Routes to Dissipation in the Ocean: The 2D/3D Turbulence Conundrum

Peter Müller, James McWilliams, Jeroen Molemaker

May 3, 2005

Turbulence research has reached a mature stage, even though at the most fundamental level turbulence itself remains a mystery. The phenomenology and physical mechanisms of the separate regimes of two- and three-dimensional, geostrophic, stratified, gravity-wave, and boundary-layer turbulence are all now fairly completely established. They describe well the nature of geophysical turbulence at various space and time scales and at various locations. However, an unsolved fundamental problem remains: how do these different regimes of turbulence coexist and connect in the ocean? How and where does the energy of the general circulation cascade from the large climatic scales, where most of it is generated, to the small scales, where all of it is dissipated? In particular, how is the dynamical transition made from an anisotropic, 2D-like, geostrophic cascade at large scales—with its strong inhibition of down-scale energy flux—to the more isotropic, 3D-like, down-scale cascades at small scales. This essay discusses these problems, highlighting the physical laws that prevent a simple down-scale energy cascade and the possible connections among the different regimes of geophysical turbulence.

1 Introduction and Overview

The general circulation of the ocean is driven by surface fluxes of momentum, heat, and fresh water at large, basin-wide horizontal scales on the order of 1000 km or larger and at long time scales of the order on 1 yr and longer. The circulation itself has comparably large and long scales, as well as intermediate ones associated with equatorial zonal and lateral boundary currents and various mesoscale instabilities. In steady state the large-scale input of variance into the velocity, temperature, and salinity fields of the circulation must be dissipated at small scales by molecular friction, heat conduction, and salt diffusion. For molecular friction the dissipation scale is the Kolmogorov scale, on the order of 10^{-2} m. (Table 1 gives definitions and typical oceanic values for the parameters discussed in this paper.) We presently do not know the processes by which these variances and energy cascade across the eight orders of magnitude separating the generation and dissipation scales. This cascade may be different for different quantities and in different regions. In the ocean there exist many smaller- and faster-scale motions, such as inertial and internal-gravity waves, that are generated at least partly on smaller and shorter scales. These motions must also cascade their variances down to the dissipation scales (*i.e.,* a down-scale cascade), thereby enhancing the overall dissipation rate. The down-scale cascade of these wave motions may or may not be coupled to that of the general circulation, and it is not yet clear how much these waves contribute to the dissipation of the general circulation.

The energy of the general circulation overwhelmingly resides in the (available) potential energy reservoir, and the kinetic energy of currents is only a small fraction. However, except for *gravitational instability*—due to a positive vertical gradient of potential density, which occurs only rarely in the ocean outside of the surface boundary layer—the down-scale transfer of energy occurs mainly through shear instabilities of various types and thus essentially involves the larger-scale kinetic energy. The total kinetic energy budget for the general circulation is dominated by its creation through wind work on large scales and destruction ultimately through its turbulent dissipation on very small scales, each at a level of about $\rho_0 H_0 \epsilon \sim 5 \times 10^{-3}$ W m^{-2} in column integral (or $\rho_0 \epsilon \sim 10^{-6}$ Wm^{-3} locally, hence $\epsilon \sim 10^{-9}$ m^{2}s^{-3}; see Table 1). In numerical general circulation models, this dissipation is effected through eddy viscosity terms in the interior and bottom boundary drag. There is a secondary contribution through conversion to potential energy (whose sign is even uncertain, but probably a net loss of kinetic energy) and various even lesser contributions (Oort *et al.*, 1994; Wunsch, 1998; Toggweiler & Samuels, 1998). Although wind work leads to available potential energy conversion by Ekman pumping at large scales, this seems to be nearly canceled by reverse conversion on the mesoscale. The first steps in the interior, down-scale energy cascade of the general circulation are the processes of *barotropic and baroclinic instability.* These generate mesoscale eddies with a spatial scale on the order of the internal Rossby radius of deformation, which has an average size of about 50 km. These instabilities thus provide down-scale cascades of energy from the basin-wide generation scale to the internal deformation radius. But any further down-scale cascade is inhibited by the fact that mesoscale eddies, like the general circulation, are very anisotropic flows, constrained by rotation and stratification to be nearly horizontal (*i.e.,* nearly 2D in this sense). These motions approximately satisfy hydrostatic and geostrophic momentum balances, and their dynamical evolution is governed by the potential-vorticity equation. The conditions that assure that this occurs are that the Rossby number Ro and the Froude number Fr, defined by

$$Ro \;=\; V/fL \quad \text{and} \quad Fr \;=\; V/NH, \tag{1}$$

are not too large. Here V is a characteristic horizontal velocity magnitude, L and H are characteristic horizontal and vertical shear scales, and f and N are characteristic Coriolis (Earth's rotation) and buoyancy (density stratification) frequencies.

The best known example of such balanced dynamics are the Quasigeostrophic Equations, but at larger Ro and Fr, geostrophic horizontal momentum balance changes to the more general gradient-wind balance. The advective dynamics of the Balance or Quasigeostrophic Equations (called *geostrophic turbulence*) generally cause an up-scale turbulent cascade of energy toward larger scales, hence away from dissipation by molecular viscosity at small scales. This inverse energy cascade occurs along with a down-scale cascade of enstrophy (the variance of potential vorticity) to its dissipation at small scales (Charney, 1971). This behavior is analogous to *2D turbulence*. In the enstrophy-cascading inertial range, Ro and Fr are invariant as the scale decreases, at least in the quasigeostrophic limit, $Ro, Fr \to 0$. There is much less experience with balanced turbulence at finite values of Ro and Fr compared with quasigeostrophic turbulence, but available solutions indicate that the former is only moderately more dissipative of energy (Yavneh *et al.*, 1997). It remains an open question whether finite Ro and Fr avoid increasing in the down-scale enstrophy cascade or whether they increase and thus generate inconsistencies with the underlying assumptions of momentum balance. Nevertheless, our present understanding is that balanced dynamics do not provide an efficient route to energy dissipation in the oceanic interior.

At smaller scales stratified shear flows become unstable once the Richardson number,

$$Ri \;=\; \frac{N^2}{(\partial \mathbf{u}/\partial z)^2}, \tag{2}$$

Figure 1: Conceptual diagram of energy flow for the general circulation: I. baroclinic instability; II. dissipation in boundaries layers; III. unbalanced instabilities; IV. interaction with inertia-gravity waves; V. breaking inertia-gravity waves.

falls below 0.25, or equivalently, once the Froude number $Fr \sim Ri^{-1/2}$ exceeds 2 (Miles, 1961; Howard, 1961). This Kelvin-Helmholtz instability is observed to set in at vertical scales on the order of 10 m and to have horizontal scales on the order of 50 m. This instability starts a down-scale energy cascade—called *stratified turbulence*—where the overturning eddies work against buoyancy, mix vertically, and increase the mean potential energy at the same time as they are being dissipated by molecular friction.

Asymptotic analysis of the regime of stratified turbulence as $Fr \to 0$, shows that the leading-order dynamics are equivalent to having independent 2D flows in each vertical layer (Lilly, 1983) and that the 3D structure is consistent with the Balanced Equations (McWilliams, 1985). This prompted the conjecture that there could be significant up-scale energy cascade in this regime, but many subsequent numerical solutions have shown that the great majority of the cascade is down-scale, primarily due to the growth of large vertical shear (*e.g.,* Herring & Metais, 1989). An important issue in the down-scale energy cascade in stratified turbulence is the relative fractions lost to dissipation and converted to potential energy. There are even examples (Holloway, 1983; Carnevale *et al.*, 2001) that show circumstances where the vertical buoyancy flux in stratified turbulence is positive, converting potential energy to kinetic energy and locally unmixing the water column. Once the down-scale cascade reaches the Ozmidov or buoyancy scale (on the order of 1 m), the stratification becomes unimportant, and it continues the cascade as *3D, isotropic turbulence*, down-scale through a Kolmogorov inertial range to dissipation. If the stratification ever becomes gravitationally unstable, then *convective turbulence* provides an efficient down-scale cascade to dissipation, but this rarely happens in the oceanic interior.

There is a scale gap of roughly three orders of magnitude between the anisotropic, balanced motions at large scales and the isotropic, small-scale motions that complete the down-scale energy cascade to dissipation (Table 1), which needs to be bridged for equilibrium kinetic-energy balance in the ocean. Phenomenologically, this gap is populated by intermediate-scale motions that include submesoscale vortices, fronts, vorticity and material filaments, inertia-gravity waves, and barotropic and baroclinic tides. All of these intermediate-scale motions are among the solutions of the Boussinesq Equations where the constraints imposed by rotation and stratification are felt less strongly. The flows become more ageostrophic, more non-hydrostatic, and more 3D isotropic as the scale decreases. Here we are interested in the processes that can transfer kinetic energy from the large scales across the intermediate scales to the small scales. We will consider three possible routes:

Inertia-Gravity Wave Route: Preexisting inertia-gravity waves interact with large- and mesoscale motions and catalyze a cascade of their energy to dissipation scales.

Instability Route: At sufficiently high Ro and Fr numbers, mesoscale balanced flows become unstable and transfer energy to unbalanced motions that cascade energy down-scale, through either the inertia-gravity wave spectrum or stratified turbulence or otherwise.

Boundary Route: Balanced-flow dissipation mainly occurs in surface and bottom boundary layers, either as a down-scale *boundary-layer turbulent cascade* or through inertia-gravity wave generation by flow over topography, with relatively little dissipation of balanced energy in the interior of the ocean.

The various routes to dissipation are shown in Figure 1, which depicts the possible pathways of energy from the large-scale general circulation to the small-scale flows that have a well known down-scale energy cascade toward viscous dissipation. Our focus is on the gap that exists between these large- and small-scale flow regimes. The diagram shows only the horizontal length scale of the dynamical regimes that are discussed. Associated with the various instabilities and cascades are also transformations of the vertical length scale H. It generally decreases with horizontal length scale L such that their aspect ratio is $O(f/N) \ll 1$ at large scales and O(1) at small scales. The two most important non-dimensional parameters are Ro and Fr, which also generally increase with decreasing H, L such that $Ro, Fr \ll 1$ at large scales, and $Fr = 2$, $Ro \gg 1$ at the scales of the Kelvin-Helmholtz instability.

None of these routes to dissipation has yet received strong theoretical or observational confirmation. But it is essential that we determine and understand the routes to dissipation. Simply inserting horizontal and vertical eddy viscosity coefficients in models of the oceanic general circulation and tuning their values to some desired outcome is hazardous and diminishes the fidelity of the model. Any circulation model that aspires to predictive capabilities must base its parameterizations on a dynamical understanding of the parameterized processes. In this paper we discuss each of these hypothesized routes in turn, but we stop short of attempting a quantitative partitioning of the requisite dissipation among them or of discussing their proper parameterization forms.

2 The Inertia-Gravity Wave Route

Inertia-gravity waves are ubiquitous in the ocean, and they are generated in a variety of ways. Boundary currents over topographic features in the stratified ocean can generate inertia-gravity waves that propagate into the interior. In particular, barotropic tidal flows across bottom topography generate internal-gravity waves of tidal frequency, the internal tides. Fluctuations in the atmospheric wind stress cause inertial oscillations in the surface mixed layer. Part of their energy leaks into the ocean interior as near-inertial internal-gravity waves, and boundary-layer turbulence directly generates inertia-gravity waves that propagate into the interior. Though not fully understood in detail, these energy inputs are believed to instigate a down-scale energy cascade by wave-wave interactions—called *gravity-wave turbulence*—until the waves reach the critical Richardson number $Ri = 0.25$ (at a wavenumber scale k_{KH}, typically $O(2 \times 10^2\ \mathrm{m}^{-1})$) and then break down. Then stratified and/or 3D, isotropic turbulence cascades the energy to its dissipation at molecular scales. It is still not entirely clear how much the inertia-gravity wave spectrum interacts with the large-scale circulation, but at least partly, it just coexists with an independent energy flux from generation to dissipation without influencing the energy pathways of the general circulation.

Conservation of Potential Vorticity

The exchange of momentum and energy between internal-gravity waves and geostrophic or more generally balanced flows is severely constrained by the conservation of potential vorticity. Potential vorticity is defined by

$$q = (2\boldsymbol{\Omega} + \nabla \times \mathbf{u}) \cdot \nabla b, \tag{3}$$

where $\boldsymbol{\Omega}$ is Earth's rotation vector; $\mathbf{u}$ the fluid velocity; b the buoyancy field, $b = g[1 - \rho/\rho_o]$; and ρ_o the mean density. For a hydrostatic, geostrophic flow, $\mathbf{u_g}$, the horizontal component of $\boldsymbol{\Omega}$ and the vertical velocity can be neglected. Hence,

$$q = (f + \zeta_g)N^2 - \frac{\partial v_g}{\partial z}\frac{\partial b_g}{\partial x} + \frac{\partial u_g}{\partial z}\frac{\partial b_g}{\partial y}, \tag{4}$$

where $\zeta_g = \partial v_g/\partial z - \partial u_g/\partial z$ is the vertical component of the geostrophic relative vorticity; f the local Coriolis frequency; and N^2 is the buoyancy frequency squared,

$$N^2 = \partial(\tilde{b} + b_g)/\partial z, \tag{5}$$

which includes both the background stratification $\tilde{b}$ and anomalies b_g associated with the geostrophic flow. The vertical geostrophic shear and the horizontal buoyancy gradients are related through the thermal-wind relations so that the geostrophic potential vorticity anomaly can be written, using (2), as

$$\delta q = \zeta_g N^2 + f\frac{\partial b_g}{\partial z} - f\left[\left(\frac{\partial u_g}{\partial z}\right)^2 + \left(\frac{\partial v_g}{\partial z}\right)^2\right] \sim fN^2(Ro - Fr^2), \tag{6}$$

demonstrating that potential vorticity anomalies are associated with both horizontal and vertical geostrophic shears as well as the product of anomalous stratification and Coriolis frequency.

On the other hand, internal-gravity waves have no potential-vorticity fluctuations associated with them since the fluctuation terms in (3) cancel. Conservation of potential vorticity restricts the nonlinear interactions: in the absence of molecular dissipation processes, internal-gravity waves cannot modify the potential vorticity on fluid parcels. This does not prevent an exchange of momentum and energy as the geostrophic flow redistributes its potential vorticity. These exchanges are reversible. With the inclusion of molecular processes, internal-gravity waves can affect the geostrophic flow irreversibly at sites of internal wave breaking, due to the enhancement of dissipation and mixing at these sites. Breaking can result from either random superposition or critical-layer absorption. These constraints are part of the so-called non-acceleration theorem (Eliassen and Palm, 1961; Andrews and McIntyre, 1976), which applies to a broad range of phenomena including all wave-mean flow interactions.

Critical-Layer Absorption

Consider the propagation of an internal-gravity wave group in a time-independent shear flow $u_g(z)$. The internal-wave Eulerian frequency ω and intrinsic frequency ω_0 are related by the Doppler shift,

$$\omega \;=\; \omega_0 + k_x u_g. \tag{7}$$

A wave propagation theory, based on a scale separation between the waves and their more slowly varying shear and stratification environment, asserts that

$$\frac{d\omega}{dt} \;=\; 0 \quad \text{and} \quad \frac{d\omega_0}{dt} \;=\; -k_x v_z \frac{du_g}{dz}, \tag{8}$$

where v_z is the vertical group velocity and d/dt is the total time derivative following the group velocity (Müller, 1976). The Eulerian frequency is constant because of the time-independence of the geostrophic flow. The intrinsic frequency changes as the wave group propagates and encounters different mean flow velocities and hence different Doppler shifts. Conservation of wave action implies that the action $\mathcal{N}$ of a wave group does not change, whereas its energy $\mathcal{E} \;=\; \mathcal{N}\omega_0$ changes according to

$$\frac{d\mathcal{E}}{dt} \;=\; -\mathcal{N} k_x v_z \frac{du_g}{dz}. \tag{9}$$

This energy exchange is reversible (*e.g.,* if the direction of the energy exchange is reversed because the wave group travels in the opposite direction, $v_z \to -v_z$, after reflection at the surface or bottom). The energy equation can also be written as

$$\mathcal{E}(t) \;=\; \frac{\omega_0(t)}{\omega_0(t_0)}\mathcal{E}(t_0). \tag{10}$$

The reversibility is broken when the wave group reaches a critical layer where it is dissipated. A critical layer will be encountered when the intrinsic frequency approaches the Coriolis frequency, $\omega_0(t) \to f$ (Bretherton, 1969). Note that this condition differs from the one for a classical, non-rotating, shear-flow critical layer, $\omega_0(t) \to 0$. A critical layer will thus be encountered once the wave group has transversed a velocity change $\delta u_g \;=\; (\omega_0(t_0)-f)/k_x$, differing again from $\delta u_g \;=\; c \;=\; \omega_0/k_x$ in the non-rotating case. While for much of the internal-wave spectrum in the ocean, $\omega_0 \gg f$ (so that this difference does not matter), a large fraction of the spectrum is near-inertial; it is this part of the spectrum that is most sensitive to critical-layer absorption.

Kunze and Müller (1989) consider an internal wave field radiating down from the surface into geostrophic flows. The spectrum of the waves is taken to be the Garrett-Munk spectrum (Munk, 1981), except for the assumed vertical asymmetry of only downward-propagating waves. As these waves propagate downward, they produce a net momentum-flux divergence due to the increasing fraction of the internal-wave spectrum lost to critical-layer absorption with depth. Specifically, they find

$$\frac{\partial\langle u'w'\rangle}{\partial z} \;=\; -c\frac{du_g}{dz}, \tag{11}$$

where $c \;=\; O(1.5\times10^{-5})\ \mathrm{ms}^{-1}$ for typical values of the buoyancy frequency and spectral energy level. One might be tempted to equate this momentum-flux divergence with an eddy-viscosity term, $A_v d^2u_g/dz^2$, and estimate an eddy viscosity, $A_v \;=\; -\langle u'w'\rangle/(du_g/dz) \approx 2\times10^{-2}\ \mathrm{m}^2\mathrm{s}^{-1}$. However, (11) is not an eddy-viscosity relation: the flux divergence is proportional to the gradient of the geostrophic flow and not its curvature. Even more disconcerting for our purposes is the fact that energy is transferred from the wave field to the balanced mean flow, a result that also follows directly from (10),

$$\mathcal{E}(t) \;=\; \frac{f}{\omega_0(t_0)}\mathcal{E}(t_0) \;\le\; \mathcal{E}(t_0). \tag{12}$$

As an internal-gravity wave approaches a critical layer, it loses some of its energy to the mean flow before the remainder gets dissipated in the critical layer.

Mixing-Length Theory

There are many routes to internal-wave dissipation. Before a wave group reaches its critical level, its amplitude rises to a level where non-linear wave-wave interactions can no longer be ignored. Wave-wave interactions also cascade energy to wave-breaking scales and are thus able to overcome the reversibility of inviscid, wave-mean flow interaction. Müller (1976) considers this case by applying mixing length arguments to the problem. Specifically, he considers a vertically symmetric wave spectrum, like the Garrett-Munk spectrum. When two waves of this spectrum with opposite vertical wavenumber at time $t = 0$, $k_z(t=0) = k_z^0$ and $k_z'(t=0) = -k_z^0$, propagate in a (constant) vertical shear, their wavenumbers change to

$$k_z(t) = -k_z^0 + \delta k_z \quad \text{and} \quad k_z'(t) = -k_z^0 + \delta k_z \neq -k_z(t), \tag{13}$$

where

$$\delta k_z = -t k_x du_g/dz. \tag{14}$$

A vertical shear hence forces an initially symmetric spectrum to become asymmetric. The asymmetric wave pair supports a vertical momentum flux,

$$\begin{aligned} \mathcal{F}_{xz} &= \mathcal{N} k_x v_z(k_z^0 + \delta k_z) + \mathcal{N} k_x v_z(-k_z^0 + \delta k_z) \\ &\approx 2\mathcal{N} k_x \frac{\partial v_z}{\partial k_z} \delta k_z. \end{aligned} \tag{15}$$

Nonlinear wave-wave interactions (Müller *et al.*, 1986), scattering at density fine structure (Mysak and Howe, 1976), and other processes have been shown to attenuate asymmetries. If this relaxation back to symmetry can be characterized by a relaxation time τ_R, then the shear-induced asymmetry is arrested at

$$\delta k_z = -\tau_R k_x du_g/dz, \tag{16}$$

resulting in a momentum flux,

$$\mathcal{F}_{xz} = -2\mathcal{N} k_x^2 \frac{\partial v_z}{\partial k_z} \tau_R \frac{du_g}{dz}, \tag{17}$$

proportional to the shear. The factor of proportionality defines the contribution of this wave pair to an internal-wave induced eddy viscosity. Integration over all possible pairs will give the overall eddy viscosity. This integration requires knowledge of the wavenumber dependences of $\mathcal{N}(\mathbf{k})$ and $\tau_R(\mathbf{k})$. Both dependencies were not well known when Müller (1976) made the original estimate with the rather large value of $A_v = 5 \times 10^{-1}$ m^{2}s^{-1}. Even today the nature and efficiency of the relaxation mechanisms are not sufficiently well established to arrive at reliable estimates. The important issue for the routes to dissipation is that the energy transfer from the large-scale flow to the internal wave field depends on the dissipation. In Müller's theory it depends on the relaxation time τ_R as a parameterization of dissipative effects. This dependence is a consequence of potential vorticity conservation: without dissipation there is no irreversible secular energy exchange between the mean flow and the internal-gravity waves. It might thus be that the wave field's route to molecular dissipation ultimately determines the energy cascade from balanced flows to waves. This is in contrast to 3D turbulence, where the down-scale energy flux is determined by the large-scale eddies and is largely independent of the value of the molecular viscosity. Reducing the latter merely pushes the Kolmogorov scale to smaller values and allows for finer scales in the velocity field.

Observational evidence of the internal wave-mean flow interaction is mixed. Estimates for internal-wave induced vertical eddy viscosities vary widely depending on location, time and other circumstances (*e.g.*, $A_v \leq 10^{-1}$ m^{2}s^{-1} [Frankignoul, 1976]; $\leq 10^{-2}$ m^{2}s-1 [Frankignoul & Joyce, 1979; Ruddick & Joyce, 1979]; $\approx 10^{-4}$ m^{2}s^{-1} [Lueck & Osborn, 1986; Peters *et al.*, 1988]), and they have not yet led to a coherent picture. The reason may be that the internal-wave field becomes horizontally isotropic and vertically symmetric when averaged over sufficiently large-space and long-time scales and thus has vanishing eddy viscosities. The possible anisotropies and asymmetries, which support horizontal and vertical momentum fluxes and their correlations with mean-flow gradients, appear only when short-time segments are considered, and their estimates have rather broad and often uncertain confidence limits.

3 The Instability Route

Balanced Dynamics

The essential basis for the approximation of balanced dynamics is velocity anisotropy. In a rotating, stratified fluid away from boundaries, the evolution from general initial conditions or forcing by the process called geostrophic (or balanced) adjustment leads to a local anisotropy with $u, v \gg w$, while radiating away inertia-gravity waves. Here z and w are the coordinate and velocity in the vertical direction, anti-parallel to gravity, and (x, y) and (u, v) are their horizontal counterparts. The condition for this to occur are that Ro and Fr are not too large. Under these conditions the vertical momentum balance is approximately hydrostatic,

$$\phi_z \;\approx\; b, \tag{18}$$

where $\phi \;=\; p/\rho_o$ is the geopotential function; the divergence of the horizontal momentum balance is approximately gradient-wind balance,

$$\nabla^2 \phi \;\approx\; \nabla \cdot f \nabla \psi + 2[\psi_{xx}\psi_{yy} - \psi_{xy}^2], \tag{19}$$

where ∇ is the horizontal gradient operator; and the horizontal velocity is weakly divergent and thus can be approximately represented by a streamfunction,

$$u \;\approx\; -\psi_y \quad \text{and} \quad v \;\approx\; \psi_x. \tag{20}$$

The consistent maximal simplification of the incompressible Boussinesq Equations consistent with these approximations and the conservation of either energy—in Cartesian coordinates, (x, y, z)—or enstrophy—in isopycnal coordinates, (X, Y, b)—is called the Balance Equations (Lorenz, 1960; Gent & McWilliams, 1984). The Balance Equations contain no inertia-gravity wave solutions, so they are often taken as a dynamical-systems model for the (advective) "slow manifold". They have been shown in many analyses to accurately represent the observed state and evolution of large-scale flows in the atmosphere and ocean. For example, they are often used for initialization of weather forecasts, even for hurricanes (Daley, 1991). An important reason for this accuracy is that inertia-gravity wave generation by balanced motions is known to be very weak when Ro and Fr are small but only somewhat weak when only Fr is small (Spall & McWilliams, 1992; Ford *et al.*, 2000).

Limits of Balance

An analysis for the solvability of the Balance Equations is made in Yavneh *et al.* (1997) and McWilliams *et al.* (1998). To be able to integrate forward in time from a balanced state, several necessary conditions must be satisfied everywhere within the domain. Where these are violated, there is a change of type of the Partial Differential Equation system, and the initial-and boundary-value problem becomes ill-posed because the problem of determining time-derivatives becomes singular and prevents further time integration. For the Balance Equations with isopycnal coordinates and $f \geq 0$, the conditions for loss of balance are the following:

1. Change of sign of vertical stratification, $N^2 = \partial b / \partial z$;
2. Change of sign of absolute vorticity, $A = f + \zeta^{(z)} = f + v_X - u_Y$ (where the horizontal derivatives denoted by capital letters are in isentropic coordinates);
3. Change of sign of $A - |S|$ (where $S^2 = (u_X - v_Y)^2 + (v_X + u_Y)^2$ is the variance of the horizontal strain rate).

None of these conditions can happen in the quasigeostrophic limit, since $A, A - |S| \to f + \mathcal{O}(Ro)$ and $N \to N_o + \mathcal{O}(Ro)$, where $N_o(z)$ is the resting-state stratification. Note the greater susceptibility of anticyclonic regions (*i.e.,* with $\zeta^{(z)}/f < 0$) in the second and third conditions. Furthermore, note the greater susceptibility to the third condition, since $A - |S| \leq A$. The first and second conditions also are related to the potential vorticity, $q = AN^2$. Since potential vorticity and buoyancy are conserved on parcels, except for mixing effects, there is an evolutionary inhibition for an unforced flow to spontaneously develop a violation of the first and second conditions. However, there is no such constraint with respect to the third condition, which indicates another sense in which there may be a greater susceptibility to the third condition. Thus, for Ro and Fr values of $O(1)$, especially in anticyclonic regions, the dynamical evolution cannot remain wholly balanced.

Unbalanced Instabilities

The only linear instabilities of a steady current in the quasigeostrophic limit are of the *inflection-point* type, requiring a change in sign of the mean potential vorticity gradient due to horizontal and/or vertical curvature in the flow (*i.e.,* barotropic and/or baroclinic). Inflection-point instability continues to occur at larger Ro and Fr values, where it can be expected to remain balanced over some range of these parameters; thus, its onset conditions have nothing to do with the limits of balance. In contrast, *gravitational (or convective) instability* occurs for unstable stratification, $N^2 < 0$ (Chandrasekhar, 1961), which is the first condition for loss of balance above. And *centrifugal (or inertial) instability* occurs for the second condition, $A < 0$, for axisymmetric or parallel mean flows (Ooyama, 1966; Hoskins, 1974). The growing modes for each of these latter instability types exhibit significant departures from balanced dynamics and lead to highly dissipative turbulent energy cascades, but they also have onset conditions that are rarely realized on the scales of the oceanic general circulation.

Motivated by the unfamiliar form of the third condition for loss of balance, $A - |S| < 0$, we have recently solved several rotating, stratified, shear-flow instability problems in the regime where it is satisfied and the first and second conditions are not satisfied (*i.e.,* $A, N^2 > 0$). In each of these problems we find an instability for anticyclonic, balanced mean flows with finite Ro, Fr in the neighborhood of the third condition for loss of balance. This instability has a growth rate, σ, which vanishes in the quasigeostrophic limit, roughly exponentially fast,

$$\sigma \;\propto\; e^{-c/|Ro|} \;\text{ as } Ro \to 0^-, \tag{21}$$

for some constant c independent of Ro. Three of these problems are for barotropic mean flows that do not satisfy the inflection-point criterion: elliptical flow in an unbounded domain (McWilliams & Yavneh, 1998); Taylor-Couette flow between co-rotating cylinders (Molemaker *et al.*, 2000; Yavneh *et al.*, 2001); and a boundary-trapped jet (McWilliams *et al.*, 2003). The fourth problem is the classical baroclinic instability problem (Eady, 1949; Stone, 1966, 1970; Nakamura, 1988), which we will use to illustrate this *anticyclonic ageostrophic* instability behavior. We believe it is generically likely to occur for the general circulation, albeit only weakly where $Ro, Fr \ll 1$.

Consider the linear instability of a mean zonal current $U(z)$ in geostrophic balance in a rotating (f), stably stratified (N), inviscid fluid. This is an anticyclonic flow since it has a negative potential vorticity anomaly relative to the rotating, stratified resting state. To help focus on the primary issue, we assume that f is spatially uniform, although effects from $f(y)$ are geophysically important for large-scale currents. Since our focus is departure from balanced dynamics, we will use the incompressible Boussinesq Equations as the fundamental fluid dynamics. This problem is a non-geostrophic, non-hydrostatic version of Eady (1949). The dynamics is governed by Ro and the ratio of stratification and rotation frequencies (f/N). This problem has been studied extensively, but recently new insight has been obtained from the perspective of loss of balance (Molemaker *et al.*, 2003). In the quasigeostrophic limit, the classic baroclinic instability is the only instability type. Increasing Ro values progressively opens the possibility for different types of unstable modes. The value $Ro = -1/\sqrt{2}$, corresponds to the previously discussed

$A - |S| = 0$ limit of the Balanced Equations, and indeed an *ageostrophic anticyclonic instability* occurs in this neighborhood. Although its onset does not occur with a sharp boundary, separating stable and unstable behavior, its growth rate becomes significant in relation to the classical mode for large values of $|Ro|$. As in the barotropic examples mentioned above, the growth rate disappears exponentially fast for $Ro \to 0$ and is absent in the Quasi-geostrophic Equations. While the maximum growth rates of the ageostrophic anticyclonic mode are smaller than the geostrophic growth rate in the calculated range of Ro, it is important to note that the classical mode is only unstable at length scales larger than the first baroclinic deformation radius, and at smaller scales the ageostrophic mode is the only unstable one. So we can call the latter a submesoscale instability in anticipation of how it is likely to occur in the ocean. Even larger values of $Ro < -1$ open up the possibility of centrifugal instability; this is also an unbalanced instability which has its fastest growth rate at very small spatial scales selected by molecular viscosity. The sequence of possible instabilities is summarized in Table 2. At even larger $Ro = -2$ (*i.e.,* $Ri = 1/4$), Kelvin-Helmholtz instability occurs, and as $Ro \to -\infty$ (*i.e.,* $Ri < 0$), gravitational instability occurs, but these regimes are beyond the regime in which balanced dynamics has any relevance.

The ageostrophic anticyclonic mode can be interpreted as a resonance between a neutral shear mode and an inertia-gravity wave mode; so it is a fundamentally unbalanced phenomenon, even for $|Ro| \ll 1$. Since this instability is the first unbalanced one encountered for increasing Ro, it could be an important mechanism for energy to leak from balanced to unbalanced dynamics. In Figure 1 this route to dissipation appears as the arrow III. Once a balanced flow reaches the ageostrophic anticyclonic instability wavenumber $k_{aa} = f/V\sqrt{2}$, there is energy transfer to smaller scales, leading to dissipation *via* Kelvin-Helmholtz instability, stratified turbulence, and 3D turbulence. The prevalence of the instability in an oceanic context is not established yet, nor is the subsequent evolution of the unstable flow; for example, does it continue to depart from balanced dynamics or relax back toward it, as some kind of a geostrophic adjustment process? The important point is that the ageostrophic anticyclonic instability offers a route to dissipation that puts less demanding requirements on the Ro and Fr numbers of the mesoscale balanced flows than do the classical gravitational, centrifugal, and Kelvin-Helmholtz instabilities. Still unknown are whether balanced turbulence increases its Ro and Fr values during the down-scale cascade, developing an increased susceptibility to unbalanced instability, and if so how efficient is the ensuing energy transfer to unbalanced motions.

4 The Boundary Route

Energy Flux To the Boundary

The instability and inertia-gravity wave routes extract energy from the mean flow in the interior of the ocean. There is also energy extraction in bottom boundary layers, either due to boundary-layer turbulent drag or due to inertia-gravity wave generation by scattering from flow over topography. Each of these near-boundary processes extracts its energy from the current near the bottom. Yet the bulk of the circulation's energy resides in the interior. Thus, in addition to uncertainties about the boundary energy-extraction rates (below), another difficulty with these boundary routes providing sufficient dissipation is that a large spatial energy flux toward the boundaries would be required to continuously deplete the interior energy reservoir and supply the boundary extraction processes. Presumably, this energy flux would take the form of a mean $p\mathbf{u}$ flux by the large- and mesoscale currents themselves. Since dynamically induced pressure fluctuations are small compared to ambient pressure values, field measurements of this flux are difficult and model solutions seem more accessible. As yet this $p\mathbf{u}$ has not been detected, which is a necessary step to assess the boundary route to dissipation.

Planetary Boundary Layers

Turbulent boundary layers dissipate energy at a rate $c_D V^3$ where c_D is the drag coefficient O(10^{-3}) and V the near-boundary interior velocity. For boundary-layer dissipation to account for all dissipation (*i.e.,* , $c_D V^3 \sim H_o \epsilon$), the needed boundary velocity is $V \sim 2 \times 10^{-1}$ ms^{-1}, which is much larger than usually found in the abyssal ocean.

Boundary Wave Generation

For the energy of the general circulation, an irreversible down-scale transfer may also occur at the bottom through generation of internal-gravity or lee waves generated by large-scale flow over topography, with subsequent propagation into the interior. Estimates of the internal wave energy flux in some locations (*e.g.,* Polzin *et al.*, 1997) are as high as the required average value of $\rho_0 H_0 \epsilon = 5 \times 10^{-3}$ Wm^{-2}, but only for the very rough topography occurring in parts of the Brazil Basin. This degree of topographic roughness occurs in only a fraction of the oceanic basins and may not be sufficiently widespread to achieve such a large flux for a global average. Furthermore, the principal cause of the internal-wave generation at this site is believed to be by barotropic tidal flows, which does not extract energy from the general circulation.

5 Summary

In steady state the general circulation of the ocean must, and somehow manages to, dissipate the kinetic energy generated by atmospheric forcing. In numerical models of the general circulation, this dissipation is accomplished to a significant degree through eddy viscosity coefficients that extract kinetic energy from the interior, large-scale circulation, implicitly a transfer to the unresolved subgrid scale and further on to molecular dissipation — along with other contributions from conversion to potential energy and boundary dissipation. Despite the widespread use of this conventional parameterization scheme, the underlying physical transfer mechanisms are not yet well understood. The conservation of potential vorticity severely restricts and constrains any simple down-scale energy cascade for balanced flows because an enstrophy cascade to smaller scales generally accompanies an energy cascade to larger scales. Interactions of the general circulation with inertia-gravity waves and other submesoscale, unbalanced motions, both in the oceanic interior and near the bottom, seem to offer the only significant available routes to dissipation. It is of great importance that we further investigate the ocean's routes to dissipation, so that we can trust the foundations and predictions of general circulation models.

Acknowledgments JCM and MJM received support from the National Science Foundation grant OCE 96-33681 and the Office of Naval Research Grant N00014-98-1-0165. PM received support from the office of Naval Research grant N00014-96-1-0989.

References

Andrews, D.G., & M.E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.* **33**, 2031-2048.

Bretherton, F.P., 1969: Momentum transport by gravity waves. *Quart. J. Roy. Met. Soc.* **95**, 213-243.

Carnevale, G.F., M. Briscolini, & P. Orlandi, 2001: Buoyancy- to inertial-range transition in forced stratified turbulence. *J. Fluid Mech.* **427**, 205-239.

Chandrasekhar, S., 1961: *Hydrodynamic and Hydromagnetic Stability.* Oxford University Press.

Charney, J.G., 1971: Geostrophic turbulence. *J. Atmos. Sci.* **28**, 1087-1095.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press.

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

Eliassen, A. & E. Palm, 1961: On the transfer of energy in stationary mountain waves. *Geofys. Publ.* **22**, 1-23.

Ford, R., M.E. McIntyre, & W.A. Norton, 2000: Balance and the slow quasi-manifold: some explicit results. *J. Atmos. Sci* **57**, 1236-1254.

Frankignoul, C., 1976: Observed interaction between oceanic internal waves and mesoscale eddies. *Deep-Sea Res.* **23**, 805-820.

Frankignoul, C., & T.M. Joyce, 1979: On the internal wave variability during the internal wave experiment (IWEX). *J. Geophys. Res.* **84**, 769-776.

Gent, P.R., & J.C. McWilliams, 1984: Balanced models in isentropic coordinates and the shallow water equations. *Tellus* **36A**, 166-171.

Herring, J.R., & O. Metais, 1989: Numerical experiments in forced stably stratified turbulence. *J. Fluid Mech.* **202**, 97-115

Holloway, G., 1983: A conjecture relating oceanic internal waves and small-scale processes. *Atmosphere-Ocean* **21**, 107-122.

Hoskins, B.J., 1974: The role of potential vorticity in symmetric stability and instability. *Q. J. Royal Met. Soc.* **100**, 480-482.

Howard, L., 1961: Note on a paper of John Miles. *J. Fluid Mech.* **10**, 509-512.

Kunze, E., & P. Müller, 1989: The effect of internal waves on vertical geostrophic shear. In: *Parameterization of small-scale processes*, 'Aha Huliko'a Hawaiian Winter Workshop, Hawaii Institute of Geophysics, Special Publication, 271-285.

Lilly, D.K., 1983: Stratified Turbulence and the Mesoscale Variability of the Atmosphere. *J. Atmos. Sci.* **40**, 749-761.

Lorenz, E.N., 1960: Energy and numerical weather prediction. *Tellus* **12**, 364-373.

Lueck, R.G., & T.R. Osborn, 1986: The dissipation of kinetic energy in a warm-core ring. *J. Geophys. Res.* **91**, 803-818.

McWilliams, J.C., 1985: A note on a uniformly valid model spanning the regimes of geostrophic and isotropic, stratified turbulence: balanced turbulence. *J. Atmos. Sci.* **42**, 1773-1774.

McWilliams, J.C., & I. Yavneh, 1998: Fluctuation growth and instability associated with a singularity of the Balance Equations. *Phys. Fluids* **10**, 2587-2596.

McWilliams, J.C., I. Yavneh, M.J.P. Cullen, & P.R. Gent, 1998: The breakdown of large-scale flows in rotating, stratified fluids. *Phys. Fluids* **10**, 3178-3184.

McWilliams, J.C., M.J. Molemaker, & I. Yavneh, 2003: Ageostrophic, anticyclonic instability of a barotropic boundary current. *Phys. Fluids*, submitted.

Miles, J., 1961: On the stability of homogeneous shear flows. *J.Fluid.Mech.*, **10**, 496-508.

Molemaker, M.J, J.C. McWilliams, & I. Yavneh, 2000: Instability and equilibration of centrifugally-stable stratified Taylor-Couette flow. *Phys. Rev. Lett.* **86**, 5270-5273.

Molemaker, M.J., J.C. McWilliams, & I. Yavneh, 2003: Ageostrophic baroclinic instability and loss of balance. *J. Phys. Ocean.*, submitted.

Müller, P., 1976: On the diffusion of momentum and mass by internal gravity waves. *J. Fluid Mech.* **77**, 789-823.

Müller, P., G. Holloway, F. Henyey, & N. Pomphrey, 1986: Nonlinear interactions among internal gravity waves. *Rev. Geophys.* **24**, 493-536.

Munk, W.H., 1981: Internal waves and small scale processes. In: *Evolution of Physical Oceanography*, B.A. Warren & C. Wunsch, Eds., The MIT Press, 264-291.

Mysak, L.A., & M.S. Howe, 1976: A kinetic theory for internal waves in a randomly stratified ocean. *Dyn. Atmos. Oceans* **1**, 3-31.

Nakamura, N., 1988: Scale selection of baroclinic instability—Effects of stratification and nongeostrophy. *J. Atmos. Sci.* **45**, 3253-3267.

Oort, A.H., L.A. Anderson, & J.P. Peixoto, 1994: Estimates of the energy cycle of the oceans. *J. Geophys. Res.* **99**, 7665-7688.

Ooyama, K., 1966: On the stability of the baroclinic circular vortex: a sufficient condition for instability. *J. Atmos. Sci.* **23**, 43-53.

Peters, H., M.C. Gregg, & J.M. Toole., 1988: On the parameterization of equatorial turbulence. *J. Phys. Ocean* **93**, 1199-1218.

Polzin, K.L., J.M. Toole, J.R. Ledwell, & R.W. Schmitt, 1997: Spatial variability of turbulent mixing in the abyssal ocean. *Science* **276**, 93-96.

Ruddick, B.R., & T.M. Joyce, 1979: Observations of interaction between the internal wave field and low-frequency flows in the North Atlantic. *J. Phys. Oceanogr.* **9**, 498-517.

Spall, M.A., & J.C. McWilliams, 1992: Rotational and gravitational influences on the degree of balance in the shallow-water equations. *Geophys. Astrophys. Fluid Dyn.* **64**, 1-30.

Stone, P.H., 1966: On non-geostrophic baroclinic instability, *J. Atmos. Sci.* **23**, 390-400.

Stone, P.H., 1970: On non-geostrophic baroclinic instability: Part II. *J.*

Atmos. Sci. **27**, 721-726.

Toggweiler, J.R., & B. Samuels, 1998: On the ocean's large-scale circulation near the limit of no mixing. *J. Phys. Ocean.* **28**, 1832-1852.

Wunsch, C., 1998: The work done by the wind on the oceanic general circulation. *J. Phys. Ocean.* **28**, 2332-2340.

Yavneh, I., A.F. Shchepetkin, J.C. McWilliams, & L.P. Graves, 1997: Multigrid solution of rotating, stably stratified flows: The balance equations and their turbulent dynamics. *J. Comp. Phys.* **136**, 245-262.

Yavneh, I., J.C. McWilliams, & M.J. Molemaker, 2001: Non-axisymmetric instability of centrifugally stable, stratified Taylor-Couette flow. *J. Fluid. Mech.* **448**, 1-21.

Table 1: Oceanic Parameters

Coriolis frequency	f	$7 \cdot 10^{-5}$ s^{-1}
Tidal frequency	M_2	$1.5 \cdot 10^{-4}$ s^{-1}
Buoyancy frequency	$N = \frac{1}{H_0}\int_{-H_0}^{0} dz\ N_0 e^{z/h}$, $N_0 = 5.2 \cdot 10^{-3}$ s^{-1}, $h = 1.3$ km, $H_0 = 5$ km	$2 \cdot 10^{-3}$ s^{-1}
Vertical scale, wavenumber	H, k_v	
Horizontal scale, wavenumber	L, k	
Horizontal velocity scale	V	$5 \cdot 10^{-2}$ ms^{-1}
Rossby number	$Ro = \frac{V}{fL}$	
Froude number	$Fr = \frac{V}{NH}$	
Richardson number	$Ri = \frac{N^2H^2}{V^2} = Fr^{-2}$	
Internal deformation radius	$R_d = \frac{N}{f} H_c$, $H_c = b = 1.3$ km	50 km
Ageostrophic anticyclonic instability wavenumber	$k_{aa} = \frac{1}{\sqrt{2}}\frac{f}{V}$	10^{-3} m^{-1}
Kelvin-Helmholtz wavenumbers	m_{KH} k_{KH}	10^{-1} m^{-1} $2 \cdot 10^{-2}$ m^{-1}
Kinetic energy dissipation rate	ϵ $\rho_0\epsilon$ column integral global integral	10^{-9} m^{2}s^{-3} 10^{-6} Wm^{-3} $5 \cdot 10^{-3}$ Wm^{-2} $15 \cdot 10^{11}$ W
Ozmidov (buoyancy)wavenumber	$k_O = \left(\frac{N^3}{\epsilon}\right)^{1/2}$	3 m^{-1}
Kolmogorov (dissipation) wavenumber	$k_K = \left(\frac{\epsilon}{\nu^3}\right)^{1/4}$	10^2 m^{-1}
Molecular viscosity	ν	10^{-6} m^{2}s^{-1}

Table 2: Instability Regimes and Anticyclonic (negative) Rossby Number

Ro	Regime	Instability
$\to 0$	quasigeostrophy	inflection-point baroclinic
$-1/\sqrt{2}$	loss of balance: $A - \|S\| \; = \; 0$	ageostrophic anticyclonic
-1	loss of balance: $A \; = \; 0$	centrifugal