Chapter 4

ENERGETICS OF THE GLOBAL OCEAN:
THE ROLE OF MESOSCALE EDDIES

HIDENORI AIKI
Application Laboratory,
Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

XIAOMING ZHAI
Centre for Ocean and Atmospheric Sciences,
School of Environmental Sciences, University of East Anglia, Norwich, UK

RICHARD J. GREATBATCH
GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany

This article reviews the energy cycle of the global ocean circulation, focusing on the role of baroclinic mesoscale eddies. Two of the important effects of mesoscale eddies are: (i) the flattening of the slope of large-scale isopycnal surfaces by the eddy-induced overturning circulation, the basis for the Gent-McWilliams parametrization; and (ii) the vertical redistribution of the momentum of basic geostrophic currents by the eddy-induced form stress (the residual effect of pressure perturbations), the basis for the Greatbatch-Lamb parametrization. While only point (i) can be explained using the classical Lorenz energy diagram, both (i) and (ii) can be explained using the modified energy diagram of Bleck as in the following energy cycle. Wind forcing provides an input to the mean KE, which is then transferred to the available potential energy (APE) of the large-scale field by the wind-induced Ekman flow. Subsequently the APE is extracted by the eddy-induced overturning circulation to feed the mean KE, indicating the enhancement of the vertical shear of the basic current. Meanwhile, the vertical shear of the basic current is relaxed by the eddy-induced form stress, taking the mean KE to endow the eddy field with an energy cascade. The above energy cycle is useful for understanding the dynamics of the Antarctic Circumpolar Current. On the other hand, while the source of the eddy field energy has become clearer, identifying the sink and flux of the eddy field energy in both physical and spectral space remains major challenges of present-day oceanography. A recent study using a combination of models, satellite altimetry, and climatological hydrographic data shows that the western boundary acts as a “graveyard” for the westward-propagating eddies.

1. Introduction

The mechanical energy input to the ocean by atmospheric winds is a major energy source for driving the large-scale ocean circulation and maintaining the abyssal stratification. The wind power input to the general circulation in the global ocean is estimated to be \( O(1) \) TW (1 TW = \( 10^{12} \) watts, Wunsch, 1998; Wunsch and Ferrari, 2004; Hughes and Wilson, 2008, Scott and Xu, 2009), with a significant fraction being supplied by the synoptic winds (Zhai et al., 2012). It is now clear that regions where power is actually injected into the geostrophic
circulation can differ significantly from those places where power is generated, owing to the lateral energy transfer in the Ekman layer (Kuhlbrodt et al. 2007; Roquet et al. 2011). The energy flux into the geostrophic interior can be viewed as work done by the Ekman flow against the horizontal pressure gradient, thereby helping to maintain, by means of Ekman pumping/suction, the observed thermocline structure and sea surface height anomaly associated with the gyre circulations. The available potential energy (APE) built up by the large-scale Ekman pumping/suction of the main thermocline is subsequently released by the generation of mesoscale eddies through baroclinic instability of the mean flow (Charney 1947; Eady 1949; Gill et al. 1974).

The ocean general circulation models (OGCMs) used for climate projections do not, in general, resolve mesoscale eddies explicitly, requiring that their effect be parametrized. An important impact of mesoscale eddies is an eddy-induced transport (analogous to the Stokes drift associated with surface gravity waves). There are two alternative approaches to parametrizing this transport. The first approach follows Gent and McWilliams (1990; hereafter GM90) and Gent et al. (1995), and is referred to as the tracer approach. In this approach, the momentum equation carried by a model solves for the development of the Eulerian mean velocity, with an eddy-induced advection term being added to the tracer equation of the model. This eddy-induced, additional velocity is parametrized in such a way as to release the APE stored in the large-scale density field, mimicking the principal aspect of baroclinic instability. A basis for the tracer approach is the classical energy diagram of Lorenz (1955) that is derived from equations averaged in \( z \) coordinates.

The second approach follows Greatbatch and Lamb (1990; hereafter GL90) and Greatbatch (1998), and is referred to as the momentum approach. In this approach, the momentum equation carried by a model solves for the development of the total transport velocity (i.e. the sum of the Eulerian mean velocity and the eddy-induced additional velocity), with an eddy-induced vertical stress term being added to the momentum equation of the model (with no additional vertical diffusion applied for tracers). This eddy-induced additional vertical stress is parametrized in such a way as to vertically redistribute the geostrophic momentum of the large-scale current, mimicking another aspect of baroclinic instability. A basis for the momentum approach is the energy diagram of Bleck (1985) that is derived from equations averaged in density coordinates, and is fundamentally different from the classical energy diagram of Lorenz (1955).

The additional vertical stress term in the GL90 parametrization represents the layer thickness form stress (the residual effect of pressure perturbations) that originates from the thickness-weighted-mean (TWM) momentum equations in density coordinates (Andrews 1983; de Snoo and Bennett 1993). The layer thickness form stress is closely related to theories for the dynamics of the Antarctic Circumpolar Current (ACC) that apply in the latitude band of the Drake Passage, where there are no continental boundaries available to establish the traditional Sverdrup balance (cf. Nowlin and Klink 1986; Gnanadesikan and Hallberg 2000; Hughes and De Cuevas 2001). It is generally thought that the layer thickness form stress

---

1The advection term of the tracer equations of the model is written in terms of the total transport velocity (i.e. the sum of the Eulerian mean velocity and the eddy-induced additional velocity). It should be noted that the total transport velocity is available as a prognostic quantity in the model adopting the momentum approach. See Eqs. (4) and (8b).
is responsible for transferring the wind-induced momentum from the upper layers of the ocean to the bottom layers so that in the zonal mean the wind stress at the surface is balanced by the topographic form stress associated with the pressure difference across ridges (cf. Johnson and Bryden 1989; Rintoul et al. 2001; Olbers and Visbeck 2005). Using the GL90 parametrization, the layer thickness form stress at the middepths of the Southern Ocean may be scaled (see the end of Appendix A) as \( (0.4 \text{ m}^2/\text{s})(1000 \text{ kg/m}^3)(1.0 \text{ m/s})/(2000 \text{ m}) = 0.2 \text{ N/m}^2 \), which is of the same order of magnitude as the zonally averaged wind stress over the Southern Ocean. Locally, however, the layer thickness form stress can be much larger, as found in a diagnosis of the output of an eddying OGCM simulation by Aiki and Richards (2008). These authors found that the layer thickness form stress in the region of the Drake Passage was as large as \( 4.0 \text{ N/m}^2 \). One way to examine the above-mentioned theory for the dynamics of the ACC, as well as the machinery of the GL90 parametrization, is to use the energy diagram of Bleck (1985) which contains an energy conversion path associated with the layer thickness form stress. This energy conversion path is not explicitly contained in the classical energy diagram of Lorenz (1955), and is not suitable for examining the role of the layer thickness form stress in the ACC.

Since the 2000s, the output of a series of eddy-resolving simulations for the global ocean circulation has become available to the community. However, identifying the life cycle of mesoscale eddies remains one of the major challenges of present-day oceanography. While the source of the eddy field energy has become clearer, the sink and flux of the eddy field energy are yet to be characterized in both physical and spectral space (e.g. Vallis and Hua 1987; Klein et al. 2008; Qiu et al. 2008). The eddy energy flux in physical space is closely related to the westward motion of planetary eddies (e.g. Yamagata 1982; Williams and Yamagata 1984; Cushman-Roisin et al. 1990; Eden et al. 2007), with the consequence that the western boundary is a major sink for eddy energy (Zhai et al. 2010), suggesting that this region is also one of enhanced diabatic mixing in the ocean (e.g. Tandon and Garrett 1996; Eden and Greatbatch 2008; Walter et al. 2005; Stober et al. 2009).

This article is organized as follows. In Sec. 2 we explain the energy cycle for large-scale flows in the global ocean, using the energy diagram of Bleck (1985), which includes the role of layer thickness form stress. In Sec. 3 we explain (the state-of-the-understanding regarding) the sources, sinks, and fluxes of the mesoscale eddy energy, using the energy diagram of Lorenz (1955). Section 4 provides a summary. The alternative use of the two energy diagrams in Secs. 2 and 3 is due to a tradeoff between the physical convenience (to distinguish the adiabatic and diabatic processes) of the TWM framework in density coordinates and the mathematical convenience of the Eulerian mean framework in \( z \) coordinates.

2. Maintenance of the Mean Field Energy

Despite several theoretical studies in various research areas of atmosphere and ocean dynamics (cf. Rhines and Young 1982; Andrews 1983; Johnson and Bryden 1989), the vertical redistribution of momentum by layer thickness form stress and associated energy conversions have been little investigated based on output from high-resolution OGCMs. This is more or less a result of the four-box energy diagram of Lorenz (1955) being exclusively used in previous studies. For example, the pioneering paper by Holland and Lin (1973) included a diagnosis of the energetics of eddies and wind-driven circulation as simulated by a two-layer model. Holland and Lin (1975) used an energy diagram which is similar to that of Lorenz (1955) and hence there is no term representing the layer thickness.
form stress. As will be shown in this section, it is possible to revise the definition of the mean and eddy kinetic energies (KEs) in Holland and Lin (1975), with the consequence that the revised energy diagram involves an energy conversion term representing the role of the layer thickness form stress.

2.1. Revise the Energy Diagram of Holland and Lin (1975)

We consider a standard two-layer model consisting of incompressible water of uniform density in each layer. Let $h_i$ be the thickness of each layer ($i = 1$ and $2$ represent the upper and lower layers respectively), and $\mathbf{V}_i = (U_i, V_i)$ be the horizontal velocity vector in each layer. Table 1 presents a list of the symbols used in the text. The sea surface is assumed to be rigid and the bottom depth $H_i(x, y) > 0$ may vary in the horizontal space, and hence $b_1 + b_2 = H_i$ (Fig. 1). APE and KE in each layer are given by

$$P = \frac{\rho_0}{H} (h_1 - h_2)^2, \quad (1a)$$

$$K_i = \frac{\rho_0}{2} h_i |V_i|^2, \quad (1b)$$

where $\rho_0$ is the reference density of sea water, $g^* = g(\rho_s - \rho_1)/\rho_0$ is the reduced gravity.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{h}$</td>
<td>Thickness of $i$th layer ($&gt; 0$)</td>
</tr>
<tr>
<td>$\bar{\mathbf{A}}$</td>
<td>Unweighted time mean in $i$th layer</td>
</tr>
<tr>
<td>$\bar{\mathbf{A}}'$</td>
<td>Thickness-weighted time mean in $i$th layer</td>
</tr>
<tr>
<td>$\mathbf{A}'' = \mathbf{A} - \bar{\mathbf{A}}$</td>
<td>Deviation from the unweighted mean, compared in $i$th layer ($\mathbf{A}'' = 0$)</td>
</tr>
<tr>
<td>$\mathbf{A}''' = \mathbf{A} - \bar{\mathbf{A}}$</td>
<td>Deviation from the thickness-weighted mean, compared in $i$th layer ($\mathbf{A}''' = 0$)</td>
</tr>
<tr>
<td>$\mathbf{A}_t = \partial \mathbf{A}/\partial t$</td>
<td>Time gradient in $i$th layer</td>
</tr>
<tr>
<td>$\nabla_i \mathbf{A}$</td>
<td>Lateral gradient in $i$th layer</td>
</tr>
<tr>
<td>$\tau_{\rho} = \partial h/\partial \rho$</td>
<td>Thickness in density coordinates ($&lt; 0$)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Unweighted time mean in density coordinates</td>
</tr>
<tr>
<td>$\bar{\tau}$</td>
<td>Thickness-weighted time mean in density coordinates</td>
</tr>
<tr>
<td>$\mathbf{A}''' = \mathbf{A} - \bar{\mathbf{A}}$</td>
<td>Deviation from the unweighted mean, compared at fixed $\rho$ ($\mathbf{A}''' = 0$)</td>
</tr>
<tr>
<td>$\mathbf{A}'''' = \mathbf{A} - \bar{\mathbf{A}}$</td>
<td>Deviation from the thickness-weighted mean, compared at fixed $\rho$ ($\mathbf{A}'''' = 0$)</td>
</tr>
<tr>
<td>$\mathbf{A}_x = \partial \mathbf{A}/\partial x$</td>
<td>Time gradient in density coordinates $</td>
</tr>
<tr>
<td>$\nabla_i \mathbf{A}$</td>
<td>Lateral gradient in density coordinates $\mathbf{A} = \nabla_i \mathbf{A} + (\nabla_i \mathbf{A})_{h_i}$</td>
</tr>
<tr>
<td>$\mathbf{A}$</td>
<td>Eulerian time mean in $z$ coordinates</td>
</tr>
<tr>
<td>$\mathbf{A} = \mathbf{A} - \bar{\mathbf{A}}$</td>
<td>Deviation from the Eulerian mean, compared at fixed $z$ ($\mathbf{A}^1 = 0$)</td>
</tr>
<tr>
<td>$\mathbf{A}_x = \partial \mathbf{A}/\partial x$</td>
<td>Time gradient in $z$ coordinates</td>
</tr>
<tr>
<td>$\nabla_i \mathbf{A}$</td>
<td>Three-dimensional gradient in $z$ coordinates</td>
</tr>
<tr>
<td>$\nabla_i \mathbf{A} = (\nabla_i \mathbf{A}, \partial \mathbf{A}/\partial z)$</td>
<td>Three-dimensional gradient in $z$ coordinates</td>
</tr>
<tr>
<td>$\mathbf{v}$</td>
<td>Vertical component of velocity</td>
</tr>
<tr>
<td>$\mathbf{H}$</td>
<td>Horizontal component of velocity</td>
</tr>
<tr>
<td>$u, v$</td>
<td>Vertical component of velocity</td>
</tr>
<tr>
<td>$w$</td>
<td>Horizontal component of velocity</td>
</tr>
<tr>
<td>$\mathbf{V}$</td>
<td>Three-dimensional velocity</td>
</tr>
<tr>
<td>$\mathbf{V} = (V, \mathbf{w})$</td>
<td>Three-dimensional velocity</td>
</tr>
<tr>
<td>$\mathbf{V}(x, y, z)$</td>
<td>Three-dimensional velocity</td>
</tr>
<tr>
<td>$\mathbf{V} = (V, \mathbf{w})$</td>
<td>Total transport velocity</td>
</tr>
<tr>
<td>$\mathbf{V} = (V, \mathbf{w})$</td>
<td>Bolus velocity $= (V, \mathbf{w})$</td>
</tr>
<tr>
<td>$\mathbf{V} = (V, \mathbf{w})$</td>
<td>Quasi-Stokes velocity $= (V, \mathbf{w})$</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>Global background density ($\mathbf{V} \rho_s = 0$)</td>
</tr>
<tr>
<td>$\rho_0$</td>
<td>Global background density ($\mathbf{V} \rho_s = 0$)</td>
</tr>
<tr>
<td>$p_g$</td>
<td>Global background density ($\mathbf{V} \rho_s = 0$)</td>
</tr>
<tr>
<td>$p = \int_0^h g(\rho - \rho_1) , dz$</td>
<td>Hydrostatic pressure</td>
</tr>
<tr>
<td>$\delta = \int_0^h g(\rho - \rho_1) , dz + g_0 h = \rho \partial_\mathbf{A} + g_0 \mathbf{A} = \nabla \delta + g_0 \mathbf{A}$</td>
<td>Montgomery potential $= \nabla \delta + g_0 \mathbf{A}$</td>
</tr>
</tbody>
</table>
associated with the difference of density $\rho_i$ in the upper and lower layers, and $h_i$ is the reference thickness of the upper layer. As in Holland and Lin (1975), we use a low-pass time filter to decompose the layer thickness into the mean and perturbation components, resulting in the mean and eddy APEs being defined as

$$h_i \equiv \bar{h}_i + h_i''',$$  

$$P_i = \frac{\rho_0}{2} g^* (H_1 - \bar{h}_1)'^2 + \sum_{p=0}^{P_{max}} \frac{\rho_0}{2} h_i'''V_i''^2,$$  

(2a)

(2b)

(for $i = 1, 2$), where the overbar and the triple prime symbols denote the time filter and the associated deviation in each layer. On the other hand, KE in each layer is a third-moment quantity, as shown by Eq. (1b). Central to our argument is the question of how to separate KE into that associated with mean and eddy components. Holland and Lin [1975; see their Eq. (14)] defined the mean KE as $(\rho_0/2)\bar{h}_iV_i|^2$ and the eddy KE as $(\rho_0/2)(h_i''V_i''|^2 - \bar{h}_iV_i''|^2)$, the latter of which is not a positive-definite expression. A clearer definition of the mean and eddy KEs is given by Bleck (1985) as follows:

$$V_i \equiv \bar{V}_i + V_i'',$$

$$K_i = \frac{\rho_0}{2} \bar{h}_iV_i'\bar{V}_i'^2 + \sum_{p=0}^{P_{max}} \frac{\rho_0}{2} h_i''V_i''^2,$$  

(3a)

(3b)

(for $i = 1, 2$), where $\bar{A}$ is the TWM and $A'' \equiv A - \bar{A}$ is the associated deviation of an arbitrary quantity $A$. It should be noted that each of the mean and eddy KEs in (3b) is clearly a positive-definite quantity. The TWM velocity $\bar{V}_i \equiv \bar{h}_i\nabla_i\bar{V}_i$ is written as the sum of the unweighted mean velocity $\bar{V}_i$ and the so-called bolus velocity

---

*The mathematical symbols of the present study are also applicable to equations in density coordinates for a continuously stratified fluid (Table 1). It should be noted that the deviation from the unweighted mean in density coordinates ($A'' \equiv A - \bar{A}$ compared at constant density) is slightly different from the deviation from the Eulerian mean in $z$ coordinates ($A'' \equiv A - \bar{A}$ compared at constant depth, to be used in the next section) for an arbitrary quantity $A$.\"
\( \mathbf{V}^2 = \mathbf{h}_i \cdot \mathbf{v}_i \),
\( \hat{\mathbf{V}_i} = \mathbf{v}_i + \mathbf{V}_i^0 \).

(4)

Usually, the unweighted mean velocity represents the large-scale geostrophic circulation as well as the Ekman flow at the top and bottom of the ocean, with the bolus velocity representing the eddy-induced circulation. Equation (3b) states that the mean KE, as defined by \( \rho_0/2 \mathbf{V}_j^2 \), includes the effect of the bolus velocity. Because of this modified definition for the mean and eddy KEs, the energy diagram shown below is different from the Lorenz (1955) energy diagram.

2.2. Exact Energy Equations for a Two-Layer Ocean Model

Equations for the thickness and the horizontal velocity vector in each layer are written by

\[ \partial_t h_1 + \nabla \cdot (h_i \mathbf{V}_i) = 0, \]

(5a)

\[ \rho_0(\partial_t h_1 + \mathbf{V}_1 \cdot \nabla \mathbf{V}_1 + \rho_0 g \hat{\mathbf{v}} \times \mathbf{V}_1) = -\nabla \phi_1, \]

(5b)

\[ \phi_2 = \phi_1 + \rho_0 g^2 (H_1 - h_1), \]

(5c)

where \( \partial_t \) and \( \nabla \) are the temporal and lateral derivative operators in each layer (i.e. operators in density coordinates), \( f \) is the Coriolis parameter, and \( \phi_1 \) is the anomaly of the Montgomery potential (MP) in each layer. Using Eqs. (5a)–(5c), one can derive equations for the APE, KE, and the MP flux divergence in each layer;

\[ \partial_t P = -\rho_0 g^2 (H_1 - h_1) \partial_t h_1 \]

(6a)

\[ = \phi_1 \partial_t h_1 + \phi_2 \partial_t h_2, \]

\[ \partial_t K_i + \nabla \cdot (\mathbf{V}_i K_i) = -h_i \mathbf{v}_i \cdot \nabla \phi_i, \]

(6b)

\[ \nabla \cdot (\phi_2 h_i \mathbf{V}_i) = -\partial_t \partial_t h_i + h_i \mathbf{V}_i \cdot \nabla \phi_i, \]

(6c)

where Eq. (5c) and \( \partial_t h_1 = -\partial_t h_2 \) (the rigid lid approximation) have been used to derive (6a). All terms on the right hand side cancel out once Eqs. (6a)–(6c) are summed to yield a conservation equation for \( P + K_1 + K_2 \).

We now consider equations for the general circulation. Using (5a), we rewrite the momentum equation (5b) to a flux divergence form:

\[ \rho_0(\partial_t (h_i \mathbf{V}_i) + \nabla \cdot (h_i \mathbf{V}_i \mathbf{V}_i) + f \mathbf{v}_i \times h_i \mathbf{V}_i) = -h_i \nabla \phi_i. \]

(7)

Application of a low-pass temporal filter to Eqs. (5a) and (7) yields

\[ \partial_t \hat{h}_i + \nabla \cdot (\hat{h}_i \hat{\mathbf{V}_i}) = 0, \]

(8a)

\[ \rho_0[\partial_t (h_i \hat{\mathbf{V}_i}) + \nabla \cdot (h_i \hat{\mathbf{V}_i} \hat{\mathbf{V}_i}) + f \hat{\mathbf{v}}_i \times h_i \hat{\mathbf{V}_i}] = -h_i \hat{\nabla} \phi_i - h_i h_2^0 \hat{\phi}_2^0 - \rho_0 \hat{\nabla} (h_i \hat{h}_i^0 \hat{\mathbf{V}_i} \hat{\mathbf{V}_i}), \]

(8b)

\[ \hat{\phi}_2 = \hat{\phi}_1 + \rho_0 g^2 (H_1 - \hat{h}_1), \]

(8c)

where the Reynolds stress term \( -\rho_0 \nabla (h_i \hat{\mathbf{V}_i} \hat{\mathbf{V}_i}) \) is associated with only the lateral redistribution of momentum. The form stress term \( -h_i \hat{\nabla} \phi_i \) on the other hand, can be rewritten as

\[ -h_i \hat{\nabla} \phi_i = -(1/2) \nabla \partial_t P, \]

(9a)

\[ -(1/2) \nabla \hat{\mathbf{P}_{\text{eddy}}}, \]

(9b)

where Eq. (8c) and \( h_i^0 + h_i^0 = 0 \) have been used to derive Eq. (9b), and \( \hat{\mathbf{P}_{\text{eddy}}} \) is given by Eq. (2b). The first terms of Eqs. (9a) and (9b)

\[ ^{\text{Let } P^o \text{ be the sea surface pressure and } z (\leq 0) \text{ the water depth. At all depths in the upper layer (i.e. } h_1 < z < 0), \text{ the MP is written by } [\mathbf{p}^o - \rho_0 g \hat{h}_1 + \rho_0 \mathbf{g} z] + \mathbf{g} h_1 = \mathbf{p}^o - \rho_0 g \hat{h}_1 = \mathbf{p}^o - \rho_0 g z. \text{ The difference of the MP in the upper and lower layers is } -\rho_0 g h_1^0, \text{ and hence its anomaly is written by } \rho_0 g^2 (H_1 - h_1) = \phi_2 - \phi_1, \text{ which is Eq. (5c).} \]
represents the transfer of momentum between the upper and lower layers. GL90 suggested to parameterize these terms as a vertical viscosity term, noting that mesoscale eddies relax the vertical shear of the basic current in thermal wind balance. The first terms of Eqs. (9a) and (9b) are also closely related to the pressure-based expression of the vertical component of the Eliassen–Palm flux (Andrews 1983; Lee and Leach 1996; Greatbatch 1998; Aiki et al. 2015). The second terms on the right hand side of Eqs. (9a) and (9b) may be regarded as the lateral gradient of an eddy-induced pressure. Aiki and Richards (2008) have confirmed that the net effect of the lateral gradient term is one order in terms of magnitude smaller than the vertical redistribution term.

Using Eqs. (8a)–(8c), one can derive equations for the mean APE, the mean KE, and the mean MP flux divergence in each layer:

\[
\begin{align*}
\partial_t \overline{\text{APE}} &= -\rho \overline{g} \hat{h}_1 \overline{\nabla} \cdot \overline{\nabla} V + \phi \overline{\nabla} \cdot \overline{\nabla} \phi, \\
\partial_t \overline{\text{KE}} &= \nabla \cdot \left( \overline{\nabla} \cdot \overline{\nabla} \phi \right), \\
\nabla \cdot \left( \phi \overline{b}_1 \overline{V}_i \right) &= -\phi \overline{\nabla} \cdot \overline{\nabla} \phi + \phi \overline{\nabla} \cdot \overline{\nabla} V + \overline{\nabla} \cdot \overline{\nabla} \phi,
\end{align*}
\]

where Eqs. (8c) and \(\partial_t \overline{T}_1 = -\partial_t \overline{T}_2\) have been used to derive Eq. (10a). The last two terms of Eq. (10b) represent the roles of the form stress and the Reynolds stress, and account for interaction with the eddy field energy. The other terms on the right hand side cancel out once the sum of Eqs. (10a)–(10c) is taken to yield an equation for the mean field energy \(\overline{\text{APE}} + \overline{\text{KE}} + \overline{\text{MP}}\).

It should also be noted that the first term on the right hand side of Eq. (10b), as well as

\[
\partial_t \overline{\text{APE}} = -\phi \overline{\nabla} \cdot \overline{\nabla} \phi,
\]

the last term of Eq. (10c), have been written using Eq. (4) as \(-\overline{\nabla} \cdot \overline{\nabla} \phi = -\overline{\nabla} \cdot \left( \overline{\nabla} \phi + \overline{\nabla} \phi \right)\), which will prove useful later in the article.

Finally, we consider equations for mesoscale eddies. Subtraction of Eqs. (10a)–(10c) from the low-pass-filtered version of Eqs. (6a)–(6c) yields equations for the eddy APE, the eddy KE, and the MP flux divergence in each layer:

\[
\begin{align*}
\partial_t \overline{\text{APE}}_{\text{eddy}} &= -\rho \overline{g} \hat{h}_1 \overline{\nabla} \cdot \overline{\nabla} V_{\text{eddy}} + \phi \overline{\nabla} \cdot \overline{\nabla} \phi_{\text{eddy}}, \\
\partial_t \overline{\text{KE}}_{\text{eddy}} &= \nabla \cdot \left( \overline{\nabla} \cdot \overline{\nabla} \phi_{\text{eddy}} \right), \\
\nabla \cdot \left( \phi_{\text{eddy}} \overline{b}_1 \overline{V}_{i_{\text{eddy}}} \right) &= -\phi_{\text{eddy}} \overline{\nabla} \cdot \overline{\nabla} \phi_{\text{eddy}} + \phi_{\text{eddy}} \overline{\nabla} \cdot \overline{\nabla} V_{\text{eddy}} + \overline{\nabla} \cdot \overline{\nabla} \phi_{\text{eddy}},
\end{align*}
\]

An energy diagram for Eqs. (10a)–(11c) is illustrated in Fig. 2a. In order to clarify the structure of the energy cycle, equations for the energy interaction, (10c) and (11c), are here written separately from the budget of the APE and KE (in contrast to traditional oceanic studies, where equations for the pressure flux divergence...
Fig. 2. A four-box energy diagram based on energy equations averaged in (a) density coordinates as shown by Eqs. (10a)–(11c) and (b) $z$ coordinates as shown by Eqs. (17a)–(19c). (a) is appropriate to the approach taken by Greatbatch and Lamb (1990), and (b) to the approach taken by Gent and McWilliams (1990). The cyan arrows represent the energy conversion route associated with the wind-induced Ekman flow. The green arrows represent the energy conversion route associated with the continental drag Ekman flow. The red arrows represent the main energy conversion route associated with baroclinic mesoscale eddies. The yellow arrows represent the energy conversion route associated with the Reynolds stress. Note that $\nabla \phi = \nabla_2 p$ (the lateral gradient of the MP in density coordinates is identical to the horizontal gradient of hydrostatic pressure in $z$ coordinates).

have been merged into equations for the KE. In the eddy field of Fig. 2a, the quantity $- \hat{V} \cdot h''' \nabla \phi'''$ is independently connected to both the eddy APE and the eddy KE; this is due to Eq. (11c). In particular, the direct connection between the eddy APE and the mean KE involves both the density surface perturbation and the layer thickness form stress. The situation in the eddy field is consistent with the result of Bleck (1985) and Iwasaki (2001) derived from the mass-weighted-mean equations for a continuously stratified non-Boussinesq fluid (see Chassignet and Boudra (1988) and Røed (1999) for oceanic applications). In addition to the form stress term, the Reynolds stress term connects the mean and eddy KEs, which is as in the Lorenz (1955) energy diagram and is relevant to the role of relative vorticity in barotropic instability, baroclinic instability, and geostrophic turbulence (cf. Thompson 2010; Berloff and Kamenev 2013; Chapman et al. 2015).

To summarize, although it has been little mentioned in previous studies, the simplest framework for understanding the role of layer
thickness form stress and associated energy conversions is a classical two-layer model with a modified definition of the mean and eddy KEs. Extension of the above formulation to a continuously stratified fluid with arbitrary bottom topography is straightforward as done by Aiki and Yamagata (2006) and Aiki and Richards (2008). These authors have further shown that the energy equations are robust in the presence of ocean boundaries; in particular, the structure of the energy diagram is unchanged even in the presence of (i) finite-amplitude perturbations associated with mesoscale eddies, as well as (ii) density surface outcropping (i.e. intersecting) at the top and bottom boundaries of the ocean.

2.3. A Model Diagnosis in Density Coordinates

In order to illustrate the energy cycle in the new energy diagram, we use a set of three-day snapshots from a high-resolution (0°×0.1°) near-global hindcast simulation. This simulation is called the OFES (OGCM For the Earth Simulator) hindcast run, and was forced by daily mean wind stress from the QuikSCAT data and daily-mean heat and fresh water fluxes from the NCEP/NCAR reanalysis (Masumoto 2010). A set of three-day snapshots archived throughout the years 2004–2007 of the OFES hindcast run was used for the diagnosis (a total of 483 three-dimensional snapshots of the global ocean). These snapshots were originally in z coordinates. The model has 54 depth levels, with the discretization varying from 5 m at the surface to 330 m at the maximum depth of 6065 m. Each snapshot was first mapped onto density coordinates in each vertical column. As in Aiki and Richards (2008), we use 80 density layers defined by the potential density referenced to sea surface pressure. This density was chosen in an attempt to overview the global ocean with most of the KE distributed above the main thermocline. In the present diagnosis, the low-pass temporal filter (overbar) was set to a four-year mean for 2004–2007. We have determined the global distribution of energy conversion terms concerning the budget of the mean KE in Eq. (10b) and Fig. 2a. These are 

\[ -\nabla \cdot \left( \hat{V} \nabla \cdot \hat{V} \right) - \nabla \cdot \left( \hat{V} \nabla \hat{\varphi} \right) - \nabla \cdot \left( \hat{\varphi} \hat{V} \right), \]

and 

\[ -\rho_b \hat{V} \cdot \nabla \left( \nabla \cdot \hat{\varphi} \right) \]

for each of which the depth integral in each vertical column is plotted in Fig. 3 with negative (positive) values indicating a decrease (increase) in the mean KE.

Whilst significant positive and negative signs are evident in the work of the isopycnal mean velocity, 

\[ -\nabla \cdot \nabla \hat{\varphi} \] (Fig. 3a), the working rate in the model, integrated globally, is −0.77 TW (negative), which indicates conversion of the mean KE to the mean APE. This quantity mainly reflects the wind-induced Ekman flow steepening the slope of isopycnal surfaces near the sea surface. Strictly speaking, the work of the isopycnal mean velocity includes also the effects of continental-drag Ekman flow (associated with the sum of the horizontal friction term and the bottom friction term in the model), which is represented by the positive signals near the western boundaries of the Pacific, Atlantic, and Indian Oceans, in particular the Greenland Sea, the upstream of the Kuroshio and the Gulf Stream, and the Agulhas Current. These positive signals indicate conversion of the mean KE to the mean APE to the mean KE, which is then dissipated by the boundary friction at the continental slope of the ocean. (See Aiki et al. 2011b for the global distribution of the energy dissipation by the boundary friction in the 0.1°×0.1° OFES simulation; the working rate by the boundary friction integrates to −0.32 TW in the modelled global ocean, not shown. See also Trossman et al. 2013.)

The work done by the bolus velocity, 

\[ -\nabla \hat{V} \cdot \nabla \hat{\varphi} \], is shown in Fig. 3b. This quantity is clearly positive over each region of high eddy activity in the global ocean: the mean APE is extracted to provide an input to the mean KE when the eddy-induced overturning circulation relaxes the slope of isopycnal surfaces.
Fig. 3. The depth integral of the energy conversion rates (W/m²) associated with (a) the unweighted mean velocity $\mathbf{V} \cdot \nabla \Phi$ representing the effect of the Ekman flow, (b) the bolus velocity $\mathbf{V}_B \cdot \nabla \Phi$ representing the effect of the eddy-induced overturning circulation, (c) the layer thickness form stress $-\mathbf{V} \cdot h'''' \nabla \Phi'''$ representing the effect of the vertical redistribution of momentum by eddies, and (d) the Reynolds stress $-\rho_0 \mathbf{V} \cdot \left( \nabla \cdot (h'' V'' V'') \right)$ representing the effect of the lateral redistribution of momentum by eddies. The sign is referenced to the bud- get of the mean kinetic energy, positive indicating an increase in the mean kinetic energy. The estimate is based on the three-day snapshot of the OFES hindcast simulation with application of a four-year mean in density coordinates.

The working rate in the modeled global ocean integrates to $+0.87$ TW. The fact that the global distribution is mostly positive indicates that this quantity provides a suitable basis for mesoscale eddy parameterization in climate ocean models, and is consistent with GM90.

The work done by the layer thickness form stress, $-\mathbf{V} \cdot h'''' \nabla \Phi'''$, is shown in Fig. 3c, where negative values indicate an energy cascade to the eddy field associated with the relaxation of the vertical shear of the geostrophic currents. The horizontal distribution of the work done by the layer thickness form stress is partly anticorrelated with the work done by the bolus velocity in Fig. 3b. Nevertheless, there are also positive signals, for example in the Kuroshio Extension region. The working rate in the modeled global ocean integrates to $-0.48$ TW. Note that this is the combined effect of the first and second terms on the right hand side of Eqs. (9a) and (9b). If we had plotted the work done by only the first (the vertical redistribution) term of Eqs. (9a) and (9b) which is associated with the GL90 parametrization, the sign is clearly negative in each region of the global ocean, as is confirmed by Aiki and Richards (2008).

The work done by the Reynolds stress, $-\rho_0 \mathbf{V} \cdot \left( \nabla \cdot (h'' V'' V'') \right)$, is shown in Fig. 3d. The working rate in the model-global ocean integrates to $-0.15$ TW, suggesting an energy cascade to the eddy field associated with the relaxation of the lateral shear of the geostrophic currents. This might indicate the effect of barotropic instability, for example at the western boundary of the Indian Ocean, and in the Kuroshio and the Gulf Stream regions. However, in the Gulf Stream extension region, the signal is clearly positive, indicating acceleration of the mean current by the eddies (Held and Andrews 1983; Greatbatch et al. 2010; Waterman and Jayne 2011), in which the meridional radiation of Rossby waves could play a role (Thompson 1971).

To summarize, the equilibrium of the mean KE in the wind-driven ocean circulation is
maintained by the following energy cycle. Wind forcing provides an input to the mean KE, which is then transferred to the mean APE by the wind-induced Ekman flow (the role of the isopycnal mean velocity \( \mathbf{V} \)). Nevertheless, the net mean APE does not change, because the eddy-induced overturning circulation (the role of the bolus velocity \( \mathbf{V}^b \)) extracts some of the mean PE and feeds the mean KE, which is subsequently drained by the work of the layer thickness form stress term \( \frac{1}{H} \frac{d}{dz} H \) endowing the eddy field with an energy cascade. The result of the above diagnosis is complementary to Aki and Richards (2008), in that (i) we also identify the work of the Reynolds stress term, (ii) our statistics are based on a four-year mean of three-day snapshots instead of monthly means of three-day snapshots, and (iii) we use the output of a hindcast simulation forced by daily-mean atmospheric forcing instead of monthly mean of three-day snapshots instead of monthly.

3. Maintenance of the Eddy Field Energy

Identifying the life cycle of mesoscale eddies is one of the major challenges of present-day oceanography. Theories for mesoscale eddies develop through diagnoses of field observations, satellite data, and numerical simulations. A practical approach is to work in \( z \) coordinates (where \( z \) is the geopotential height). This is a trade-off between the mathematical convenience of the Eulerian mean framework in \( z \) coordinates and the physical convenience (to distinguish the adiabatic and diabatic processes) of the TWM framework in density coordinates. Also, many of the widely used numerical models are formulated in \( z \) coordinates, an example being the OFES model discussed in the previous section.

Let \( \mathbf{U} = (\mathbf{V}, \mathbf{u}) \) and \( \mathbf{V}' = (\mathbf{V}', \mathbf{l}) \), where \( \mathbf{V}' \) is the horizontal gradient operator in \( z \) coordinates (Table 1). Let a low-pass time filter in \( z \) coordinates be denoted by an overbar with superscript \( z \), and \( A' = A - \overline{A} \) (compared at constant depth) for an arbitrary quantity \( A \). It is obvious that \( \nabla_z \mathbf{U}' = 0 \), and \( \mathbf{U}' \cdot \mathbf{n} = 0 \), where \( \mathbf{n} \) is a unit vector normal to the sloping bottom boundary. This is convenient for constructing theories for the life cycle of mesoscale eddies. In contrast, matching the deviation in density coordinates (\( A'' = A - \overline{A} \), compared at constant density) onto the mean depth of each isopycnal yields \( \nabla_z \mathbf{U}'' \neq 0 \) and \( \mathbf{U}'' \cdot \mathbf{n} \neq 0 \). This concern is the reason why McDougall and McIntosh (2001) suggested to replace the set of the isopycnal mean velocity and the bolus velocity (each of which is three-dimensionally divergent and does not satisfy a no-normal-flow boundary condition) in Gent et al. (1995) with the set of the Eulerian mean velocity and the quasi-Stokes velocity, each of which is three-dimensionally nondivergent and satisfies a no-normal-flow boundary condition (Appendix B). The exact definition of the quasi-Stokes stream function is

\[
\Psi_{\text{Q-S}}^z = \int_{-h_b}^{0} (\mathbf{V} - \nabla_z) dz,
\]

where \( h_b(x, y) > 0 \) is the bottom depth. It can be easily shown that \( \Psi_{\text{Q-S}}^z \) vanishes at the top and bottom boundaries of the ocean without relying on extra assumptions (see the pile-up rule of Aiki and Yamagata (2006)), which is in contrast to theories relying on diabatic mixing in the surface mixed layer (Plumb and Ferrari 2005, Ferrari et al. 2008). In this section, we briefly explain energy equations averaged in \( z \) coordinates and then present the state-of-the-art understanding of the life cycle of mesoscale eddies in the classical Lorenz (1955) energy diagram.

3.1. Energy Equations Averaged in \( z \)-coordinates

Equations for the incompressibility, density, and momentum in \( z \) coordinates (neglecting diabatic
and viscous effects for simplicity) read
\[ \nabla \cdot \mathbf{U} = 0, \]  
\[ (\partial_t + \mathbf{U} \cdot \nabla) \rho = 0, \]  
\[ \rho_0[(\partial_t + \mathbf{U} \cdot \nabla) \mathbf{V} + \mathbf{f} \times \mathbf{V}] = -\nabla \mathbf{p}, \]  
\[ 0 = -\partial_t \mathbf{p} - g(\rho - \rho^b), \]
where \( \rho \) is the combined sea surface and hydrostatic pressure \( (\nabla \mathbf{p} = \nabla \rho \) is understood). The quantity \( \rho^b(z) \) is the global background density which varies only in the vertical direction, and might be obtained by sorting all water parcels in the global ocean. Potential energy (PE) and KE are here defined by
\[ \mathcal{P} = g(\rho - \rho^b)z; \]  
\[ \mathcal{K} = \frac{\rho}{2} \mathbf{V}^2, \]
Using Eqs. (13a)–(13d), one may derive equations for PE, KE, and the pressure flux divergence:
\[ \partial_t \mathcal{P} + \nabla \cdot (\mathbf{U}(\mathcal{P} - \nabla \mathbf{b})) + g(\rho - \rho^b)\mathbf{w}, \]  
\[ \partial_t \mathcal{K} + \nabla \cdot (\mathbf{U}\mathcal{K}) = -\nabla \cdot \mathbf{V}\mathbf{p}, \]  
\[ \nabla \cdot (\mathbf{U}\partial_t \mathbf{P}) = \nabla \cdot \mathbf{V}\partial_t \mathbf{p} - g(\rho - \rho^b)\mathbf{w}, \]
where the expression (17a) omits the triple and higher product of perturbation quantities for simplicity, and
\[ \Psi^{\alpha \beta} = -\frac{\rho}{\rho^2} \nabla \cdot \left( \frac{1}{2} \nabla \rho \right) \]
\[ \rho^b(z) \] is the MP associated with the background density. Equation (15a) for PE is exact and applicable to finite-amplitude density variations as well as sloping-bottom boundary conditions (Kang and Fringer 2010, Aiki et al. 2011a). For each of \( \mathcal{P} \) and \( \mathcal{K} \), the low-pass-filtered energy can be written as the sum of the mean and eddy field energies (Appendix C),
\[ \mathcal{P} = g(\mathcal{P}_0 - \mathcal{P}^b)z + \frac{g}{2} \rho^2 \frac{\mathcal{P}}{\rho^b}; \]
\[ \mathcal{K} = \frac{\rho}{2} \mathbf{K}^2 + \frac{\rho}{2} \mathbf{V}^2, \]
As shown below, the exact definition of the mean and eddy PE in Eq. (16a) allows us to derive the mean and eddy PE equations (19a) and (17a) without relying on any approximations, which is in contrast to previous studies where the mean and eddy PE are defined by \( (g/2)(\mathcal{P} - \mathcal{P}^b)^2/(\mathcal{P}^b)^2 \) and \( (g/2)\mathcal{P}/(\mathcal{P}^b)^2 \), respectively (Böning and Budish 1993, Ollers et al. 2012). See Appendix D for details. Using \( (\partial_t + \mathbf{U} \cdot \nabla) \rho = -\mathbf{V} \cdot \nabla \mathbf{p} - \mathbf{V} \mathcal{P} \) and \( (\partial_t + \mathbf{U} \cdot \nabla) \mathcal{P} = -\mathbf{V} \cdot \mathbf{V} \mathcal{P} - \mathbf{V} \mathcal{K} \), one may derive equations for the eddy PE, the eddy KE, and the pressure flux divergence in the eddy field to read
\[ \partial_t \mathcal{P}^{\text{eddy}} + \nabla \cdot (\mathbf{U}\mathcal{P}^{\text{eddy}}) = g\mathcal{P}^{\text{eddy}} + (-\Psi^{\alpha \beta} g \nabla \rho \mathcal{P} + \Psi^{\alpha \beta} \mathcal{P}^{\text{eddy}}) \]
\[ \partial_t \mathcal{K}^{\text{eddy}} + \nabla \cdot (\mathbf{U}\mathcal{K}^{\text{eddy}} + \mathbf{U}\mathcal{K}) = -\nabla \cdot \mathbf{V} \mathcal{P}^{\text{eddy}} + \rho_0 \nabla \cdot \nabla \mathbf{p} \]
\[ \nabla \cdot (\mathbf{U}\partial_t \mathbf{P}^{\text{eddy}}) = \nabla \cdot \mathbf{V} \partial_t \mathbf{p}^{\text{eddy}} - g\rho^{\text{eddy}}, \]
where the expression (17a) omits the triple and higher product of perturbation quantities for simplicity, and
\[ \psi^{\alpha \beta} = -\frac{\rho}{\rho^2} \nabla \cdot \left( \frac{1}{2} \nabla \rho \right) \]
turns out to be identical to the approximated expression of \( \psi_{\text{shallow}} \) in McDougall and McIntosh (2001). They transformed Eq. (12) to Eq. (18) using an approximation \( z'''' \sim -\rho''/\rho^2 \) which is applicable to only depths away from the top
and bottom boundaries of the ocean. This is why Eq. (18) does not vanish at the top and bottom boundaries of the ocean, in contrast to Eq. (12).

\[ \partial \phi V^{\text{mean}} + \nabla \cdot \left[ \left( \phi V^{\text{mean}} + g\bar{\varphi} \right) + \nabla \rho' g \frac{z}{2} \right] = g \left( \rho' \cdot -\bar{\varphi} \right) - \left( -\Psi' \cdot g \nabla \bar{w}' + \tau_{\text{eddy}} \right), \quad (19a) \]

\[ \partial \rho' + \nabla \cdot \left( \nabla \rho' \right) = -\rho \left[ \nabla \cdot \left( \nabla \bar{v} \right) \right] \quad (19b) \]

\[ \nabla \cdot \left( \nabla \bar{v} \right) = -\nabla \cdot \left( \nabla \bar{v} \right) = -g \left( \rho' \cdot -\bar{\varphi} \right) \] (19c)

An energy diagram based on Eqs. (17a)–(17c) is illustrated in Fig. 2b, and it is a slightly improved version of the classical Lorenz (1955) energy diagram.

When one is considering baroclinic mesoscale eddies that extract the mean PE (to feed the eddy PE in the Lorenz diagram), the sign of \( \Psi' \cdot g \nabla \bar{v} \) should be negative. Using integration by parts, the depth integral of \( \nabla \cdot \left( \nabla \rho' \right) \) approximates that of \( \nabla^{2} \rho ' \), which has already been shown in Fig. 3b, and is clearly positive in each region of the global ocean. This result of the model diagnosis justifies the principle of the GM90 parametrization, which is to make \( \Psi' \cdot g \nabla \bar{v} \) negative in each vertical column. There are several possibilities for parametrizing \( \Psi' \). The solution of GM90 is \( \Psi' \propto -\nabla \bar{v} \), which has been explained in Sec. 1 (see also Appendix A). The solution of Aiki et al. (2004) is \( \Psi' \propto \nabla \bar{w} \), which enables the stream function to be closed at the top and bottom boundaries without resorting to tapering. Ferrari et al. (2010) suggested to blend the above two solutions. It is also of interest to note that Eq. (12) of Eden et al. (2007) is the quasi-geostrophic equivalent of Eq. (17a).

Subtraction of Eqs. (17a)–(17c) from the low-pass-filtered version of Eqs. (15a)–(15c) yields equations for the mean PE, the mean KE, and the pressure flux divergence in the mean field:

\[ \partial \varphi V^{\text{mean}} + \nabla \cdot \left[ \left( \varphi V^{\text{mean}} + \bar{\varphi} \right) + \nabla \rho' g \frac{z}{2} \right] = \left( -\Psi' \cdot g \nabla \bar{w}' + \tau_{\text{eddy}} \right), \]

Equation (13) in their paper goes on to provide a formula for deriving the GM90 diffusivity for parametrizing the quasi-Stokes velocity.

3.2. Eddy Energy Generation and Vertical Eddy Energy Flux

The APE built up by the large-scale Ekman pumping/suction of the main thermocline is hypothesized to be subsequently released by the generation of eddies through baroclinic instability of the mean flow (e.g. Gill et al. 1974; Wunsch 1998). This hypothesized energy route has been supported by a series of idealized numerical model experiments (e.g., rectangular basin, flat bottom, and simplified surface forcing) by Radko and Marshall (2003), who found that the downward energy flux from the Ekman layer into the main thermocline is largely balanced by lateral eddy transfer, with small-scale mixing making only a minor contribution. Recently, Zhai and Marshall (2013) re-examined this problem using a realistic eddy-permitting model of the North Atlantic Ocean driven by climatological monthly-mean forcing.

The model used in their study is the MIT General Circulation Model (Marshall et al. 1997). The model spans the Atlantic Ocean between...
Fig. 4. (a) The wind power input to the North Atlantic Ocean (W m$^{-2}$). (b), (c), and (d) show the rate of APE released by baroclinic instability ($-g \rho' w' z$) at 61 m, 900 m, and 1875 m, respectively. The color scale is saturated in order to reveal regions of moderate eddy activity. (Adapted from Zhai and Marshall 2013.)
remarkable to see that the values are overwhelm-
ingly positive, especially away from the tropical
region, meaning that the eddies act as a sink for
the APE built up by the large-scale wind action
almost everywhere in the model. Interestingly,
besides the western boundary region where the
eddies are known to have a major role to play,
the eddies are also found to systematically flat-
ten isopycnals in regions such as the eastern and
southern rims of the subtropical gyre (Fig. 4b).
Integrating over the North Atlantic Ocean, the
rate of APE extracted by baroclinic instability
is found to be about 0.11 TW, approximately
80% of the total wind energy input. The model-
ing study by Zhai and Marshall (2013) thus adds
further support to the hypothesis that the wind
energy input to the gyre circulations is largely
balanced by eddy generation through baroclinic
instability — consistent with our earlier finding
when discussing Fig. 3 (see also Fig. 4 of Eden
et al. 2007).

As noted by Zhai and Marshall (2013), the
eddy KE generation, \(-g \rho \rho' \left( \frac{\partial w}{\partial z} \right)\), is mainly con-
fined to the upper ocean in the subtropical gyre,
but has a much deeper structure in the subpolar
gyre, with its maximum strength at about 2000
m depth on average. This difference in the verti-
cal structure of eddy KE production has impli-
cations for the vertical fluxes of eddy energy.

Figure 5 shows the vertical eddy energy flux,
\(\rho \rho' \left( \frac{\partial w}{\partial z} \right)\), at different depths and latitudes, which
represents the redistribution of eddy energy in

---

Fig. 5. The vertical eddy energy flux \(\rho \rho' \left( \frac{\partial w}{\partial z} \right)\times 10^{-2}\ W\ m^{-2}\) at (a) 61 m, (b) 900 m, (c) 39°N, and (d) 59°N.
(Adapted from Zhai and Marshall 2013.)
the vertical [see Eq. (17c)]. Large downward eddy energy fluxes are found near the surface in the North Brazil Current and Caribbean Sea, and deeper down in the water column in the western boundary current and its extension. In contrast, the vertical eddy energy flux is mostly upward in the subpolar region, especially in the Labrador Sea. This difference in the direction of $p'w'z'$ is further illustrated in Figs. 5c and 5d, where $p'w'z'$ is mainly downward under the Gulf Stream at 39°N, but predominantly upward in the Labrador Basin and Irminger Basin at 59°N — consistent with the depths of eddy KE sources at these two latitudes. The vertical eddy energy flux has received relatively little attention to date, although the equivalent vertical energy flux associated with internal waves/tides has been widely diagnosed in both observations and numerical models. This issue apparently merits further investigation.

3.3. The Eddy Energy Sink and Its Large-Scale Effect

In equilibrium, the energy flux into the eddy field has to be dissipated. However, our understanding of the fate of ocean eddies remains rather poor. Potential candidate processes that may remove eddy energy from the ocean include: bottom frictional dissipation (e.g. Sen et al. 2008; Wright et al. 2012), damping by air-sea momentum and heat fluxes (e.g. Duhaut and Straub 2006; Zhai and Greatbatch 2006, 2007; Greatbatch et al. 2007; Shuchburgh et al. 2011), generation of lee waves over rough topography (e.g. Marshall and Naveira Garabato 2008; Nikurashin and Ferrari 2011; Scott et al. 2011), the “Neptune effect” along sloping topography (e.g. Holloway 1992; Dukowicz and Greatbatch 1999; Greatbatch and Li 2000; Adcock and Marshall 2000), and interactions with the internal wave field through loss of balance and Rossby wave deformation (e.g. Straub 2003; Molemaker et al. 2005; Bühler and McIntyre 2005). To date, there is still no consensus on which mechanism(s) provides the dominant eddy energy sink.

Ocean eddies are observed to propagate westward at long Rossby wave speeds (e.g. Chelton et al. 2007), but what happens to the eddies when they encounter the western boundary has been unclear. Recently, using a combination of models, satellite altimetry, and climatological hydrographic data, Zhai et al. (2010) showed that the western boundary acts as a “graveyard” for the westward-propagating eddies. The model used is a nonlinear reduced-gravity model on a β plane, with a lateral resolution of 3.5 km. Model integrations initialized with both a single eddy and a random sea of eddies have been conducted to examine the eddy energy budget near the western boundary. These model experiments suggest the following picture: upon encountering the western boundary, the available potential energy associated with the eddies is converted into kinetic energy of the reflected short Rossby waves and smaller eddies, the majority of which is dissipated near the western boundary. In the reduced-gravity model, there is only one baroclinic mode and the only energy sink for eddies is lateral viscous dissipation, whereas in the ocean much of the eddy energy is likely scattered into high-wave-number vertical modes that rapidly dissipate (e.g. Dewar and Hogg 2010).

Using satellite altimetry and climatological hydrographic data, Zhai et al. (2010) computed the divergence of the depth-integrated linear eddy energy fluxes associated with the first baroclinic mode, which is presumably balanced by eddy energy sources and sinks. Readers are referred to Zhai et al. (2010) for a detailed method description and error analysis (see Kunze et al. 2002 for an application of a similar method to diagnosing internal wave energy dissipation). Figure 6 shows the sources and sinks of eddy energy in the first baroclinic mode.
mode. Note that regions shallower than 300 m deep are shaded in light gray and excluded in the calculation. The most striking features are the ubiquitous eddy energy source (in red) in the interior and the eddy energy sink (in blue) near the western boundary in each ocean basin. Therefore, in agreement with the model results, the satellite altimetry data also point to the western boundaries as an important region of energy loss for ocean eddies. Furthermore, the total eddy energy sink near the western boundaries poleward of 10° of latitude is estimated to be approximately 0.1–0.3 TW, representing a significant fraction of wind power input to the ocean general circulation. Wright et al. (2012) have recently estimated the bottom dissipation of eddy energy at the Atlantic zonal boundaries using ocean current meter data and found that dissipation at the western boundary is significantly larger than that at the eastern boundary, adding support to the idea that the western boundary acts as a “graveyard” for ocean eddy energy.

Some of the dissipating processes that remove eddy energy may lead to enhanced diapycnal mixing in the western boundary regions (i.e. the dissipated eddy energy is converted back to the mean PE), which can have important implications for the large-scale ocean circulation and climate. Saenko et al. (2012) recently conducted a suite of sensitivity experiments with an ocean general circulation model assuming that the eddy energy that converges at the western boundary is scattered into high-wave-number vertical modes, resulting in locally enhanced diapycnal mixing (Dewar and Hogg 2010). When diapycnal mixing in their model is maintained only by the tidal energy dissipation (tidal experiment; Fig. 7a), the abyssal overturning circulation and stratification are too weak. When diapycnal mixing in their model is maintained by both the tidal and eddy energy dissipation (tidal+eddy experiment; Fig. 7c), the abyssal overturning circulation and stratification become stronger and closer to the observations. Furthermore, mixing associated with the eddy energy dissipation is found to be able to drive a relatively strong overturning in the abyss (∼5 Sv) on its own (eddy experiment; Fig. 7b). These sensitivity experiments highlight the importance of including mixing associated with the eddy energy dissipation especially near the western boundary in future ocean climate models.
4. Summary

We have reviewed the state-of-the-understanding of the energy cycle of the global ocean circulation, focusing on the role of baroclinic mesoscale eddies. The presence of two independent but equivalent prescriptions for the energy cycle — one based on the energy diagram of Lorenz (1955) and the other based on the energy diagram of Bleck (1985) — provides complementary ways of understanding the role of mesoscale eddies in ocean circulation theories, such as parametrization used in climate OGCMs, the dynamics of the Antarctic Circumpolar Current (ACC), and the life cycle of mesoscale eddies. Two of the important effects of mesoscale eddies are: (i) the flattening of the slope of large-scale isopycnal surfaces by the eddy-induced overturning circulation, the basis for the GM90 parametrization; and (ii) the vertical redistribution of the momentum of basic geostrophic currents by the layer thickness form stress (the residual effect of pressure perturbations), the basis for the GL90 parametrization. While only point (i) can be explained using the classical Lorenz (1955) energy diagram, both point (i) and point (ii) can be explained using the energy diagram of Bleck (1985), which is based on the modified definitions of the mean and eddy KEs in density coordinates or layer models.

With the role of the layer thickness form stress in the ACC in mind, the energy cycle of the wind-driven circulation in the global ocean is explained as follows. Wind forcing provides an input to the mean KE, which is then transferred to the available potential energy (APE) of the large-scale field by the wind-induced Ekman flow. In some regions the APE is extracted by the eddy-induced overturning circulation to feed the mean KE, as expressed by \(- \nabla \cdot \psi \cdot \mathbf{V} \cdot \mathbf{g} \) in Fig. 2a. This route is found to be in the mean field, and it provides an input to the mean KE (i.e. acceleration of the mean current), in contrast to the corresponding conversion term in the classical Lorenz energy diagram, as represented by \(- \nabla \cdot \mathbf{V} \cdot \nabla \rho \) in Fig. 2b. It is also noted in Fig. 2a that the mean KE \(\rho_0 / 2 \nabla \mathbf{V}^2 \) will leak to the eddy field by \(- \nabla \cdot \mathbf{V} \cdot \nabla \rho / \rho \), a result of the vertical redistribution of momentum (i.e. deceleration of the mean current) by the layer thickness form stress.
While the source of the eddy field energy has become clearer, identifying the life cycle of mesoscale eddies remains one of the major challenges of present-day oceanography. The sink and flux of the eddy field energy are yet to be characterized both physically and spectrally. The eddy energy flux in physical space is closely related to the westward motion of planetary eddies, with the consequence that the western boundary is a major sink for eddy energy. Potential candidate processes that may remove eddy energy from the ocean include: bottom frictional dissipation, damping by air-sea momentum and heat fluxes, and interactions with the internal gravity wave field through loss of balance and Rossby wave deformation.

Acknowledgments

H. A. would like to express the deepest gratitude to his supervisor, Emeritus Professor Toshio Yamagata, for introducing the wonder of mesoscale eddies and the valuable guidance at the University of Tokyo and JAMSTEC. He also thanks Prof. Masaki Ishiwatari for helpful comments on Appendix D. R. J. G. is grateful to Prof. Yamagata for his friendship of over 30 years and to GEOMAR for continuing support. X. Z. would like to thank David Marshall for the many helpful discussions.

Appendix A. Tracer and Momentum Approaches for the Mesoscale Eddy Parametrization

Most OGCMs currently maintained at the world’s climate centers adopt the tracer approach. Indeed, the inclusion of the eddy-induced advection of tracers represented a major advance in ocean modeling, allowing coarse-resolution OGCMs used in climate studies to maintain a sharper and more realistic thermocline and removing spurious overturning cells in the Southern Ocean (Danabasoglu and McWilliams 1994; Griffies et al. 2000; Gent 2011). The GM90 parametrization presents an eddy-induced velocity based on the principle that mesoscale eddies flatten the large-scale slope of isopycnals, which may be interpreted as a lateral diffusion of the depth of each isopycnal surface (Fig. 1). Letting $\kappa$ denote this diffusion coefficient, the eddy-induced overturning stream function associated with the GM90 parametrization can be written as $-\kappa$ times the slope of the isopycnals [cf. Eq. (18) of Gent et al. 1995].

The momentum approach is much less widely used in the community, despite being advocated by a number of authors (Greatbatch 1998; Wardle and Marshall 2000; Zhao and Vallis 2008). An example applied to a global ocean model is the study by Ferreira and Marshall (2006). Let $\nu$ denote the coefficient of the GL90 viscosity; then the vertical flux of momentum associated with the layer thickness form stress is parametrized by: $\rho_0 \nu$ times the vertical shear of the total transport velocity. The two approaches, GM90 and GL90, are equivalent when the main component of the total transport velocity is the geostrophic Eulerian mean velocity and the coefficients of the GM90 diffusivity and the GL90 viscosity are related as $\kappa/N^2 = \nu/f^2$ where $N = \sqrt{(g/\rho_0) - \partial z^2}$ is the buoyancy frequency. Using the former condition, GL90 and McWilliams and Gent (1994)

The total transport velocity may consist of three kinds of velocity: (i) the geostrophic component of the Eulerian mean velocity, (ii) the Ekman component of the Eulerian mean velocity, and (iii) the eddy-induced additional transport velocity (the quasi-Stokes velocity). The Ekman velocity is almost absent in the midlatitude ocean interior, i.e. outside the equatorial region and the surface and bottom boundary layers, and the eddy-induced velocity is usually small compared to the geostrophic velocity of the large-scale flow (Aiki and Yamagata 2006).
have shown that
\[
\frac{\partial}{\partial z} \left( \nu \frac{\partial \mathbf{V}}{\partial z} \right) = -\mathbf{z} \times \frac{\partial}{\partial z} \left( \nu \left( \frac{\partial \mathbf{V}}{\partial z} \right) \mathbf{z} \right) = f \times \frac{\partial}{\partial z} \left( \nu \left( \frac{\partial \mathbf{V}}{\partial z} \right) \mathbf{z} \right) = f \times \frac{\partial}{\partial z} \left( \nu \frac{\partial \mathbf{V}}{\partial z} \right) - \kappa \left( \frac{\partial \mathbf{V}}{\partial z} \right)
\]

where the last line may be interpreted as the Coriolis force associated with the eddy-induced velocity as given by the GM90 parametrization. A typical value for the GM90 diffusivity is \( \kappa = 1000 \text{m}^2/\text{s} \) (i.e. operating laterally to flatten the slope of isopycnals), and a typical value for \( N^2/f^2 \) at midlatitudes is 400, for which the corresponding GL90 viscosity is \( \nu = 0.4 \text{m}^2/\text{s} \), several orders of magnitude greater than the coefficient of vertical eddy viscosity associated with three-dimensional turbulence.

Then, in order to scale the eddy-induced vertical stress in the ACC (Sec. 1), we have assumed the vertical shear of the geostrophic Eulerian mean velocity to be \((1.0 \text{m/s})/(2000 \text{m})\).

**Appendix B. The Quasi-Stokes Velocity and Gent and McWilliams (1990)**

In density coordinates for a continuously stratified fluid, Eq. (23a) may be rewritten as
\[
\nabla \tau + \mathbf{V} \cdot \left( \nabla \phi \mathbf{V} \right) = 0,
\]

where \( z \) is the height of an isopycnal surface and \( \phi_z = \partial \phi / \partial p \) is the thickness between adjacent isopycnal surfaces. The subscript \( \tau \) and the symbol \( \nabla \) are the temporal and lateral gradient operators in density coordinates, respectively, and \( \mathbf{V} \) is the unweighted isopycnal mean and \( \nabla \phi \mathbf{V} \) is the TWM for an arbitrary quantity \( \phi \) (Table 1). Mapping the above equation to \( z \) coordinates referencing the low-pass filtered height of each isopycnal surface yields
\[
\nabla \cdot \left( \mathbf{V} + \mathbf{\Psi} \tau + \mathbf{V} \nabla \tau \right) = 0,
\]

which leads to \( \nabla \cdot \mathbf{V} + \partial_z \mathbf{V} = -\left( \nabla \cdot \mathbf{V} + \partial_z \mathbf{V} \right) \); \( \partial_z \mathbf{V} \neq 0 \). Again with Eq. (21b) in mind, the quasi-Stokes velocity is defined as the difference between the total transport velocity and the unweighted isopycnal mean velocity to read
\[
\mathbf{V}^B \equiv \mathbf{V} - \mathbf{\Psi} \tau ,
\]

\[
\mathbf{w}^B \equiv \mathbf{\tau} \tau + \mathbf{V} \nabla \tau - \mathbf{\Psi} \tau ,
\]

which leads to \( \nabla \cdot \mathbf{V} + \partial_z \mathbf{V} = \left( \nabla \cdot \mathbf{V} + \partial_z \mathbf{V} \right) \neq 0 \). Thus, one can introduce an overturning vector stream function for the quasi-Stokes velocity, as in (12), to read \( \mathbf{w}^B = \partial_z \mathbf{\Psi}^\text{WST} \) and \( \mathbf{w}^B = -\nabla \cdot \mathbf{\Psi}^\text{WST} \) (McDougall and McIntosh 2001). Substitution of these into (23b) yields
\[
\mathbf{\tau} \tau + \mathbf{V} \nabla \tau - \mathbf{\Psi} \tau = -\nabla \cdot \mathbf{\Psi}^\text{WST} ,
\]

where the advection term on the left hand side is based on the Eulerian mean velocity and

\[\text{A} \equiv \nabla \cdot \mathbf{\tau} \tau ,\]

\[\text{B} \equiv \mathbf{\tau} \tau .\]

\(^{6}\text{Jacobson and Aiki (2006) presented an alternative proof for Eq. (21b) using a one-dimensional analog of the hybrid Lagrangian-Eulerian coordinates of Andrews and McIntyre (1978).}\)
the right hand side has been written using the lateral-divergence operator in density coordinates. Substitution of the GM90 stream function (see Appendix A) into the last term of Eq. (24) yields $\nabla (\kappa \nabla \tau)$, which is why the GM90 parametrization is formally referred to as “the lateral diffusion of the depth of each isopycnal surface” (Gent et al. 1995; Gent 2011) rather than the so-called thickness diffusion.

### Appendix C. Interpreting the Mean Potential Energy of the Present Study

Using a Taylor expansion in the vertical direction, we express the Eulerian mean of an arbitrary quantity $A$ (averaged in $z$ coordinates) in terms of quantities averaged in density coordinates:

$$\bar{A} = \hat{A} + \left[ \frac{\partial \bar{A}}{\partial z} \right]_{\bar{\rho}} z + \left[ \frac{\partial^2 \bar{A}}{\partial z^2} \right]_{\bar{\rho}} z^2 + \cdots,$$

where $z = z^m + z^\rho$ is the height of an isopycnal surface and $z^\rho = \partial z/\partial \rho$ is the thickness between adjacent isopycnal surfaces. The triple and higher product of perturbation quantities have been omitted in Eq. (25). The depth integral of Eq. (25) yields $\int_{-H_\rho}^{H_\rho} \bar{A} \, dz = \int_{-H_\rho}^{H_\rho} \hat{A} \, dz$,

$$g \left[ \bar{\tau} \right] = g \left[ \hat{\tau} \right] + g \left[ \frac{\partial \bar{\tau}}{\partial \rho} \right] + \left[ \frac{\partial^2 \bar{\tau}}{\partial \rho^2} \right] + \cdots,$$

which is consistent with the pile-up rule (Aiki and Yamagata 2006). In other words, the depth integral of the second term on the last line of Eq. (25) vanishes with $z^\rho = 0$ at $\bar{\rho} = 0$ and $-H_\rho$, no matter how the bottom is sloped. We substitute $A = g\bar{\tau}$ into Eq. (25):

$$g \left[ \bar{\tau} \right] = g \left[ \hat{\tau} \right] + g \left[ \frac{\partial \bar{\tau}}{\partial \rho} \right] + g \left[ \frac{\partial^2 \bar{\tau}}{\partial \rho^2} \right] + \cdots.$$
no matter how the bottom is sloped. To summarize, \( g\bar{\omega} \) is at the heart of the mean PE in Eq. (16a). While the total (mean plus eddy) PE is associated with the Eulerian mean density \( \overline{\rho} \), the eddy PE is associated with the density difference \( \overline{\rho} - \bar{\rho} \).

Appendix D. Approximate Equations for the Available Potential Energy in Previous Studies

An Eulerian approximation for the available potential energy (APE) has been widely used in previous studies,

\[
g \frac{(\rho - \overline{\rho})^2}{2 \left(-\partial_z \overline{\rho}^b\right)} \quad (28)
\]

where the global background density \( \overline{\rho}^b = \overline{\rho}(z) \) varies only in the vertical direction. Application of an Eulerian time mean to (28) yields

\[
g \frac{(\rho - \overline{\rho})^2}{2 \left(-\partial_z \overline{\rho}^b\right)} + g \frac{(\overline{\rho} - \overline{\rho}^b)^2}{2 \left(-\partial_z \overline{\rho}^b\right)} + 2 \frac{(\overline{\rho} - \overline{\rho}^b)^2}{2 \left(-\partial_z \overline{\rho}^b\right)} \quad A_{\text{mean}}^{\text{APE}} + A_{\text{eddy}}^{\text{APE}} \quad (29)
\]

where the symbols \( A_{\text{mean}} \) and \( A_{\text{eddy}} \) represent traditional definitions for the mean APE and the eddy APE, respectively (Böning and Budish 2012). However, it is rather difficult to derive exact prognostic equations for the mean APE and the eddy APE in Eq. (29). A compromise in the previous studies (Olbers et al. 2012; von Storch et al. 2012) is to approximate Eq. (13b) as

\[
(\partial_t + \nabla \cdot \nabla_2)A_{\text{mean}} + g \frac{\nabla \cdot \nabla_2 (\rho^b - \overline{\rho}^b)}{\left(-\partial_z \overline{\rho}^b\right)} = -g \frac{\nabla \cdot \nabla_2 \overline{\rho}^b}{\partial_z \overline{\rho}^b} \frac{\rho}{\overline{\rho}^b} g \overline{\rho}^b + g \overline{\rho}^b (\overline{\rho}^b - \overline{\rho}^b) \quad \text{A}_{\text{mean}}^{\text{APE}},
\]

\[
(\partial_t + \nabla \cdot \nabla_2)A_{\text{eddy}} + g \frac{\nabla \cdot \nabla_2 (\rho^b - \overline{\rho}^b)}{\left(-\partial_z \overline{\rho}^b\right)} = g \frac{\nabla \cdot \nabla_2 \overline{\rho}^b}{\partial_z \overline{\rho}^b} g \overline{\rho}^b + g \overline{\rho}^b (\overline{\rho}^b - \overline{\rho}^b) \quad \text{A}_{\text{eddy}}^{\text{APE}}.
\]

Fig. 8. Views of (a) the raw density \( \rho(z,t) \), showing the vertical fluctuation of a density surface in a two-density fluid; (b) the Eulerian mean density \( \overline{\rho}(z,t) \), which is given by the fixed-height temporal average; and (c) the mean-height density \( \rho(z,t) = \bar{\rho}(z) \) (2006), as it is given by the distribution of the third line has been derived using \( \bar{\rho} = 1/\rho \). The quantity \( \bar{\rho} \) is referred to as the mean-height density \( \bar{\rho} \). As is shown in Fig. 8. Usually the mean-height density \( \bar{\rho} \) represents a sharper stratification than the Eulerian mean density \( \overline{\rho} \).

Using Eq. (26), we interpret the mean PE in Eq. (16a) as

\[
\overline{\rho} = \overline{\rho}^b + \overline{\rho}^b = \overline{\rho}^b + \rho - \frac{1}{2} \frac{\left(-\partial_z \overline{\rho}^b\right)}{\left(-\partial_z \overline{\rho}^b\right)}
\]

where the first term on the last line is based on the mean-height density \( \bar{\rho} \). The second term on the last line of Eq. (27) vanishes when the depth integral is taken in each vertical column.
which contain three concerns. First, the advection term of Eqs. (30a) and (30b) is based on only the horizontal component of velocity, which is not suitable for calculating the volume budget in a domain with a sloping bottom. Second, the baroclinic conversion term [i.e. the first term on the right hand side of each of (30a) and (30b)] depends on the global background density $\rho_{gb}$. Third, the definition of the eddy APE in Eq. (29) depends on the global background density $\rho_{gb}$.

The above three concerns have been resolved in the PE equations (17a) and (19a) of the present study.

References


Kuhlbrodt, T.; A. Griseel, M. Montoya, A. Levermann, M. Hofmann, and S. Rahmstorf,


