside the tropics, as shown in Figure 2, the mean QGPV gradient due to vertical shear of the horizontal mean velocity is 20–100 times larger than the planetary gradient, β, and so the orientation of the flow with respect to the planetary vorticity gradient is largely irrelevant for the linear instabilities. Nevertheless, we consider this current location as before. The linear instability calculation and the format of the Figures 6a–d is identical to the previous two examples. Note first that the fastest growth rate, 0.035 days^{-1}, is smaller by less than one half that of the Gulf Stream location showed in the first example. The maxima are oriented along the direction of the mean shear near the surface.

The vertical structures of the modes corresponding to the labeled peaks in Figure 6c are shown in Figure 6d. Notably both peaks here have their maxima at or within 200 m of the surface. The largest instability has no amplitude at depths below 2000 m, while the second (larger-scale) instability again has amplitude at all depths. In broad terms, the structure of the instability is similar to the other locations, albeit with the main instability more surface intensified (due to weaker stratification and weaker shears), and no small-scale surface modes.

d. The Mid North Atlantic: 60W, 23.5N

This last location, in the Sargasso Sea, is chosen as a representative example of the kind of very small scale, surface instability that dominates the low latitudes. The mean structure and resulting instability are shown in Figure 7, with the same format as for the previous example locations. The mean flow at this location is characterized by a southwestward shear concentrated in the upper thousand meters, with a reversal in U in the upper 200 m (a turning velocity field). The instability is dominated by surface instabilities, but also exhibits one larger-scale mode that is peaked at 500 m depth. Note that the growth rate in panel (c) extends to $\pm 100 K_1$ in each direction, and that in panel (d) the vertical structures for the two prominent growth rate peaks are shown only for the top 2000 m. The maximum growth
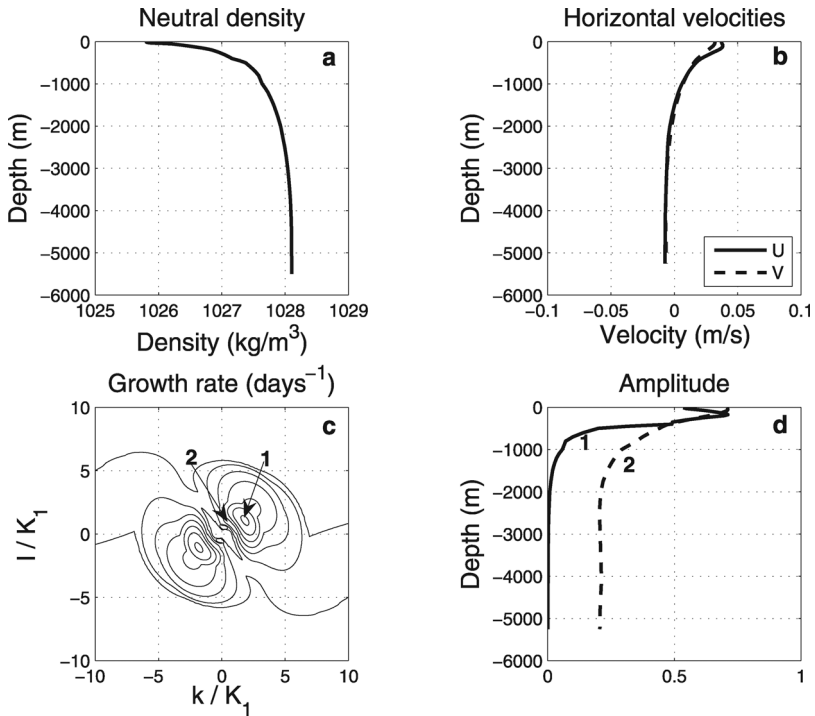
Figure 6. Same as Figure 4, but for 172E, 44N, in the western Pacific, just north of the Kuroshio extension. The peaks have growth rates and locations as follows. '1': $\omega_i = 0.036$ days^{-1}, $K/K_1 = (1.8, 1.1)$; and '2': $\omega_i = 0.015$ days^{-1}, $K/K_1 = (0.16, 0.51)$.

rate is almost as large as that in the Gulf Stream location analyzed in Section 5a, yet it is confined to a very small portion of the water column. The degree to which such spatially confined but quickly growing instabilities can affect the mean state will be addressed in the next section.

6. Analysis of the global dataset

The global dataset is analyzed as follows. For each location, a discrete grid of k and ℓ values in the upper-half spectral plane is formed, with domain $-k_{max} \leq k \leq -k_{min} \bigcup k_{min} \leq k \leq k_{max}$ and $\ell_{min} \leq \ell \leq \ell_{max}$. The discrete values of k and ℓ are distributed logarithmically between their respective minimum and maximum values, with $k_{min} = \ell_{min} = 0.1K_1$ and $k_{max} = \ell_{max} = 100K_1$, where K_1 is estimated by latitude using the functional fit proposed by Chelton *et al.* (1998). At each location, 31 values of ℓ and 62 values of k are used, for a total grid size of 1922 wavenumbers.

If the horizontal spatial scale of the fastest growing wave at each location in the ocean is plotted without filtration, the field is dominated by values more than an order of magnitude
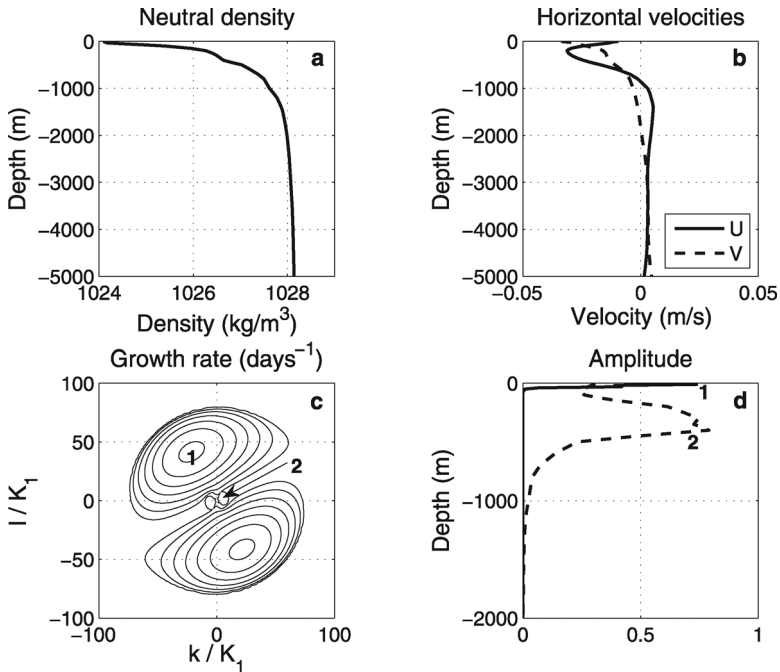
Figure 7. Same as Figure 4, but for 60W, 23.5N, in the Sargasso Sea. Note that the range of nondimensional wavenumbers shown in panel c is -50 to 50, a much larger range than in the previous three growth rate figures. Also note that the vertical scale in panel d runs only from 0 to -300 m. The peaks have growth rates and locations as follows. '1': $\omega_i = 0.072$ days^{-1}, $K/K_1 = (-12, 23)$; and '2': $\omega_i = 0.017$ days^{-1}, $K/K_1 = (2.2, 0.48)$.

smaller than the local deformation radius (not shown, but see thin solid line in Fig. 11 for zonal average). Similarly, the temporal scales of the fastest growing waves are characterized in an average sense by values on order days (also not shown). This result was foreshadowed by Figure 3 and the local analyses of Section 5—the fastest growth at many locations is dominated by surface modes. The relevance of the sea of instabilities, however, depends on the degree to which each can affect the mean state. Given an unstable wave, how much energetic conversion from the mean state can the wave generate? We address this question before plotting growth rates and scales.

a. The baroclinic conversion rate

A growing wave will lead to a conversion of mean available potential energy (MAPE) to eddy kinetic energy (EKE). The rate of conversion can be calculated from the energy budget equation, obtained by multiplication of the quasigeostrophic equation of motion (3.1) by $-\rho_0 \psi$ and integration over the local spatial domain. The energy

conversion results from the advection of eddy potential vorticity by the mean shear, and is given by

$$G = \rho_0 \int_{-H}^{0} \overline{\psi U \cdot \nabla q} \, dz = \rho_0 \int_{-H}^{0} \frac{f^2}{N^2} \frac{\partial U}{\partial z} \cdot \overline{\nabla \psi \frac{\partial \psi}{\partial z}} \, dz,$$

where integration by parts is used in x and y (boundary terms and the horizontal vorticity fluxes vanish due to the double periodicity of the domain), and in z (where $\overline{\psi_z \psi_{xz}} = \overline{\psi_z \psi_{yz}} = 0$ by periodicity again). Note that G has units Wm^{-2}. For a given unstable wavenumber $K = (k, \ell)$, we can write the streamfunction as

$$\psi = \exp(\omega_i t) |\hat{\psi}| \cos[kx + \ell y - \omega_r t + \theta(z)].$$

Substituting this expression into the expression for G above gives, after horizontal integration over one cycle,

$$G = \int_{-H}^{0} R(z) dz \tag{6.1}$$

where

$$R = \exp(2\omega_i t) \frac{\rho_0}{2} \frac{f^2}{N^2} \frac{d\theta}{dz} |\hat{\psi}|^2 K \cdot \frac{dU}{dz}.$$

The amplitude $|\hat{\psi}|$ that arises from the instability calculation has units $[UL]$ but has no meaningful value, since it is the response to an infinitesimal perturbation. We follow GGS and normalize $\hat{\psi}$ such that it corresponds to a maximum eddy velocity of

$$V_e = \max_z [K |\hat{\psi}| \exp(\omega_i t)],$$

thus

$$R = \frac{V_e^2 \rho_0}{2} \frac{f^2}{N^2} \frac{d\theta}{dz} \left(\frac{|\hat{\psi}|}{|\hat{\psi}|_{\max}} \right)^2 \frac{K}{K^2} \cdot \frac{dU}{dz}. \tag{6.2}$$

In all cases, we set $V_e = 0.1$ m s^{-1}. A useful value to which G can be compared is the estimated rate at which energy is imparted to the oceanic mean state by the winds: 10^{-3} Wm^{-2} = 1 mWm^{-2} (GGS).

The conversion rate G is calculated at each horizontal location for the fastest growing mode and displayed in Figure 8a, in units of mWm^{-2}. Only values of G that exceed 0.5 mWm^{-2} are plotted, despite that lower values are present. The depth at which the integrand R in (6.2) has its maximum, for the same instabilities, is displayed in Figure 8b. The emergent pattern is that the surface-trapped instabilities that dominate the low-latitudes have very little baroclinic conversion potential, while the higher-latitude instabilities associated with the ACC, the Gulf Stream and the Kuroshio all have conversion maxima deeper in the water column, and much greater energetic conversion potential.
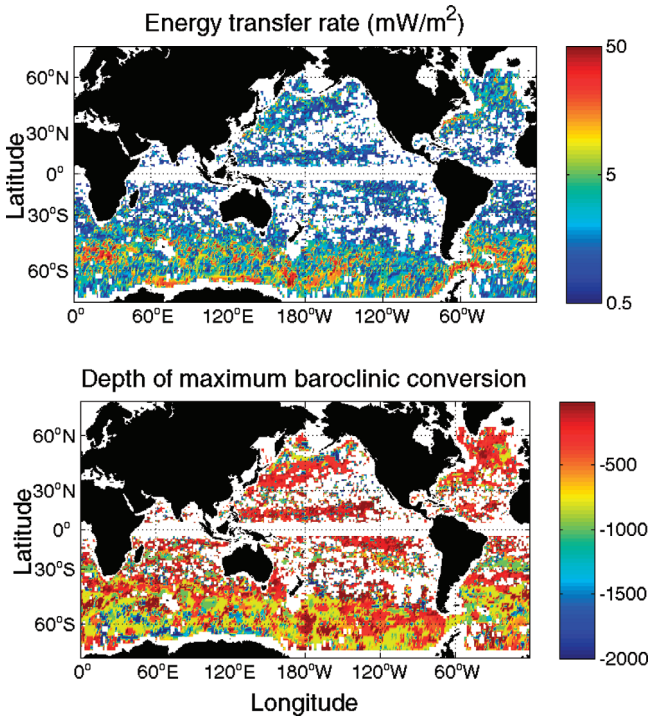
Figure 8. (a) The baroclinic conversion rate (6.1) (in $\mathrm{mWm^{-2}}$) for the fastest growing wave at each location, and (b) the depth of maximum conversion (the depth at which R of equation 6.2 is maximum) for the fastest growing wave (in m). Plotted values in each panel are subjected to the condition that the baroclinic conversion rate $G > 0.5\,\mathrm{mWm^{-2}}$ at each location.

In order to focus on the energetically important instabilities, we re-compute the linear instabilities at each location, but this time condition the search for the fastest growing modes on the requirement that G exceed $0.5\,\mathrm{mWm^{-2}}$. The inverse growth rates and scales of fastest growth filtered in this way are shown in Figure 9a–b, respectively. At many locations, slower but more energetically significant instabilities were missed by searching only for the fastest growth rates; at other locations, we now find no significant instabilities.

The results of the filtered calculation give a cleaner picture of the most energetically important baroclinic instabilities. A posteriori, the filtering operation has also selected the instabilities that are most most likely to have been accurately calculated in this QG framework. The smaller-scale, faster instabilities near the surface will typically involve ageostrophic effects not represented here (Boccaletti *et al.*, 2007).

The filtered maximum growth rates and their scales are plotted in nondimensional units in Figure 10a–b, respectively. Specifically, growth rates are compared to the APE timescale estimate σ^{-1} of (2.3), and scales of fastest growth are compared to local first deformation radii R_1. Apparently, the inverse Eady timescale is a good estimate of the energetically
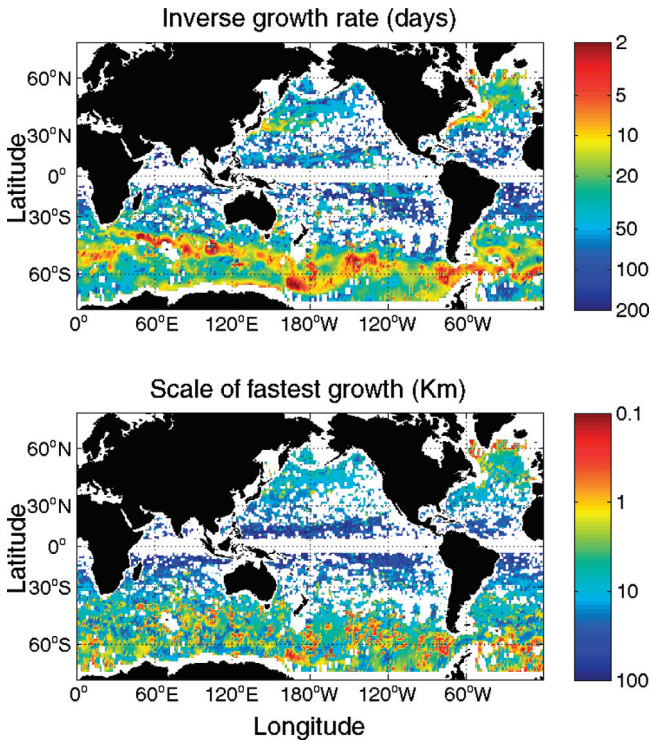
Figure 9. (a) Inverse growth rate (in days) for fastest growing wave at each location, and (b) scale of fastest growing wave (in km). A logarithmic color scale is used for each panel. Plotted values in each panel are subjected to the condition that the baroclinic conversion rate $G > 0.5$ mWm^{-2} at each location.

significant growth rates[3]. The first radius of deformation, on the other hand, is not a good estimate for the spatial scale of maximum growth, even when the results are filtered for energetic significance. In regions where the instabilities are strong (which are also regions where eddies are most prominent), the scale of maximum growth is typically about $R_1/4$, and much smaller still in many parts of the ACC.

As an indication of the degree to which the instabilities arise from zonal versus non-zonal mean shears, one can compute the angle of fastest growth, $\arctan |\ell_{max}|/|k_{max}|$ at each location. In a large fraction of the ocean, where meridional mean shears exist, the fastest growth happens in non-zonal directions (figure not shown). The implications of such growth on non-zonal currents is discussed in Spall (2000), Arbic and Flierl (2004) and Smith (2007).

---

3. The inverse Eady timescale is not, however, a good estimate of the unfiltered growth rate, which are much greater than σ in regions dominated by surface instabilities. The figure demonstrating this is not shown.
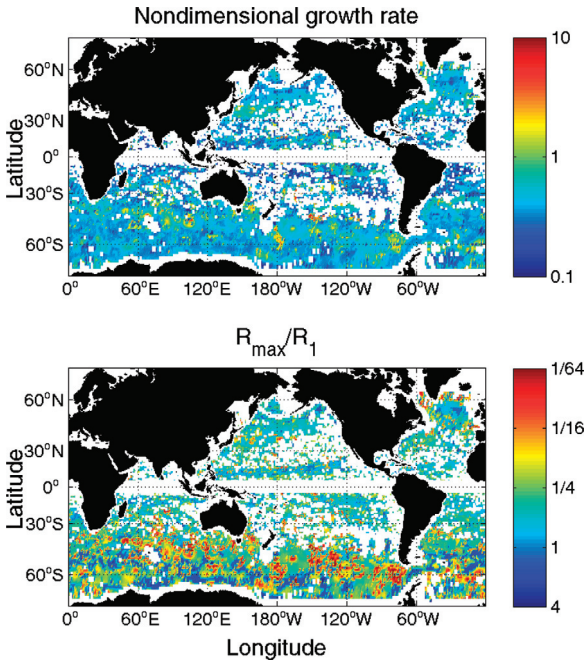
Figure 10. (a) Nondimensional inverse growth rate $f/(\sqrt{\text{Ri}}\omega_i)$ and (b) nondimensional scale of fastest growth $R_{\max}/R_1$, for results filtered by requirement that baroclinic conversion $G > 0.5\ \text{mWm}^{-2}$.

b. Comparison of linear growth and observed eddy scales

The scales of fastest linear growth can be compared directly to different estimates of eddy scale from analysis of satellite altimetry data. Stammer (1997) computes spectra of eddy kinetic energy for local regions throughout the global ocean, and summarizes the results in various ways. Eddy scales are estimated as the lag of the first zero crossing of the spatial autocorrelation function. A scatter plot of eddy scale versus deformation radius shows a correlation (see Figs. 21a and 24 of Stammer, 1997), to which a line of the form $R_{\max} = 0.8R_1 + 88$ km is fit. The scatter plot data indicates that a linear fit is only marginally justified, but nevertheless it provides a useful functional form, and demonstrates that the observed scales are everywhere larger than the deformation scale. The scale estimate as a function of deformation radius is converted to a function of latitude using the zonally averaged deformation radii computed in Section 2.

A second estimate of eddy scale can be computed from the analysis of Chelton *et al.* (2007), who use combined TOPEX/Poseidon, ERS-1 and ERS-2 data to track coherent vortices at the ocean surface. Coherent vortices are identified using a computation of the Okubo-Weiss parameter (W) from the geostrophic surface velocity field: closed contours of W with values less than $-2 \times 10^{-12}\ \text{s}^{-2}$ that last for more than 18 weeks are called coherent
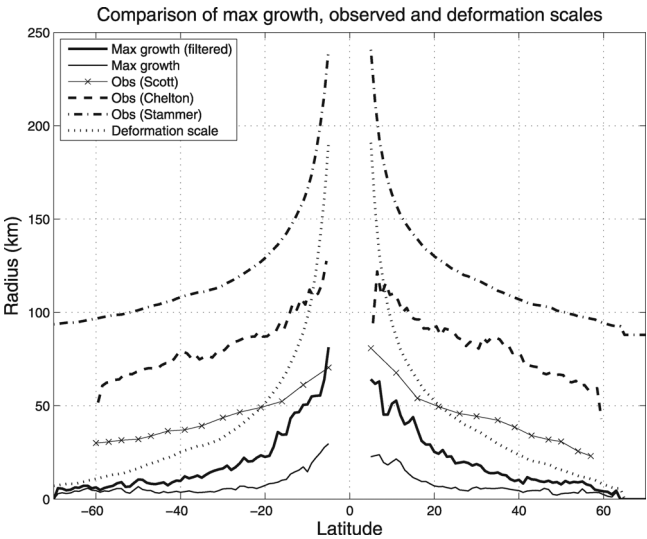
Figure 11. Zonally averaged scales (in km) of: maximum growth, filtered by condition that baroclinic conversion $G > 0.5$ mWm^{-2} (bold solid line); scale of absolute maximum growth (thin solid line); the inverse of the centroid wavenumber of surface kinetic energy spectra from analysis of satellite observations by R. Scott (line with x's); the radii of coherent vortices from the analysis of satellite observations by Chelton *et al.* (2007) (dashed line); the zero-crossing of the lag of the spatial autocorrelation function from the analysis of satellite observations by Stammer (1997) (dash-dotted line); and the first deformation radius (dotted line).

vortices, and for these they estimate their number, sizes and trajectories. Zonally averaged vortex radii data were provided by D. Chelton for the present study. A third estimate, provided by R. Scott (pers. comm.) is made by taking the inverse of the centroid of the spatial energy spectra, computed from TOPEX/Poseidon and ERS data on overlapping 10°by 10°boxes.

In Figure 11 we show the zonally averaged values for each eddy scale estimate, as well as the zonally averaged first deformation radii, and both filtered and unfiltered scales of fastest growth (see caption for additional details). Note that the estimates for eddy scale from observations differ, most likely for the following reason. The analysis by R. Scott provides a scale $R = 1/K_{\text{peak}}$, the spectral radius for the peak of the kinetic energy spectra. The analysis by Chelton *et al.*, on the other hand, provides actual radii[4] R_{vortex} of coherent vortices. The spectral transform of a field of coherent vortices will not, in general, yield a spectral peak at $K = 1/R_{\text{vortex}}$, and so we do not expect the two estimates to coincide exactly.

---

4. D. Chelton actually provided diameter values, and these were divided by 2

This discrepancy between observed eddy scales is significant, but even the smallest estimate is considerably larger than the scales of maximum linear growth, even when restricting the comparison to the energetically important instabilities at each location (the bold solid line); the scales of fastest growth are much smaller than the observed eddy scales. This is especially true at high latitudes. In the ACC, for example, deformation scales are of order 10–20 km, while the observed eddy scale is about 30–100 km. Given that the scale of maximum linear baroclinic growth is typically 1/4 the deformation radius, one can safely conclude that linear theory, while fairly accurate in predicting timescales and locations of eddy activity, fails to predict the spatial scale of the eddies. Nonlinear studies of baroclinic growth in the presence of vertical shears and density structures similar to those found in the ACC (Smith and Vallis, 2002) indicate that a strong inverse cascade should ensue, resulting in an eddy scale much larger than the scale of maximum linear growth. Thus a nonlinear cascade of energy seems necessary to explain the observed eddy statistics.

7. Discussion

Classical quasigeostrophic linear stability analysis, applied systematically to the global ocean hydrography of Gouretski and Koltermann (2004), has been shown to provide a good estimate of eddy timescales and locations of eddy activity, but a systematically too-small estimate of eddy space scales, by up to an order of magnitude at high latitudes. The discrepancy between spatial scales predicted by linear theory and the observed scales of eddies is minimized by selecting only those growing waves that can convert significant amounts of mean potential to eddy kinetic energy, but the discrepancy remains large.

Regional correlations can be appreciated by comparing Figure 7 of Stammer (1997) to Figure 1b (and 9a) in the present paper. Stammer's Figure 7 shows observed eddy kinetic energy as a function of horizontal location, filtered by spatial scale; panel (d), for instance, shows eddy kinetic energy for scales between 30 and 100 km. Given the close relation of the inverse Eady timescale (and hence growth rate) to the mean available potential energy, the degree of correlation between eddy kinetic energy and growth rate is a direct indication of the degree to which instabilities locally generate eddies. Some features in the figures are immediately apparent. Eddy kinetic energy is prominent at all scales in the ACC, Kuroshio and Gulf Stream, but strongest at large scale (30–100 km). Growth rates are also largest in these areas. Weaker eddy kinetic energy is more widespread, particularly in the western Pacific, mid-Atlantic and in the Indian Ocean, but these eddies are smaller scale—most of that weaker kinetic energy is restricted to spatial scales of less than 30 km. Consistent with this picture, we find weaker (but non-zero) growth rates throughout these areas. The smallest growth rates (and areas of zero significant energy conversion rate—see Fig. 8a) are also the areas of insignificant eddy kinetic energy. If one also considers the filtered scale of fastest growth in the present paper (Fig. 9b), it is apparent that in areas of weaker kinetic

energy, the spatial scale of the eddies is closer to the scale of maximum growth (but still generally larger).

The correlation between growth rate and eddy kinetic energy offers rationalization for the use of local closures in mesoscale eddy parameterization schemes for general circulation models. On the other hand, the results here indicate the need to include better estimates of eddy scale—the deformation scale estimate of Stone (1972), for example, should be least accurate where eddies are most important. Visbeck *et al.* (1997) find some success by using the Green (1970) estimate from the 'width of the baroclinic zone', but this is a dauntingly complex scale to estimate on-the-fly in an ocean general circulation model. The nonlinear prediction of Held and Larichev (1996), which assumes a Rhines mechanism for eddy scales, is $R_{eddy} \sim \sigma/\beta$ (almost at least—their estimate of timescale has the square-root outside the integral in Eq. 2.3). This parameterization has the advantage that it is larger where σ is large, but unfortunately predicts larger eddies at higher latitudes (in an absolute sense) due to the decrease of β with latitude. This is consistent with the analysis of Stammer (1998), who showed theHeld and Larichev (1996) theory to be a poor fit to surface data.

The dynamics that control the scale of observed ocean eddies, it must be concluded, are more complex than can be described by linear theory alone. This should not be surprising, since even weakly unstable states produce turbulence (Schneider, 2004, , for example, makes the argument that the atmosphere is thus characterized), while the oceanic mean state is clearly very unstable. The process of fully-developed baroclinic instability is not a linear process, but rather itself a turbulent process by which many scales are excited and cascades are catalyzed. Studies such as Held and Larichev (1996), Smith and Vallis (2002), Arbic and Flierl (2004) and Thompson and Young (2006) convincingly show that baroclinic instability in ocean-like environments leads to an inverse cascade of energy, and this is consistent with the observations of Scott and Wang (2005). If this is the right perspective, then a continuous feedback between the turbulence and the mean state must occur and be in balance with the large-scale forces that restore the mean to its unstable state. How this process might occur is taken up in detail by Smith and Marshall (2008), where it is found that such eddy-mean-forcing feedbacks are necessary to explain local observations in the ACC.

Instabilities with weak energetic conversion potential are suppressed in the primary analysis. These suppressed modes are typically due to instabilities of shears in or near the mixed layer, and may actually be quite important in understanding the interactions of the mixed layer with eddies, and in general the submesoscale dynamics of the upper ocean. For a more complete analysis of these issues, see Boccaletti *et al.* (2007) and Thomas *et al.* (2007). They may also generate a significant surface-trapped forward energy cascade and so may be an important part of the eddy energy cycle (see, e.g. Tulloch and Smith, 2006).

APPENDIX A

The mean geostrophic shear

The geostrophic velocity field for the entire GK04 dataset is computed as follows. Given the neutral density field $\bar{\rho} = \bar{\rho}(\lambda, \theta, z)$, where λ is longitude, θ is latitude, and z is depth ($z = 0$ is at the ocean surface, and z increases upwards), the hydrostatic pressure field is

$$P(\lambda, \theta, z) = g \int_z^0 \bar{\rho}(\lambda, \theta, z')dz' + p_0(\lambda, \theta)$$

where p_0 is the surface pressure. The numerical integral for the discrete neutral density field is calculated using Simpson's rule. The geostrophic velocity field is

$$u(\lambda, \theta, z) - u_0(\lambda, \theta) = -\frac{1}{\rho_0 f(\theta)} \frac{1}{a} \frac{\partial P}{\partial \theta},$$

$$v(\lambda, \theta, z) - v_0(\lambda, \theta) = \frac{1}{\rho_0 f(\theta)} \frac{1}{a \cos(\theta)} \frac{\partial P}{\partial \theta}$$

where $f(\theta) = 2\Omega \sin(\theta)$ and u_0 and v_0 arise from the horizontal derivatives of the unknown surface pressure p_0. The horizontal derivatives of the hydrostatic pressure field are estimated using centered finite differences applied directly to the 1/2 by 1/2 degree grid without interpolation (and so high-latitude estimates are slightly more accurate). The unknown velocity constants are irrelevant to the linear instability calculation, but for computational reasons are set at each location to remove the barotropic component of the flow, i.e.

$$u_0(\lambda, \theta) = -\int_z^0 \frac{1}{\rho_0 f(\theta)} \frac{1}{a} \frac{\partial P}{\partial \theta}(\lambda, \theta, z')dz',$$

$$v_0(\lambda, \theta) = \int_z^0 \frac{1}{\rho_0 f(\theta)} \frac{1}{a \cos(\theta)} \frac{\partial P}{\partial \lambda}(\lambda, \theta, z')dz'.$$

APPENDIX B

The vertical discretization

The discrete stretching operator. The vertical finite-difference grid used in all calculations is shown in Figure 12. On this grid, the discrete stretching operator Γ_{nm} is
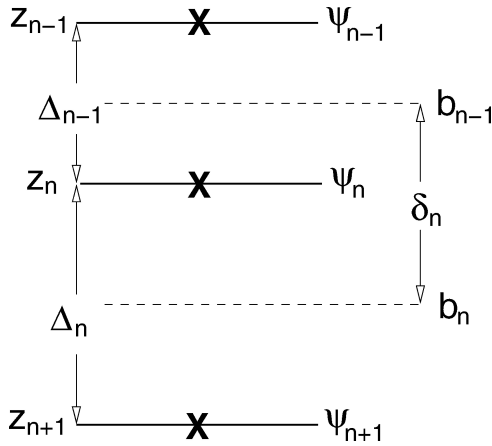
Figure 12. The grid used to represent vertical structure. Δ_n is the spacing between ψ_n and ψ_{n+1}, while δ_n is the distance between half spaces: $\delta_n = (\Delta_{n-1} + \Delta_n)/2$.

$$
\Gamma_{nm}\psi_m = \frac{f^2 \rho_0}{g} \begin{cases} \dfrac{1}{\delta_1}\left(\dfrac{\psi_2 - \psi_1}{\bar{\rho}_2 - \bar{\rho}_1}\right), & n = 1 \\[2ex] \dfrac{1}{\delta_n}\left(\dfrac{\psi_{n-1} - \psi_n}{\bar{\rho}_n - \bar{\rho}_{n-1}} - \dfrac{\psi_n - \psi_{n+1}}{\bar{\rho}_{n+1} - \bar{\rho}_n}\right), & n = 2 \,..\, \mathcal{N} - 1 \\[2ex] \dfrac{1}{\delta_\mathcal{N}}\left(\dfrac{\psi_{\mathcal{N}-1} - \psi_\mathcal{N}}{\bar{\rho}_\mathcal{N} - \bar{\rho}_{\mathcal{N}-1}}\right), & n = \mathcal{N} \end{cases} \tag{B.1}
$$

where $\bar{\rho}_n$ the background neutral density at level n, ρ_0 the average density and $\mathcal{N}$ is the total number of discrete levels. This discretization encompasses the boundary conditions (3.2b) in the manner specified by Bretherton (1966), where the infinitesimal distance below (above) the upper (lower) boundary in the δ-sheet representation is replaced by $\delta_1(\delta_N)$ in the discrete case. Therefore calculations using this discretization will have vertical errors of order δ and horizontal errors of order $N\delta/f$ (where δ is δ_1 or $\delta_\mathcal{N}$).

The approximate δ-sheet. That the above operator includes a discrete approximation of δ-sheets can be shown as follows. Referring to Figure 12, density values $\bar{\rho}$ are located at ψ points, and w and N^2 are located at b points (where $b = f\psi_z$ is the buoyancy). The QGPV equation is obtained by combining the vorticity equation

$$
\frac{D\zeta}{Dt} = f\frac{\partial w}{\partial z}
$$

and the buoyancy equation

$$
\frac{Db}{Dt} = -N^2 w,
$$

where $\zeta = \nabla^2 \psi$ and D/Dt is the horizontal advection operator. At the upper level in the discrete case,

$$\left.\frac{D\zeta}{Dt}\right|_1 = \left. f\frac{\partial w}{\partial z}\right|_1 = -f\frac{D}{Dt}\left[\frac{1}{\delta_1}\left(\frac{b_0}{N_0^2} - \frac{b_1}{N_1^2}\right)\right].$$

The boundary condition we wish to implement is $D/Dt\,(b_0/N_0^2) = 0$. Using this, the above equation can be combined to form

$$\frac{D}{Dt}\left(\zeta_1 - \frac{1}{\delta_1}\frac{f}{N_1^2}b_1\right) = 0 \longrightarrow \frac{D}{Dt}\left(\zeta - \frac{f}{N^2}b\bigg|_{-\epsilon}\delta(z)\right) = 0.$$

where the arrow denotes the continuous limit. Formally, in the continuous limit, $\epsilon \to 0$. In the discrete case, ϵ is replaced by the finite value δ_1.

Finally, using the grid specified in Figure 12,

$$\frac{b_1}{\delta_1 N_1^2} = \frac{f^2\rho_0}{g}\left(\frac{\psi_2 - \psi_1}{\rho_2 - \rho_1}\right)$$

just as implemented in (B.1). Similar arguments apply to the lower boundary.

Topographic slopes and the mean discrete PV gradient.   At each horizontal location, topographic slopes were calculated from the Smith and Sandwell global seafloor topography dataset (Smith and Sandwell, 1997). Specifically, at each point a $2° \times 2°$ section of topography, centered on the location of interest, is extracted from the dataset. A plane is then fit to the topography via linear regression, so that the bottom relief is estimated as

$$\eta(x, y) \simeq \eta_0 + \alpha^x x + \alpha^y y.$$

The discrete mean PV at the bottom level is then

$$Q_{\mathcal{N}} = \Gamma_{\mathcal{N}m}V^m x + (\beta - \Gamma_{\mathcal{N}m}U^m)y + \frac{f}{\Delta_{\mathcal{N}}}\eta.$$

Therefore

$$\nabla Q_{\mathcal{N}} = \left(\frac{f\alpha^x}{\Delta_{\mathcal{N}}} + \Gamma_{\mathcal{N}m}V^m\right)\hat{\boldsymbol{\imath}} + \left(\beta + \frac{f\alpha^y}{\Delta_{\mathcal{N}}} - \Gamma_{\mathcal{N}m}U^m\right)\hat{\boldsymbol{\jmath}}. \qquad (\text{B.2})$$

The discrete instability problem.   Letting $\hat{\psi}_i$ be the discrete amplitude vector, the generalized eigenvalue problem (3.2a) may be written

$$\omega B_{ij}\hat{\psi}_j = A_{ij}\hat{\psi}_j$$

where

$$B_{ij} = \Gamma_{ij} - K^2\delta_{ij},$$
$$A_{ij} = \left(kQ_y^m - \ell Q_x^m\right)\delta_{ijm} + (kU^m + \ell V^m)\delta_{inm}B_{nj},$$

and the δs are Kronecker tensors, equal to unity when all indices are equal, and zero otherwise. The discrete stretching operator is given by (B.1), and so the boundary conditions (3.2b) are included. The tensor products with δ in the second line generate diagonal matrices with the vectors $kQ_y^m - \ell Q_x^m$ and $kU^m + \ell V^m$ on the diagonals, respectively. In the discrete problem with $\mathcal{N}$ levels, there will be $\mathcal{N}$ eigenvectors $\hat{\psi}$ and eigenvalues ω. In order to avoid unnecessary complexification, we omitted the index for these above.

APPENDIX C

Necessary conditions for instability in arbitrarily oriented flow

The derivation of necessary conditions follows the standard route, but the resulting conditions allow more possibilities than in the zonal-flow case. Dividing (3.2a) by $\mathbf{K} \cdot \mathbf{U} - \omega$, multiplying by $\hat{\psi}^*$ and integrating in z from $-H$ to 0 gives

$$\int_{-H}^{0} \left(\frac{f}{N}\right)^2 |\hat{\psi}_z|^2 + K^2 |\hat{\psi}|^2 dz = \int_{-H}^{0} \frac{\Pi|\hat{\psi}|^2}{\mathbf{K} \cdot \mathbf{U} - \omega} dz + \left[\left(\frac{f}{N}\right)^2 \frac{\Lambda|\hat{\psi}|^2}{\mathbf{K} \cdot \mathbf{U} - \omega}\right]_{-H}^{0},$$

where (3.2b) was used to replace the $\hat{\psi}_z$ in the boundary term arising from the integral of $\hat{\psi}^*(\Gamma\hat{\psi})$. Since the left hand side is real, the imaginary part of the right-hand side must vanish, so

$$\omega_i \left\{ \int_{-H}^{0} \frac{\Pi|\hat{\psi}|^2}{|\mathbf{K} \cdot \mathbf{U} - \omega|^2} dz + \left[\left(\frac{f}{N}\right)^2 \frac{\Lambda|\hat{\psi}|^2}{|\mathbf{K} \cdot \mathbf{U} - \omega|^2}\right]_{-H}^{0} \right\} = 0.$$

Therefore, for nonzero ω_i, the term in braces must vanish identically. If $\ell = 0$, the result is exactly that of Charney and Stern. If $k = 0$, Charney-like instabilities involving Q_x require either V_z at the top be of the *same* sign, or V_z at the bottom be of opposite sign as Q_x. When k and ℓ are both non-zero, there are many ways for the term in braces to vanish. Also note that even in the absence of surface shears at the lower boundary, topographic slopes can yield instabilities through their interaction with either interior mean PV gradients or with upper-surface shears.

REFERENCES

Arbic, B. K. and G. R. Flierl. 2004. Effects of mean flow direction on energy, isotropy, and coherence of baroclinically unstable beta-plane geostrophic turbulence. J. Phys. Oceanogr., *34*, 77–93.

Boccaletti, G., R. Ferrari and B. Fox-Kemper. 2007. Mixed layer instabilities and restratification. J. Phys. Oceanogr., *37*, 2228–2250.

Bretherton, F. P. 1966. Critical layer instability in baroclinic flows. Quart. J. Roy. Meteor. Soc., *92*, 325–334.

Charney, J. 1947. Dynamics of long waves in a baroclinic westerly current. J. Meteor., *4*, 135–162.

Charney, J. G. and M. E. Stern. 1962. On the stability of internal baroclinic jets in a rotating atmosphere. J. Atmos. Sci., *19*, 159–172.

Chelton, D. B., R. A. deSzoeke, M. G. Schlax, K. E. Naggar and N. Siwertz. 1998. Geographical variability of the first baroclinic Rossby radius of deformation. J. Phys. Oceanogr., *28*, 433–460.

Chelton, Dudley B., Michael G. Schlax, Roger M. Samelson and Roland A. de Szoeke. 2007. Global observations of large oceanic eddies. Geophys. Res. Letts., *34*, L15606, doi: 10.1029/2007GL030812.

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

Gill, A. E., J. S. A. Green and A. J. Simmons. 1974. Energy partition in the large-scale ocean circulation and the production of mid-ocean eddies. Deep-Sea Res., *21*, 499–528.

Gouretski, V. V. and K. P. Koltermann. 2004. WOCE global hydrographic climatology, a technical report. Berichte des BSH, *35*, 1–52.

Green, J. S. A. 1960. A problem in baroclinic instability. Quart. J. Roy. Meteor. Soc., *86*, 237–251.

——— 1970. Transfer properties of the large-scale eddies and the general circulation of the atmosphere. Quart. J. Roy. Meteor. Soc., *96*, 157–185.

Held, I. M. and V. D. Larichev. 1996. A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta-plane. J. Atmos. Sci., *53*, 946–952.

Jackett, D. R. and J. T. McDougall. 1997. A neutral density variable for the world's oceans. J. Phys. Oceanogr., *27*, 237–263.

Killworth, P. D. and J. R. Blundell. 2007. Planetary wave response to surface forcing and to instability in the presence of mean flow and topography. J. Phys. Oceanogr. (in press).

Pedlosky, J. 1964. The stability of currents in the atmosphere and ocean. Part I. J. Atmos. Sci., *21*, 201–219.

——— 1984. The equations for geostrophic motion in the ocean. J. Phys. Oceanogr., *14*, 448–455.

——— 1987. Geophysical Fluid Dynamics, Springer, New York, 2nd edition.

Phillips, H. E. and S. R. Rintoul. 2000. Eddy variability and energetics from direct current measurements in the Antarctic Circumpolar Current south of Australia. J. Phys. Oceanogr., *30*, 3050–3076.

Phillips, N. A. 1954. Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. Tellus, *6*, 273–286.

Samelson, R. M. 1999. Note on a baroclinic analogue of vorticity defects in shear. J. Fluid Mech., *382*, 367–373.

Schneider, T. 2004. The tropopause and the thermal stratification in the extratropics of a dry atmosphere. J. Atmos. Sci., *61*, 1317–1340.

Scott, R. B. and F. Wang. 2005. Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry. J. Phys. Oceanogr., *35*, 1650–1666.

Smith, K. S. 2007. Eddy amplitudes in baroclinic turbulence driven by non-zonal mean flow: Shear dispersion of potential vorticity. J. Phys. Oceanogr., *37*, 1037–1050.

Smith, K. S. and J. C. Marshall. 2008. Evidence for deep eddy mixing in the Southern Ocean. J. Phys. Oceanogr., (submitted).

Smith, K. S. and G. K. Vallis. 2002. The scales and equilibration of mid-ocean eddies: Forced-dissipative flow. J. Phys. Oceanogr., *32*, 1699–1721.

Smith, W. H. F. and D. T. Sandwell. 1997. Global seafloor topography from satellite altimetry and ship depth soundings. Science, *277*, 1957–1962.

Spall, M. A. 2000. Generation of strong mesoscale eddies by weak ocean gyres. J. Mar. Res., *58*, 97–116.

Stammer, D. 1997. Global characteristics of ocean variability estimated from regional TOPEX/Poseidon altimeter measurements. J. Phys. Oceanogr., *27*, 1743–1769.

——— 1998. On eddy characteristics, eddy transports, and mean flow properties. J. Phys. Oceanogr., *28*, 727–739.

Stone, P. H. 1972. A simplified radiative-dynamical model for the static stability of rotating atmo-
  spheres. J. Atmos. Sci., *29*, 405–418.

Thomas, L. N., A. Tandon and A. Mahadevan. 2007. Submesoscale processes and dynamics, *in*
  Eddy-Resolving Ocean Modeling, M. Hecht and H. Hasumi, eds, AGU Monograph, Washington
  DC.

Thompson, A. F. and W. R. Young. 2006. Scaling baroclinic eddy fluxes: Vortices and energy balance.
  J. Phys. Oceanogr., *36*, 720–738.

Treguier, A. M., I. M. Held and V. D. Larichev. 1997. On the parameterization of quasigeostrophic
  eddies in primitive equation ocean models. J. Phys. Oceanogr., *27*, 567–580.

Tulloch, R. and K. S. Smith. 2006. A new theory for the atmospheric energy spectrum: Depth-limited
  temperature anomalies at the tropopause. Proc. Nat. Acad. Sci., *103*, 14690–14694.

Vallis, G. K. 2006. Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Cir-
  culation, Cambridge University Press, Cambridge, U.K.

Visbeck, M., J. Marshall, T. Haine and M. Spall. 1997. Specification of eddy transfer coefficients in
  coarse-resolution ocean circulation models. J. Phys. Oceanogr., *27*, 381–402.