PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Review

Cite this article: Thomas J. 2023 Turbulent wave-balance exchanges in the ocean. *Proc. R. Soc. A* **479**: 20220565.
https://doi.org/10.1098/rspa.2022.0565

Received: 26 August 2022
Accepted: 21 June 2023

Subject Areas:
oceanography, fluid mechanics, wave motion

Keywords:
internal gravity waves, balanced flows, turbulence

Author for correspondence:
Jim Thomas
e-mail: jimthomas.edu@gmail.com

Turbulent wave-balance exchanges in the ocean

Jim Thomas[1,2]

[1]International Centre for Theoretical Sciences, and [2]Centre for Applicable Mathematics, Tata Institute of Fundamental Research, Bangalore, India

JT, 0000-0002-5431-1619

Oceanic flows are turbulent and multi-scale in nature, and are composed of fast internal waves and slowly evolving balanced eddies. Contrary to conventional wisdom in physical oceanography, the past two decades of *in situ*, satellite altimeter and realistically forced global scale ocean model outputs have revealed that internal gravity waves can have comparable or higher energy levels than geostrophically balanced flows at 10–100 km scales in different parts of the world's oceans. These relatively recent findings have fuelled a wide range of research activities aimed at understanding how fast internal gravity waves interact with slowly evolving balanced flows, particularly with the goal of deducing whether internal waves can form an energy sink for oceanic balanced flows. In this paper, we comprehensively review theoretical, numerical and observational investigations undertaken to study internal wave-balance flow exchanges. Theoretical calculations, inspired by different wave-balance regimes seen in observational and global ocean model outputs, are used to point out that internal waves can affect balanced flow dynamics. The theoretical results are followed up by a detailed discussion of numerical results on wave-balance interactions in a broad set of parameter regimes. The numerical results reveal how different kinds of waves exchange energy with balance flow, affect energy flux across scales of balanced flow and facilitate the generation of small-scale dissipative balanced flow structures. The numerical simulation results and global internal wave energy and balanced energy maps are used to conjecture that out of the 0.8 TW of power

THE ROYAL SOCIETY PUBLISHING

going to balanced flow kinetic energy in the ocean, at least 0.1 TW could be dissipated by internal gravity waves. We therefore hypothesize that internal waves can form a non-negligible energy sink for balanced flow in the world's oceans.

1. Introduction

Oceanic mesoscale flow, with O(100) km scales, is constrained by Earth's rotation and density stratification and is the largest reservoir for total oceanic flow kinetic energy [1–3]. The mesoscale flow is an intertwined mixture of fast evolving internal gravity waves, whose time scales range from a few minutes to a day, and a slowly evolving balanced component, whose time scales range from weeks to months. Furthermore, the mesoscale flow is characterized by asymptotically small Rossby number, this dimensionless number being the ratio of inertial force to Coriolis force in the flow. The smallness of the Rossby number at mesoscales makes it possible to decompose the flow into fast internal gravity waves and the slow balanced component. Internal gravity waves are dispersive in nature and are primarily generated by atmospheric winds and gravitational tides [4,5]. On the other hand, the balanced flow, named so because it is in geostrophic balance (lateral pressure gradient balances the Coriolis force) and hydrostatic balance (vertical pressure gradient balances buoyancy force arising due to vertical density stratification), is primarily generated by atmospheric forcing [6,7].

The slow balanced flow component organizes itself into O(100) km coherent vortices known as mesoscale eddies. These coherent eddies contain close to 90% of the mesoscale kinetic energy and act as agents that transport mass and tracers across the oceans [3,8,9]. Since internal waves evolve, propagate and dissipate on a much shorter time scale than the slowly evolving balanced flow, it would seem reasonable to track the balanced component of the flow and ignore fast internal gravity waves for predicting slow long-term dynamics of the flow. Such an attempt results in a single equation model for the balanced component that is devoid of fast waves: the quasi-geostrophic (QG) equation.

Historically, the need for a reduced mathematical model that can track the evolution of the balanced component of geophysical flows was recognized more than a century back. L. F. Richardson made the first and original attempt at weather prediction by numerically integrating the full equations of motion [10,11]. Unfortunately, his numerical integration used unbalanced initial conditions and time steps larger than that allowed by the numerical stability criterion, resulting in the time integration diverging with non-physical solutions. Realizing the need to filter waves and track the slow balanced dynamics resulted in the derivation of the QG equation [12,13], which was used for the first set of successful computer weather forecasts [14,15]. In the early days of weather prediction when computational resources were limited, it was necessary to filter waves and concentrate on the slow balanced dynamics; specifically since not having to resolve fast internal waves allowed numerical integrations with larger time steps. Consequently, even when the full primitive equations were integrated, it was common practice to carefully initialize numerical integrations so as to minimize the generation of fast waves [16,17]. Such pursuits inspired the introduction of the slow manifold: an invariant manifold in state space that was devoid of fast wave activity [18,19].

Although one could define a slow manifold for an arbitrary dynamical system without much difficulty, in practice, it is extremely challenging to obtain solutions of the governing equations that remain on the slow manifold as time progresses [20–22]. Simply put, this means that even though the initial conditions can be carefully designed to be in balance, there is no guarantee that fast waves will remain suppressed as the system evolves. Extensive research in this direction based on idealized dynamical system models, asymptotic and numerical investigations of primitive equations, and laboratory experiments have pointed out that, however well-balanced initial conditions might be, fast internal gravity waves will be spontaneously excited as the flow evolves [23–27]. Nevertheless, the strength of waves and other unbalanced flow components that

are excited can be insignificantly small in the small Rossby number regime if the flow is well balanced initially (see discussions and references in [28]). Therefore, even though a perfect slow manifold that is completely devoid of fast waves might be unrealistic, a balanced flow in the small Rossby number regime can evolve in an 'almost' balanced state with insignificant fast wave activity. As a result, balanced models such as the QG equation and its variants can capture the dynamics of almost the entire flow in the small Rossby number regime, as long as wave energy levels remain small.

Seventy-five years since the first derivation of QG equation by Charney [12], *in situ* measurements, satellite altimeter datasets, and realistically forced global ocean models have revolutionized the field of physical oceanography [29–34]. One of the key features that have emerged across observations and model outputs is the dominant presence of high energy fast internal gravity waves at O(10–100) km scales in the world's oceans. Specifically, oceanic datasets of the past two decades have revealed that internal waves can have scales comparable with balanced eddy scales and can be as energetic or sometimes even more energetic than the local balanced flow [35–45]. The physical oceanographer's view has therefore changed from thinking of flows in the ocean as largely composed of large-scale balanced eddies and small-scale internal gravity waves to one where fast waves and slow balanced eddies coexist in the ocean with comparable spatial scales and energy levels (see figure 1 and caption there). This updated view of the ocean that has emerged over the past few decades has challenged the previous paradigms of QG turbulence dominating throughout the world's oceans. Given the oceanic data sets uncovered in recent times, it is imperative to understand how energetic fast internal gravity waves affect the balanced flow and modify QG turbulence phenomenology for reasons detailed below.

About 0.8 TW of power is input into the geostrophically balanced flow kinetic energy across the global oceans [3]. As mentioned earlier, the balanced flow dynamics can be captured by the QG equation and the QG equation is characterized by an upscale transfer of energy, similar to the dynamics of the two-dimensional Euler equation. The upscale energy transfer results in balanced energy accumulating at O(100 km) mesoscales, eight decades away from O(mm) viscous Kolmogorov scales where flow kinetic energy can get dissipated. Mechanisms that can form an energy sink for balanced flow are therefore actively sought after, this being a crucial ingredient for closing the overall oceanic energy budget [47,48]. Dominant mechanisms uncovered so far involve balanced eddies interacting with different forms of boundaries and boundary layers [49–56]. Quite naturally, none of these would be operational in the open ocean and interior parts of the ocean. If internal waves interact with balanced flow and facilitate a forward energy flux of balanced flow, that would be a mechanism for small-scale viscous dissipation of balanced flow without involving boundaries. Consequently, wave-induced forward flux and subsequent dissipation of balanced flow could be an energy sink for balanced flow in oceanic regions away from boundaries.

In addition to assisting in the forward flux and small-scale dissipation of balanced energy, internal waves are also often hypothesized to be capable of directly extracting energy from balanced flow [57,58]. This implies the possibility that waves might extract energy directly from large-scale balanced flow without necessarily forcing balanced energy to move to small scales. If this hypothesis is proven to be true, direct extraction of balanced energy would be an energy source for waves. Although internal waves are excited at large scales of O (100–1000) km by winds and tides, the forward energy flux and subsequent dissipation of wave energy at small viscous scales is considered to be the main mechanism for small-scale ocean mixing [59,60]. If waves gain a significant amount of energy from balanced flow, the extra wave energy gained from balanced flow would be fluxed to small scales and be dissipated along with the wave fields generated by winds and tides, leading to an additional energy source for small-scale mixing by wave breaking. Finally, a deeper understanding of the role waves play in the ocean and specifically on the balanced flow will generically help in improving large-scale ocean and climate models that are still far from resolving fast wave fields. Given that there are multiple indications of wave parametrizations making notable difference in predictions of large climate-scale models [61–65],

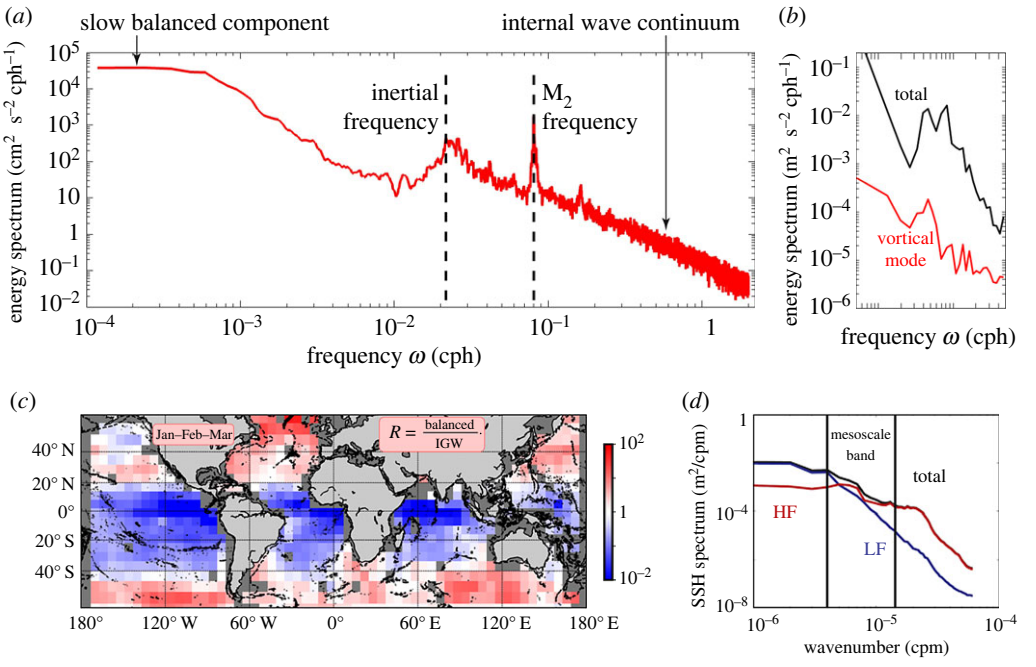

Figure 1. Figure shows fast waves and the slow balanced flow component in oceanic flows in different scenarios. Panel (*a*) shows a generic oceanic frequency spectrum from *in situ* measurements discussed in [46]. Observe the plateau of energy at low frequency end on the left, which corresponds to the slow balanced flow. The right side of the figure shows the near-inertial peak, M_2 tidal peak and a broad spectrum of high-frequency waves identified as the internal wave continuum. Panel (*b*) shows the energy spectrum in frequency space based on *in situ* measurements detailed in [43]. The black curve is the total energy spectrum and red curve denotes the vortical or geostrophic energy spectrum, implying that waves dominate the flow. Panel (*c*) is adapted from [42] with permission, copyright 2018 John Wiley and Sons, and shows the balance to wave energy ratio in the 10–50 km band based on a realistically forced global scale ocean simulation. Note that in certain oceanic regions wave energy levels can be up to two orders of magnitude higher than balanced flow and vice versa in some regions. Panel (*d*) is adapted from [35] with permission, copyright 2012 John Wiley and Sons, and shows the sea surface height (SSH) spectrum (black curve) in wavenumber space obtained from global scale ocean model simulations. The spectrum is decomposed into a high-frequency component (HF, red curve) and a low-frequency component (LF, blue curve). Note that the low-frequency component dominates at the largest scales while the high-frequency component starts overtaking the slow component at O(100 km) scales. The scale at which fast waves starts dominating over the slow balanced component is highly variable in the ocean and is sensitive to the geographical location and the season of the year.

it is imperative to develop a comprehensive understanding of the different roles played by internal waves in the ocean so as to improve the predictions of large climate-scale models.

The plan for the paper is as follows: an asymptotic analysis on wave-balance interactions is discussed in §2, results based on numerical simulations emphasizing wave-balance energy exchanges is detailed in §3, related findings from oceanic observations are discussed in §4, and comparisons between interactions in low and O(1) Rossby number flows are given in §5. The paper then concludes by discussing (a) approximate estimates for global dissipation of balanced energy due to internal wave induced and other mechanisms and (b) unresolved challenges for future studies in §6.

2. Asymptotic results on wave-balance interactions

In this section, we will use small Rossby number asymptotic analysis to derive equations that demonstrate the effect of waves on balanced flow. The leading-order slow evolution equation

for balanced flow is the QG equation that has been derived in multiple studies in the past [12,13,66]. A key feature of linear internal gravity waves is that they do not project on the leading-order part of the potential vorticity field. Consequently, the balanced flow evolves according to the QG equation unaffected by waves [67–69]. Although the balanced flow evolves unaffected by waves at leading order, higher-order nonlinear wave interactions can affect balanced flow evolution. This was pointed out using rotating shallow water equations by Thomas [70]. For the parameter regime where waves are asymptotically stronger than balanced flow, Wagner & Young [71] derived an equation for balanced flow that contains wave interaction terms. Below we combine the asymptotic analysis of [69–71] to derive an evolution equation for the slow balanced flow that is affected by fast waves.

The f-plane hydrostatically balanced Boussinesq equations in non-dimensional form are

$$\frac{\partial \boldsymbol{v}}{\partial t} + \hat{\boldsymbol{z}} \times \boldsymbol{v} + \nabla p + \epsilon \left(\frac{\partial \boldsymbol{v}}{\partial T} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} + w \frac{\partial \boldsymbol{v}}{\partial z} \right) = 0, \tag{2.1a}$$

$$\frac{\partial b}{\partial t} + N^2 w + \epsilon \left(\frac{\partial b}{\partial T} + \boldsymbol{v} \cdot \nabla b + w \frac{\partial b}{\partial z} \right) = 0, \tag{2.1b}$$

$$\frac{\partial p}{\partial z} = b \tag{2.1c}$$

and

$$\nabla \cdot \boldsymbol{v} + \frac{\partial w}{\partial z} = 0, \tag{2.1d}$$

where $\epsilon = U/fL$ is the Rossby number and $T = \epsilon t$ is a slow time variable. The above system of equations conserve potential vorticity (PV) on fluid parcels [7]. Recognizing the fact that PV has non-zero contributions even in a fluid at rest, Wagner & Young [71] introduced the concept of Available Potential Vorticity (APV). The APV equation corresponding to (2.1) in non-dimensional form is

$$\frac{\partial \Pi}{\partial t} + \epsilon \left(\frac{\partial \Pi}{\partial T} + \frac{\partial \Gamma}{\partial t} + \left(\boldsymbol{v} \cdot \nabla + w \frac{\partial}{\partial z} \right) \Pi \right) + \epsilon^2 \left(\frac{\partial \Theta}{\partial t} + \frac{\partial \Gamma}{\partial T} + \left(\boldsymbol{v} \cdot \nabla + w \frac{\partial}{\partial z} \right) \Gamma \right) + O(\epsilon^3) = 0. \tag{2.2}$$

We refer the reader to appendix A for details required to arrive at (2.1) and (2.2) along with an elaborate definition of variables that appear in (2.2).

On expanding all variables asymptotically, for example, writing $\boldsymbol{v} = \boldsymbol{v}_0 + \epsilon \boldsymbol{v}_1 + \epsilon^2 \boldsymbol{v}_2 + \cdots$ for velocity, the $O(1)$ equations we get from (2.1) are linear. We write the solution of the linear equations as the sum of slow balanced part (denoted by an overbar hereafter) and fast linear waves (denoted by a prime hereafter). For example, the velocity field is expanded as $\boldsymbol{v}_0 = \overline{\boldsymbol{v}}_0 + \boldsymbol{v}_0'$. The slow balanced flow and fast waves satisfy the linear equations:

$$\left.\begin{aligned}
\frac{\partial \boldsymbol{v}_0'}{\partial t} + \hat{\boldsymbol{z}} \times \boldsymbol{v}_0' + \nabla p_0' &= 0 \quad & \hat{\boldsymbol{z}} \times \overline{\boldsymbol{v}}_0 + \nabla \overline{p}_0 &= 0 \\
\frac{\partial b_0'}{\partial t} + N^2 w_0' &= 0 & \overline{w}_0 &= 0 \\
\frac{\partial p_0'}{\partial z} &= b_0' & \frac{\partial \overline{p}_0}{\partial z} &= \overline{b}_0 \\
\nabla \cdot \boldsymbol{v}_0' + \frac{\partial w_0'}{\partial z} &= 0 & \nabla \cdot \overline{\boldsymbol{v}}_0 &= 0.
\end{aligned}\right\} \tag{2.3}$$

In using the flow decomposition of the form $\boldsymbol{v}_0 = \overline{\boldsymbol{v}}_0 + \boldsymbol{v}_0'$, we enforce that the leading-order wave and the balanced flow fields are of the same asymptotic order. We will identify this regime as the Comparable Wave (CW) regime, meaning that waves are comparable in strength to the balanced flow. Furthermore, it is worth noting that in the above flow decomposition, the wave and the balanced flow component have no spatial scale separation, in spite of them having a time scale separation.

The leading-order equation from (2.2) is

$$\frac{\partial \Pi_0}{\partial t} = 0 \Rightarrow \Pi_0 = \overline{\Pi}_0 \quad \text{and} \quad \Pi_0' = 0. \tag{2.4}$$

The leading-order APV, Π_0, therefore does not have fast fluctuations and is a purely slow field. We proceed asymptotically to $O(\epsilon)$ and eliminate secular growth to obtain (see appendix B for details):

$$\left(\frac{\partial}{\partial T} + \overline{\boldsymbol{v}}_0 \cdot \nabla\right)\left(\Delta \overline{p}_0 + \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial \overline{p}_0}{\partial z}\right)\right) = 0. \tag{2.5}$$

The above equation is the well-known QG equation that governs the slow evolution of balanced flow on $T \sim O(1)$ time scales. Since $\overline{\boldsymbol{v}}_0$ and $\overline{p}_0$ are related through the geostrophic balance relationship in (2.3), the QG equation above is an evolution equation for a single variable, making it an attractive reduced equation that has been used to investigate geostrophic turbulence phenomenology [72–76].

As is clear from the above, the standard QG equation in (2.5) is obtained at $O(\epsilon)$ of the asymptotic analysis. However, by systematically proceeding to $O(\epsilon^2)$ in the asymptotic analysis we can derive a more accurate model that captures the long-term dynamics of the balanced flow. This procedure, detailed in appendix C, gives us

$$\frac{\partial Q_{CW}}{\partial T} + \partial[\psi_{CW}, Q_{CW}] = 0, \tag{2.6a}$$

$$Q_{CW} = Q_{G0} + \epsilon Q_{G1} + \epsilon Q_W, \ \psi_{CW} = \psi_{G0} + \epsilon \psi_{G1} + \epsilon \psi_W, \tag{2.6b}$$

$$Q_{G0} = \Delta \overline{\psi} + \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial \overline{\psi}}{\partial z}\right), \tag{2.6c}$$

$$\begin{aligned} Q_{G1} = \frac{1}{N^2}\Bigg(&\Delta\overline{\psi}\frac{\partial^2\overline{\psi}}{\partial z^2} - \left(\nabla\frac{\partial\overline{\psi}}{\partial z}\right)^2 - \frac{\Lambda''}{2N^2}\left(\frac{\partial\overline{\psi}}{\partial z}\right)^2 \\ &- \frac{1}{N^2}\frac{\partial\overline{\psi}}{\partial z}\frac{\partial}{\partial z}\left(N^2\Delta\overline{\psi} + \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial\overline{\psi}}{\partial z}\right)\right)\Bigg) \\ &- 2\partial\left[\frac{\partial\overline{\psi}}{\partial x}, \frac{\partial\overline{\psi}}{\partial y}\right] - \nabla\overline{\psi}\cdot\nabla\left(\Delta\overline{\psi} + \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial\overline{\psi}}{\partial z}\right)\right), \end{aligned} \tag{2.6d}$$

$$Q_W = \frac{1}{2}\Delta\overline{p}^S - \hat{\boldsymbol{z}}\cdot\nabla\times\overline{\boldsymbol{v}}^S + \frac{1}{N^2}\left(\overline{\zeta_0' b_{0z}'} + \overline{(\hat{\boldsymbol{z}}\times\boldsymbol{v}_{0z}')\cdot\nabla b_0'} - \frac{\Lambda''}{2N^2}\overline{b_0'^2}\right) \tag{2.6e}$$

and

$$\psi_{G0} = \overline{\psi}, \ \psi_{G1} = -\frac{1}{2}(\nabla\overline{\psi})^2 - \frac{1}{2N^2}\left(\frac{\partial\overline{\psi}}{\partial z}\right)^2, \ \psi_W = \frac{\overline{p}^S}{2}. \tag{2.6f}$$

Above, Q_{G0} and ψ_{G0} are the leading-order balanced APV and streamfunction fields, which are linear in the balanced terms, while Q_{G1} and ψ_{G1} are next-order correction terms that are composed of quadratic nonlinear balanced–balanced interaction terms. Q_W and ψ_W are, on the other hand, the contributions of quadratic nonlinear wave–wave interaction terms to the APV and the advecting streamfunction, which are in turn dependent on Stokes velocity $\overline{\boldsymbol{v}}^S$ and Stokes pressure $\overline{p}^S$ (see appendix B for mathematical expressions of Stokes fields).

In comparison to the leading-order QG equation (2.5), (2.6) is an improved higher-order accurate evolution equation for the slow balanced flow that holds for longer timescales. Although many studies have explored improved balanced-model approximations, the higher-order slow equation (2.6) in the above form with nonlinear balanced interaction terms (terms with subscript 'G1') and nonlinear wave interaction terms (terms with subscript 'W') has not been reported elsewhere to the best of the author's knowledge. Nevertheless, several reductions of (2.6) have been derived in the past.

The simplest reduction of (2.6) is obtained by setting $\epsilon = 0$; this gives us the QG equation $\partial Q_{G0}/\partial T + \partial[\psi_{G0}, Q_{G0}] = 0$ as in (2.5). On the other hand, if we drop all the terms that contain wave fields (terms with subscript 'W') in (2.6), we get a higher order version of the QG equation: $\partial(Q_{G0} + \epsilon Q_{G1})/\partial T + \partial[(\psi_{G0} + \epsilon\psi_{G1}), (Q_{G0} + \epsilon Q_{G1})] = 0$. This higher-order improved balanced equation that is unaffected by fast waves is identical to the improved QG equation derived by Zeitlin *et al*. [69] in the geostrophic adjustment setting (see their eqn (3.15)). A similar attempt to include higher-order balanced correction terms, without including internal waves was undertaken by Muraki *et al*. [77], who called their higher-order balanced equation QG+. A notable result that follows from investigating QG+ equations is that higher-order balanced correction terms can modify balanced flow dynamics. For instance, Rotunno *et al*. [78] and Hakim *et al*. [79] show that incorporating higher-order balanced contributions can affect baroclinic instability and can also introduce a cyclone–anticyclone asymmetry in vorticity structures. The asymmetric dominance of cyclonic vorticity structures is a feature observed in the ocean [80–82] and is not captured by the leading order QG equation (2.5) which maintains a cyclone–anticyclone symmetry.

Given that higher-order balanced terms can modify flow features that follow from the QG equation (2.5), we anticipate that the presence of waves in (2.6) would affect balanced dynamics, making it difficult to envision a slow manifold unaffected by fast wave fluctuations on long time scales [18–21]. In other words, for unbalanced initial conditions that can generate fast waves and slow balanced flow, the slow dynamics will be affected by fast modes and the slow dynamics' equation cannot be closed by including the slow variables alone.

Note that in the CW regime that we explored so far, where wave and balanced flow have comparable strength, the effect of waves on the balanced flow is weak since all the wave interaction terms (i.e. terms with subscript 'W') are pre-multiplied by ϵ in (2.6*b*). The wave-induced effects therefore arise as next-order asymptotic correction terms to the QG equation. A parameter regime that has received interest over the past decade is the strong wave (SW) regime where waves are stronger than balanced flow. This specific asymptotic regime was studied by Wagner & Young [71], and the appropriate slow evolution equation in this regime can be derived as we proceeded earlier, but by enforcing the leading-order fields to be purely waves and balanced flow to be an asymptotic order weaker than waves. For velocity field for example, the asymptotic expansion for this regime would be $\boldsymbol{v}_0 = \boldsymbol{v}_0' + \epsilon\overline{\boldsymbol{v}}_0 + \cdots$. Since the balanced field is $O(\epsilon)$ in this regime, the appropriate slow time scale becomes $\tau = \epsilon^2 t$. Using this and proceeding with the asymptotic analysis as before gives us the slow evolution equation

$$\frac{\partial Q_{SW}}{\partial \tau} + \partial[\psi_{SW}, Q_{SW}] = 0 \tag{2.7a}$$

and

$$Q_{SW} = Q_{G0} + Q_W, \ \psi_{SW} = \psi_{G0} + \psi_W \tag{2.7b}$$

The above equation (2.7) was derived by Wagner & Young [71]; see their eqns (4.52), (4.53) and (4.54). The same equation above was also derived by Salmon [83] using variational methods. As can be seen above, the SW equation is much simpler and more compact than the CW equation (2.6) since higher-order balanced flow interactions (terms with subscript G_1 in (2.6)) do not appear in this equation. This SW regime has also received interest in the surface gravity waves community in connection to Langmuir circulation. Equations similar to the SW equation above were derived by Craik & Leibovich [84] and Leibovich [85] to examine surface wave-vortical mode interactions.

The other extreme end of wave-balance interactions is the weak wave (WW) regime where a leading-order balanced flow exists on top of which small amplitude waves are excited. This regime was popularly studied in the past to examine how a small amplitude wave field could affect balanced dynamics on an extremely long time scale (e.g. [57,86]). As seen from (2.6), since all the wave interaction terms are pre-multiplied by ϵ, the effects of waves on balanced flow are weak when waves are comparable in strength to balanced flow. The impact of waves on balanced flow in the WW regime, where the waves are a small perturbation to a leading order balanced flow, would therefore be much less than in the CW regime.

We conclude this section by pointing out that the slow balanced flow equations (2.6) and (2.7) can be derived for other geophysical fluid dynamic models. For instance, similar slow balanced equations for the rotating shallow water model can be found in ch. 3 of [87]. It is important to note that balanced equations such as (2.6) and (2.7) are not closed, since the evolution equation for the fast wave fields have not been derived yet. Typically wave equations that hold for time scales much longer than the linear wave time period are derived assuming a spatial scale separation between waves and mean flow or by concentrating on a few discrete wave modes [88–90]. Deriving slow equations for a broad spectrum of waves when wave and balanced flow share similar spatial scales is in general extremely challenging. Furthermore, one of the benefits of having coupled wave-balance equations is that energy exchanges between the two fields can be examined. However, the accuracy of wave-balance energy transfer estimates that can be obtained from asymptotic models such as (2.6) and (2.7) coupled with asymptotically approximated wave evolution equations is *a priori* unclear, especially in turbulent oceanic flows spanning multiple spatio-temporal scales. Given these caveats, we might consider slow balanced equations such as (2.6) and (2.7) primarily as indicators of waves affecting balanced dynamics. Direct numerical integration of the governing equations is inevitable to get a firm grip on estimates on turbulent wave-balance energy exchanges. In the following section, we survey numerical results on wave-balance exchanges undertaken over the past few years.

3. Wave-balance exchanges from numerical experiments

In this section, we will review numerical investigation results on wave-balance interactions. As can be seen in the example frequency spectrum shown in figure 1*a*, near-inertial waves, internal tides and the internal wave continuum are prominent internal wave signals observed in the ocean and depending on the geographical location and season, the strength of these waves can vary significantly. Below we will examine how these different wave components interact with the slow balanced component.

(a) Low mode internal tides

The gravitational attractive force of the sun and the moon on the earth results in the excitation of barotropic or external tides on earth. Out of all the different solar and lunar tidal components, the dominant share of the total tidal power input goes into the lunar M_2 tidal component with a time scale of 12 h and 25 min [59,91]. The dominance of this tidal component is often marked by a distinguished peak in oceanic frequency spectra, as can be seen in figure 1*a*. Given that this tidal source pumps O(1 TW) of power into the ocean, mechanisms and locations of tidal dissipation has been an active research question in oceanography. Theoretical calculations going back almost a century led oceanographers to believe that the barotropic tide dissipated most of its energy in shallow oceanographic regions via bottom drag [92,93]. This conventional wisdom however changed as a result of the landmark TOPEX/POSEIDON satellite mission. Analysis of altimeter data from the TOPEX/POSEIDON mission revealed that about a third of the barotropic tidal energy gets converted to baroclinic tides in open oceanographic regions, this barotropic to baroclinic tidal conversion acting as an energy sink for barotropic tides. The scattering of the barotropic tide at bottom topographic features such as sea mounts is the primary mechanism for the generation of baroclinic tides [94–96], also known as internal tides (ITs hereafter).

The first-mode baroclinic tide contains the dominant fraction of IT energy and is often seen to propagate thousands of kilometres away from its generation site [97–99]. Similarly, the geostrophic balanced flow harbours a major share of its energy in the barotropic and the first baroclinic mode, this being inferred from theoretical calculations, observations and idealized numerical simulations [74,76,100,101]. Therefore, both the balanced flow and internal tides share comparable spatial scales of O(100 km), leading to two-way interactions between the two fields: balanced flow can affect baroclinic tides and baroclinic tides can affect balanced dynamics.

Among the studies that have examined this interaction, major share of the work has gone into exploring the effect of balanced mesoscale flow on the propagation of baroclinic tides.

As low-mode ITs propagate through the ocean, the background mesoscale eddy field can modulate, inhomogenize and scatter the waves. This process by which a tidal wave field can become incoherent due to scattering by balanced flow and the energetics associated with the process has been investigated based on *in situ* measurements, satellite altimeter datasets and numerical simulations [98,99,102–109]. These studies reveal that only a small fraction of low mode tidal energy, typically less than 5% wave energy, is lost to higher baroclinic wave modes in the low Rossby number regime as the wave propagates O(100–1000) km. The primary result of eddy-low mode tide interaction is the scattering of wave energy into other horizontal wavenumbers, thereby inhomogenizing the wave field. The scattering of ITs by the vortical mode can be intuitively seen using resonant interaction theory. In this formalism, a vortical mode can act as a catalyst in transferring energy between wave modes with the same frequency propagating in different directions [110–112]. Significant efforts have been channelled in this research direction primarily because the incoherent tidal signal makes it difficult to analyse satellite altimeter datasets and unambiguously extract the balanced flow component from the total flow field. Consequently, having access to a simpler model to track the loss of coherence of a tide propagating through an eddy field is highly useful. Motivated by this, Wagner *et al*. [114] derived an amplitude equation for capturing the scattering of low mode tides by balanced eddies. With a similar goal of developing a reduced model, Savva & Vanneste [113] derived a kinetic wave equation for IT scattering by mesoscale flows. These reduced models can be conveniently integrated with lower computational resources, making them more useful for practical purposes.

An illustration of tidal scattering from Wagner *et al*. [114] is shown in figure 2, with panels 2*a–d* showing the wave velocity magnitude and panels 2*f–i* showing the spectrum of the wave field as time increases from left to right. Observe from panels 2*a–d* how the wave field becomes incoherent as time progresses. For the internal wave dispersion relationship $\omega = \sqrt{f^2 + N^2 k_h^2/k_z^2}$, choosing the vertical mode to be the first baroclinic mode leads to all horizontal wavenumbers that lie on a circle (i.e. a constant value for k_h) to have the same tidal frequency. The scattering process leads to the transfer of energy to new horizontal wavenumbers of the same tidal frequency, this process leading to the gradual illumination of a circle in spectral space. Of course, the presence of near-resonant and non-resonant exchanges changes the circle to a ring of finite thickness. This scattering process resulting in the wave spectrum evolving towards a ring in spectral space can be seen in panels 2*f–i*. Furthermore, the spectral ring formation of the IT field is often seen in realistic oceanographic data. An example is shown in figure 2*e*, obtained using satellite altimeter data discussed in [115]. It is noteworthy that contrary to the idealized numerical result shown in panels 2*a–d* and 2*f–i* that isolates the role of balanced eddies on internal tide, oceanic tidal spectra such as the ones shown in figure 2*e* have more realistic ingredients that are difficult to separate. In addition to waves being scattered by eddies, realistic spectra such as figure 2*e* is the result of internal tides interacting with bottom topographic features and interference with other waves in the ocean. Separating and differentiating between different causes of tidal scattering from realistic datasets is still a challenge. Nevertheless, we note that the major effect of the mesoscale flow on the IT field is to scatter and inhomogenize the waves, this effect being well studied using primitive equation models and being captured by reduced mathematical models.

In addition to ITs being inhomogenized and scattered by the balanced flow, ITs can affect balanced flow dynamics. Given that the energy levels of geostrophic balanced flow and low mode ITs are highly variable over the global ocean [47,99], depending on the specific location, wave energy can be comparable to or higher than balanced energy. Inspired by these different relative energy regimes, Thomas & Yamada [116] examined how the first baroclinic mode ITs affect balanced flow dynamics and modify the two-mode QG phenomenology. Figure 3 provides the main results of [116]. Figure 3*a* shows the barotropic vorticity field in the CW regime: notice that the flow is rich with large-scale coherent vortices showing tendencies to merge and grow further. Despite waves being present, they have little effect on the turbulent dynamics of the

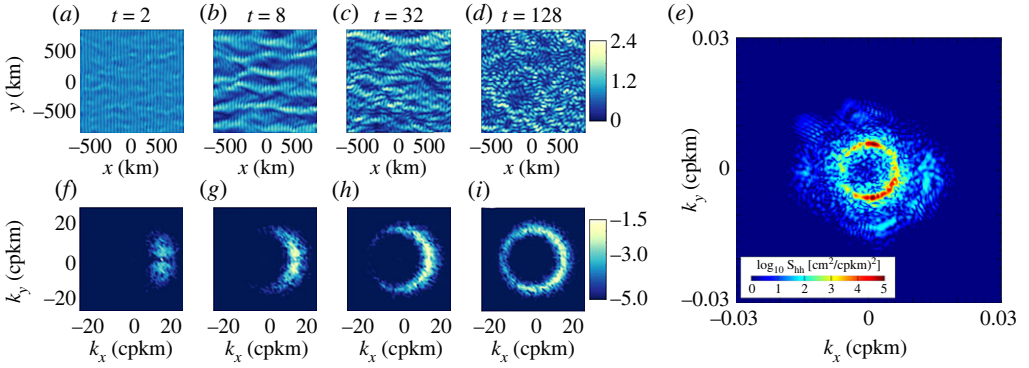

Figure 2. Panels (*a–d*) and (*f–i*) are adapted from [114] with permission, copyright 2017 Cambridge University Press, and show the scattering of the first baroclinic mode tide by a geostrophic flow. Panels (*a–d*) show the magnitude of wave velocity in m s^{-1} at different times. Notice how the wave field becomes scattered and inhomogenized as time increases from left to right. Panels (*f–i*) show the log of normalized spectrum of the wave field at different times. As the wave field gets scattered by the eddy field, the spectrum illuminates a ring. Panel (*e*) shows the spectrum of tidal flow from satellite altimeter dataset detailed in [115]. Note the bright ring-like structure in the spectrum.

balanced flow in the CW regime. The baroclinic balanced mode transfers energy to the barotropic mode in this regime, as in the QG regime. By contrast, figure 3*b* shows the barotropic vorticity field in the SW regime. High energy waves destroy large-scale coherent vortices resulting in the formation of energetic small-scale vortical structures throughout the domain. Figure 3*c* shows the energy transfers between different modes in the SW regime. The red arrows indicate all the energy transfers that appear in the presence of high energy waves, while the blue arrow indicates the energy transfer between balanced modes alone. The balanced energy can increase by a factor of 2–10 over a few hundred eddy turn over time scales due to energy transfer from waves. By contrast, both QG and CW regimes are characterized by the energy transfer between balanced modes alone, as indicated by the blue arrow. Figure 3*c* may therefore be considered as a comprehensive illustration of the barotropic mode-baroclinic mode energy transfers. If wave energy is not high enough, such as in the CW regime, or if waves are absent, as in the QG regime, transfers marked by the red arrows disappear and only the transfer indicated by the blue arrow remains. The well-known two-mode QG turbulence phenomenology may therefore be considered to be a subset of figure 3*c* obtained by dropping waves.

(b) Near-inertial waves

The near-inertial wave (NIW) band in the ocean receives about 0.3–1.5 TW of global power input due to the wind work on the upper ocean [117–119], and this power input is comparable to the global tidal power input. The high power input into the NIW band leads to a distinguished peak around the inertial frequency in typical oceanic frequency spectra, as can be seen for example in figure 1*a*. Primarily generated by atmospheric winds and storms, NIWs are ubiquitous in the upper ocean [4,120]. The extremely large horizontal scales of these waves and their small vertical scales result in NIWs having relatively low propagation speeds. Their prominent presence in the upper ocean, their low propagation speeds, and their high-energy content make NIWs attractive candidates for energy exchange with mesoscale eddy fields. Consequently, NIWs have been the most studied oceanic wave field in connection to wave-balance energy exchanges. Owing to their large horizontal scales and small vertical scales, i.e. $k_h/k_z \ll 1$, the internal waves' dispersion relationship $\omega = \sqrt{f^2 + N^2 k_h^2/k_z^2}$ simplifies as $\omega \approx f$ for NIWs, making their frequencies confined to the inertial frequency f. As a result, the wave field can be approximated as $u' + \mathrm{i}v' = A(\boldsymbol{x}, z, T)\,\mathrm{e}^{-\mathrm{i}ft}$, where t is a fast time scale of the order of the inertial period while T is a slow time scale of the

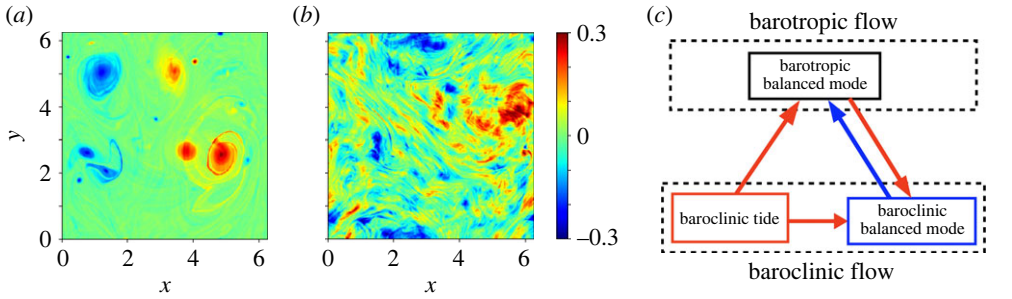

Figure 3. Figure, based on results detailed in [116], shows the effect of the first baroclinic mode tide on balanced flow. Panel (*a*) shows the non-dimensional barotropic vorticity field, i.e. the barotropic vorticity normalized by the rotation rate f, in the CW regime. Panel (*b*) similarly shows the non-dimensional barotropic vorticity field in the SW regime. Panel (*c*) shows the barotropic–baroclinic energy transfers. The blue arrow shows energy transfer from baroclinic to barotropic mode. This is the only energy transfer in the QG regime. Additionally, this is the only noticeable energy transfer in the CW regime, since wave-balance energy transfers are negligible in this regime. The red arrows show all the new transfers that appear in the presence of waves in the SW regime. Waves primarily transfer energy to the balanced flow, both barotropic and baroclinic, resulting in the destruction of large-scale coherent vortices as seen in panel (*b*).

order of eddy turn over time scale. In the low Rossby number regime assuming a time scale separation between fast NIWs and slow eddy field, using the method of multiscale asymptotics Young & Ben Jelloul [121] derived an amplitude equation for the NIW amplitude, which can be written in abstract form as

$$\frac{\partial A}{\partial T} + \mathcal{L}(A) + \mathcal{B}(\psi, A) = 0, \tag{3.1}$$

where $\mathcal{L}$ is a linear operator and $\mathcal{B}(\psi, A)$ captures the effect of the balanced flow on waves. Equation (3.1) is often identified as the Young and Ben Jelloul (YBJ) equation, and it captures the slow evolution of the NIW amplitude in the presence of a balanced eddy field. We wrote the above equation in abstract form since the YBJ equation can be written in different forms in two and three dimensions [121]. Despite being an approximate asymptotic model, the YBJ equation (3.1) is a robust model that can capture the slow advection, refraction, and scattering of NIWs by balanced flows and subsequent reduction in waves' horizontal scales, vertical propagation of NIWs and accumulation of NIWs in anticyclonic vortices [122–131].

A simple illustration of NIWs interacting with a balanced eddy field obtained by the numerical integration of the primitive equations linearized with respect to a balanced flow is shown in figure 4. The same results can be obtained by integrating the appropriate YBJ equation (see discussions in [128]). An idealized barotropic vorticity field that is composed of two cyclonic and two anticyclonic vortices is shown in figure 4*a*. Figures 4*c–e* show the NIW kinetic energy at a fixed vertical depth, 100 m below the upper ocean surface, as time progresses from left to right. The wave field is initially horizontally homogeneous, mimicking extremely large horizontal scales of NIWs. However, interaction with the eddy field shown in figure 4*a* imprints smaller horizontal scales on the wave field. This is followed by the expulsion of the waves from cyclonic vortices and accumulation of waves in anticyclonic vortices and vertical propagation of waves. Since higher concentration of NIWs are associated with anticyclonic vortices, the dominant fraction of NIW energy is transferred to interior parts of the ocean via anticyclonic vortices. This feature is often termed the inertial chimney effect and is found in large-scale ocean model outputs and idealized simulations [128,130,133]. Although the illustration in panels 4*a*, *c–e* is extremely idealized, the particular case was chosen above to clearly highlight the trapping of waves in anticyclonic vortices. Realistic oceanic geostrophic vortices are turbulent and multiscale in nature. Nevertheless, the higher concentration of NIWs in anticyclonic vortices is a generic feature observed in the ocean [132,134–136]. An example from the results based on ocean observations

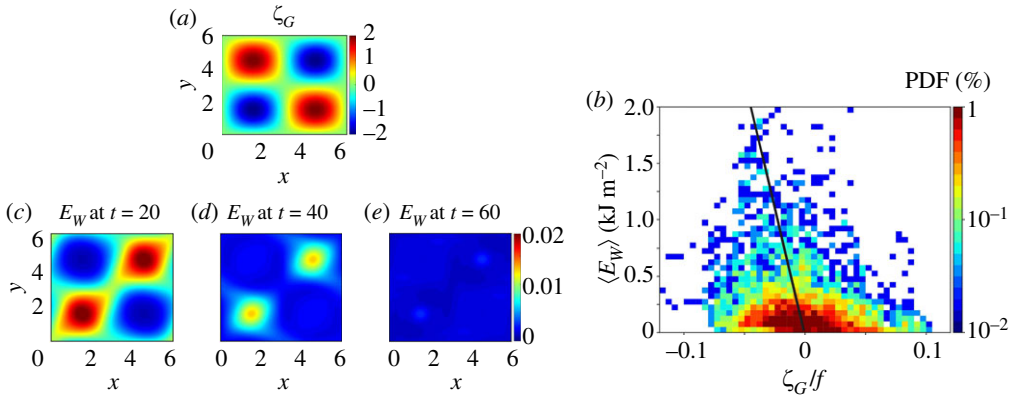

Figure 4. Panels (*a*), (*c–e*) are adapted from [128] with permission, copyright 2017 Cambridge University Press, and illustrate the vertical propagation of NIW energy, E_w, for a given geostrophic flow. Panel (*a*) shows the non-dimensional geostrophic vorticity field (vorticity normalized by the rotation rate f), ζ_G. Panels (*c*), (*d*) and (*e*) shows the horizontal structure of NIW kinetic energy (normalized by an estimate for the balanced kinetic energy) 100 m below the top surface of the ocean at three different times. Notice how NIW kinetic energy is high in anticyclonic or negative vorticity regions and low in cyclonic or positive vorticity regions. The vertical propagation of NIWs reduces the intensity of waves on the plane 100 m below the top surface as time progresses. Panel (*b*) shows the joint PDF of vertically integrated NIW kinetic energy, $\langle E_w \rangle$, and geostrophic vorticity, ζ_G, based on observational data discussed in [132]. A linear regression plot between wave energy and vorticity, shown by the black line, reveals that NIWs are correlated with negative geostrophic vorticity regions.

discussed in [132] is given in figure 4*b*, where the joint PDF of vertically integrated NIW kinetic energy and gesotrophic vorticity field is shown. The linear regression between NIW kinetic energy and geostrophic vorticity is denoted by the black line in the figure. Notice how the black line has a negative slope, indicating higher correlation between NIWs and anticyclonic vorticity regions. The trapping of NIWs in anticyclonic vorticity regions is therefore a feature that is consistently observed in the ocean, and is captured by numerical integration of the primitive equations and the YBJ equation.

In addition to using the YBJ equation for capturing the wave field's modulation by the eddy field, the asymptotic wave equation can also be coupled with balanced flow evolution equations. Recall that the slow balanced equations (2.6) and (2.7) were not closed due to the absence of wave equations that evolve on long time scales. The amplitude equation (3.1) is a wave equation that evolves on long time scales and can be coupled with the balanced flow equations given in (2.6) and (2.7) to form closed wave-balance asymptotic models. This was recognized by [137] who coupled the YBJ equation with the slow balanced equation in the SW regime, (2.7). An interesting feature of the YBJ amplitude equation (3.1) is that it conserves wave kinetic energy. However, wave potential energy is not conserved and can be exchanged with balanced flows. The coupled system of [137] therefore conserves total energy and wave kinetic energy separately. Notably, wave potential energy is proportional to $|\nabla A|^2$. Consider an NIW wave field that has extremely large horizontal scales initially. The interaction with the balanced eddy field would imprint the eddy spatial scales on the wave field as seen in figure 4 and therefore gradients of the wave field, i.e. $|\nabla A|$, would increase once the wave-balance interaction begins. This means that wave potential energy increases as time progresses. Since wave kinetic energy does not change and since the sum of wave potential energy and balanced energy is conserved in the asymptotic model, the increase in wave potential energy is accompanied by a decrease in the balanced energy. Using the coupled model, Xie & Vanneste [137] pointed this out using numerical simulations of the asymptotic model, followed by similar conclusions by Wagner & Young [138] and Rocha *et al*. [139] using reduced asymptotic models of NIW-balanced flow interactions. The results

from asymptotic models predict about 5–20% loss of balanced energy within a few tens of eddy turn over time scales as a result of wave-balance interactions.

The above mentioned mechanism is a means by which wave potential energy increases by direct extraction of balanced energy. On exploring this interaction problem with primitive equation models constrained to a barotropic and a single baroclinic mode, Thomas & Arun [140] found that NIW potential energy did increase at the expense of balanced energy, as predicted by the asymptotic model based results of [137–139]. However, contrary to the prediction of the YBJ model, NIW kinetic energy was seen to decrease with time. The drop in NIW kinetic energy was more than the increase in NIW potential energy, resulting in waves transferring energy to the balanced flow. Interestingly, almost two decades back, an idealized study using a reduced model by Gertz & Straub [141] also found that balanced flow was energized by NIWs in low balance energy regions. The phenomenon of NIWs transferring energy to balanced flow was further confirmed by three-dimensional Boussinesq equation simulations of Thomas & Daniel [142]. The numerical integration of the full set of equations, rather than the asymptotic models alone, revealed that NIWs kinetic energy can in general change as much as the changes in NIW potential energy, making NIW-balanced flow energy transfers much more non-trivial than those predicted by asymptotic models.

To get a firm grip on NIW-balanced flow energy exchanges in different regimes, Thomas & Daniel [142] integrated the Boussinesq equations with different balance-to-wave energy ratios. In general, NIW energy levels are highly variable in the world's oceans [97,117,119,143–146]. Consequently, although the asymptotic models mentioned above were derived in a SW regime where wave energy is much stronger than balanced energy, such regimes are observed in the ocean only when strong winds excite NIWs in oceanic regions where pre-existing balanced energy is weak (see for example such an event described in [147]). Given the variability in NIWs energy levels in the world's oceans, the work of Thomas & Daniel [142] was primarily aimed at exploring whether NIWs could act as an energy sink for balanced flow in the different regimes. The main results of Thomas & Daniel [142] are summarized in figure 5. Figure 5*a* shows the geostrophic vorticity field in the CW regime. In this regime balanced flow evolves as in the QG regime. Furthermore, NIWs directly extract energy from the balanced flow in the CW regime, as shown in figure 5*c*. Balanced flow loses about 10–20% energy over a few hundred eddy turn over time scales in this regime. By contrast, the SW regime is characterized by NIWs disrupting large-scale coherent vortices, as seen in figure 5*b*. Additionally, NIWs transfer energy to the balanced flow in the SW regime, as shown in figure 5*d*, with the gain in balanced energy being 5–15% over a few tens of eddy turn over timescales. Thomas & Daniel [142] found that while wave potential energy increased, as expected based on the asymptotic model based results, wave kinetic energy decreased in both CW and SW regimes. Consequently, the magnitude of change in wave kinetic energy was the key factor deciding the direction of NIW-balanced flow energy transfers. In the CW regime, wave potential energy increased more than the decrease in wave kinetic energy, resulting in waves extracting energy from the balanced flow. On the other hand, wave kinetic energy dropped more than the increase in wave potential energy in the SW regime, resulting in waves losing energy to balanced flow in the SW regime.

In addition to energy exchange with balanced flow, much of the NIW dynamics are correlated to the strength of the balanced flow. Faster vertical propagation of waves and a stronger forward energy flux of waves resulting in higher small-scale dissipation of wave energy is observed in the CW regime when compared to the SW regime. Numerical results therefore pinpoint that waves' forward flux, small-scale dissipation, and vertical propagation are proportional to the strength of the balanced flow. This finding has also been confirmed from observational results. For instance, Whalen *et al*. [148] found that small-scale dissipation of wave energy can increase by an order of magnitude in the presence of an energetic eddy field when compared with a case where the eddy field is energetically weak.

Interestingly, the SW regime of NIW-balanced flow interaction has similar features to the SW regime of IT-balance exchanges. Both wave fields transfer energy to the balanced flow and affect the formation of large-scale balanced coherent vortices. By contrast, while the CW regime for

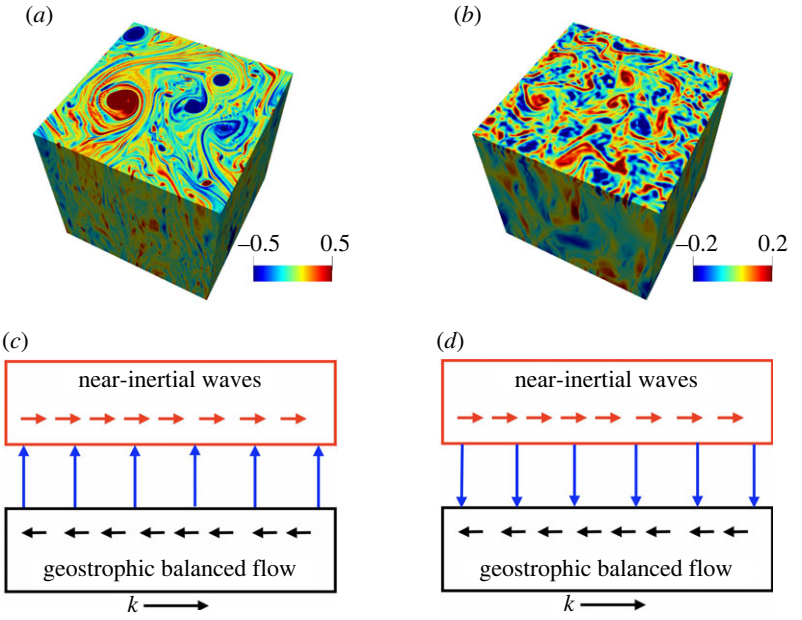

Figure 5. Figure adapted from [142] with permission, copyright 2020 Cambridge University Press, shows results of NIWs on balanced flow in different regimes. The left column, i.e. panels (*a*) and (*c*), shows the non-dimensional geostrophic vorticity field (vorticity normalized by the rotation rate *f*) and the wave-balance energy transfer pathways, respectively, in the CW regime. The right column shows the same quantities in the SW regime. In the CW regime, as shown by the blue vertical arrows in panel (*c*), NIWs directly extract balanced energy. Despite this energy transfer, the balanced flow forms large-scale coherent vortices as in the QG regime, this feature being seen in panel (*a*). By contrast, NIWs transfer energy to the balanced flow in the SW regime (blue vertical arrows in panel (*d*)) and destroys large-scale coherent vortices (see panel (*b*)), resulting in the formation of fine-scale vortical structures.

IT-balanced flow interactions is characterized by no energy transfer, the CW regime for NIW-balance exchange is composed of direct extraction of balanced energy by waves; the energy transfer accompanying the reduction in spatial scales of NIWs during wave-eddy interactions. As mentioned earlier, the SW regime for NIWs is relevant only when extremely strong winds excite waves in regions where balanced flow is relatively weak. Consequently, the CW regime is probably more relevant in the real ocean, and we anticipate NIWs to extract energy from balanced flow in this regime. Confirming this, realistic ocean model simulations set in eddy permitting and eddy resolving configurations have found that NIWs can form an energy sink for the slow balanced flow. For instance, Taylor & Straub [149,150] using realistic ocean models found that about 5–30% of power input into slow balanced flow can be drained by fast NIWs, this estimate being in agreement with results from idealized studies mentioned above.

(c) Internal wave continuum

The ocean harbours a broad spectrum of internal gravity waves in the frequency band $f \ll \omega \ll N$, identified as the internal wave continuum (IWC). Based on oceanic datasets, Garrett and Munk constructed an empirical spectral description of the IWC [151–153]. Although direct forcing of waves in the ocean is primarily via winds and tide, nonlinear wave-wave interactions and scattering by the vortical mode is thought to be responsible for the generation of this broad spectrum of waves from low-frequency waves [154–166]. Multiple oceanic observations exploring the wave spectrum have revealed that the IWC is highly variable with the strength of waves' spectrum being sensitive to the geographical location and season [167–169]. Despite the

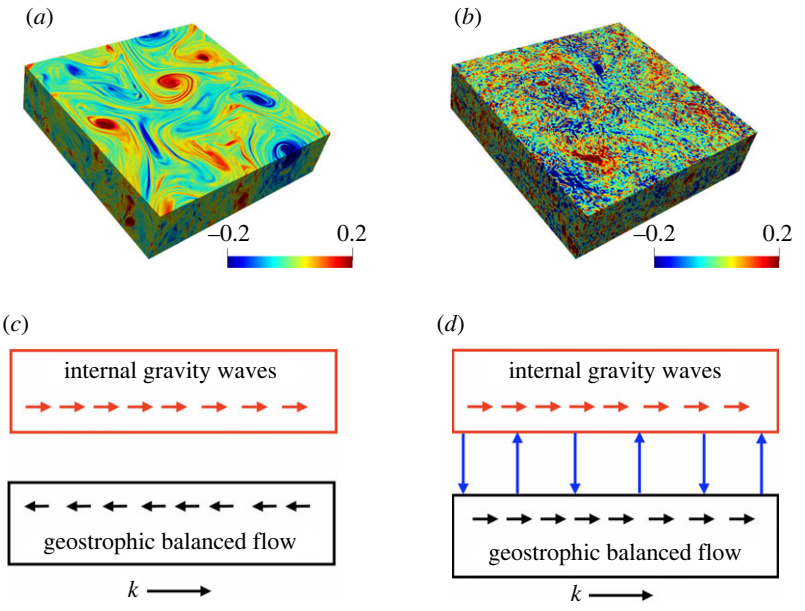

Figure 6. Figure adapted from [172] with permission, copyright 2021 Cambridge University Press, shows results of an internal wave continuum on balanced flow in different regimes. Left column, i.e. panels (*a*) and (*c*) shows the non-dimensional geostrophic vorticity field (vorticity normalized by the rotation rate *f*) and energy transfers, respectively, in the CW regime. The right column shows similar quantities for the SW regime. In the CW regime balanced flow evolves unaffected by waves with no noticeable wave-balance energy exchanges. By contrast, the SW regime is characterized by two-way wave-balance energy exchanges (see vertical blue arrows in panel (*d*)) along with waves inhibiting the formation of large-scale coherent vortices in the balanced flow (see panel (*b*)). The SW regime also features a forward flux of geostrophic energy, as indicated by the rightward pointing black arrows in panel (*d*), this being contrary to the inverse flux of balanced flow seen in the CW regime in panel (*c*).

variability, the IWC connects the high-energy low-frequency waves to high-frequency low-energy waves that break and mix the oceans. Owing to their low resolutions, ocean models of the past have had difficulties in resolving the IWC. However, recent high-resolution simulations of large-scale ocean models with realistic wave forcing have started resolving the low-frequency part of this broad waves' spectrum [39,170,171].

Oceanic datasets from *in situ* and satellite observations and global scale simulations of the past two decades have revealed that the internal waves' spectra can overtake the balanced vortical spectrum at a certain scale identified as the transition scale. An example for this can be seen in figure 1*d*: notice that the red high-frequency spectrum overtakes the blue low-frequency spectrum roughly around 100 km scales. Similar results can be seen in [35,38,40–42,171]. This transition scale, where fast waves overtake slow balanced energy spectrum, is highly variable in the world's oceans and depends on the geographical location and the season. Inspired by these findings, Thomas & Daniel [172] used numerical integration of the Boussinesq equations to explore the interactions between a broad spectrum of waves, forming an IWC, and a balanced flow and the main highlight of the study is given in figure 6.

Figure 6*a* shows the geostrophic vorticity field of the CW regime. Similar to the QG regime, the balanced flow transfers energy upscale and evolves unaffected by waves in this regime. Waves on the other hand flux their energy downscale and dissipate. As seen in the energy transfer pathways given in figure 6*c*, waves and balanced flow remain decoupled in the CW regime. Although wave and balanced flow do not exchange energy, waves' forward flux and small-scale dissipation can be an order of magnitude higher in this regime, when compared with a case where waves

interact among themselves with no background balanced flow. The SW regime is on the other hand composed of much more intense wave-balance exchanges as indicated in figure 6*b*,*d*. The broad spectrum of internal waves, as they flux energy downscale, insinuates a forward flux of balanced flow as well. Consequently, the geostrophic balanced flow exhibits a forward energy flux in the SW regime, this inhibiting the formation of coherent balanced vortices, as seen in figure 6*b*. Furthermore, waves and balanced flow continuously exchange energy in this regime, with no preferred energy transfer direction. This is in contrast to the SW regime of ITs and NIWs, where the net energy transfer is unidirectionally from waves to the balanced flow. Additionally, waves' forward flux and small-scale dissipation is less here relative to the CW regime due to the weaker balanced flow. In summary, among the different wave fields and the parameter regimes examined in the small Rossby number regime, the SW regime for a broad spectrum of waves is the only case where waves can trigger a forward flux and facilitate significant dissipation of balanced energy. Thomas & Daniel [172] find that about 30–50% of balanced energy can be dissipated at small scales over a few hundred eddy turn over time scales as a result of the balanced flow interacting with the IWC in the SW regime.

4. Challenges in quantifying wave-balance energy exchanges from observational datasets

As is clear from the previous section, a wealth of information on wave-balance energy exchanges has been obtained from numerical integration of the governing equations in different parameter regimes. By contrast, oceanic observations are spatially sparse, making it extremely challenging to examine wave-balance exchanges from observational datasets. Much of the flow fields' data is obtained from satellite altimeter and ocean moorings [173]. Satellite altimeter datasets are confined to the upper ocean and provide limited information on vertical flow features while moorings provide time series of fields at specific locations. Naturally, multiple assumptions and interpolations are required to deduce wave-balance energy exchanges from realistic datasets.

To understand the difficulties in examining wave-balance energy transfers from observations, it is worthwhile looking at the relevant equations. Time averaging the governing equations gives us the slow evolution equation written in abstract form as

$$\frac{\partial \overline{\boldsymbol{U}}}{\partial t} + \overline{\boldsymbol{N}}(\boldsymbol{U}', \overline{\boldsymbol{U}}) = 0. \tag{4.1}$$

Above $\overline{\boldsymbol{U}}$ denotes the slow field vector, which includes the velocity components and the buoyancy, while $\boldsymbol{U}'$ denotes the fast field, and $\overline{\boldsymbol{N}}$ indicates the nonlinear terms of the slow equation. Taking the dot product of (4.1) with $\overline{\boldsymbol{U}}$ gives us the slow field's energy equation

$$\frac{\partial}{\partial t}\left(\frac{\overline{\boldsymbol{U}}^2}{2}\right) = -\overline{\boldsymbol{U}} \cdot \overline{\boldsymbol{N}}(\boldsymbol{U}', \overline{\boldsymbol{U}}) = \mathrm{S}. \tag{4.2}$$

The above equation points out that slow balanced energy changes point-wise, i.e. at every point (x, y, z) in the domain, due to wave-balance interaction terms on the right-hand side of the equation. The right-hand side of the above equation, S, gives the rate of energy transfer between balanced flow and waves. Integrating (4.2) over the entire domain gives us the net energy transfer as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \frac{1}{2}\overline{\boldsymbol{U}}^2 d\boldsymbol{X} = -\int_D \overline{\boldsymbol{U}} \cdot \overline{\boldsymbol{N}}(\boldsymbol{U}', \overline{\boldsymbol{U}})\, \mathrm{d}\boldsymbol{X} = \mathrm{T}, \tag{4.3}$$

where the integration above is over the three-dimensional space with $\boldsymbol{X} = (x, y, z)$. Computing the term T above would inform us whether the balanced energy increases or decreases due to wave interactions. While above integral can be computed straightforwardly in numerical simulations where wave and balanced fields are readily available across a large number of grid points located in the domain, oceanic observations are extremely sparse, making it next to impossible to accurately compute integrals such as above. The flow fields are often available only at a

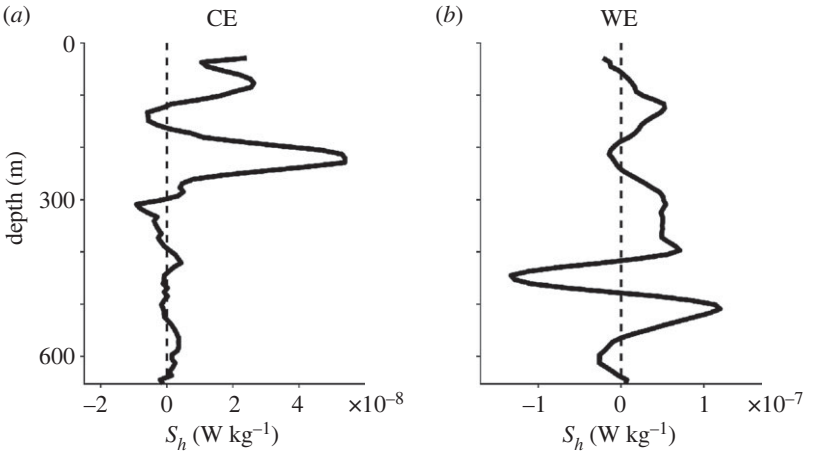

Figure 7. Figure shows S_h, the horizontal part of the energy transfer term S in (4.2). Panels (*a*) and (*b*) show the transfer term in a cold core eddy (CE) and warm core eddy (WE), respectively, plotted as a function of depth. Dashed vertical lines indicate zero transfer lines. Notice that the transfer term fluctuates about the zero line. Details of this dataset are discussed in [176].

sparse set of points associated with locations where oceanic moorings are based in or satellites can access. A practical solution to above challenge, often implemented by oceanographers, is to use time series data from a limited number of points, compute the interaction terms and then average over long periods of time, with the hope that a long-term average would approximate a spatial average. Typically, in three-dimensional homogeneous and isotropic turbulence it is common to use ergodic theory to swap ensemble averaging with time averaging (see discussions in [174]). However, in oceanic flows, especially at 10–100 km scales where large-scale coherent vortices, well defined fronts, and wave fields propagating in specific directions interact in an inhomogeneous and anisotropic environment, it is in general difficult to justify that time averaging can approximate spatial averaging.

Given the above challenges, quite often it is possible only to compute the components of the energy transfer term in (4.2) at specific locations, after making a set of simplifying assumptions to manipulate the sparse datasets obtained from *in situ* measurements. Assuming a spatial scale separation between small-scale waves and large-scale balanced flows, Muller [57] derived equations that capture interactions between waves and slow mean flows that are in geostrophic balance. The equations so derived are further simplifications of (4.2), where the interaction terms on the right-hand side of (4.2) were further decomposed into horizontal (S_h) and vertical (S_v) exchange terms. This interaction equation was used by Polzin [58] and Brown & Owens [175] to examine wave-balance energy exchanges using *in situ* datasets.

Recently Qi *et al*. [176] computed the energy transfer terms described above for NIWs interacting with balanced flows. Figure 7 shows the horizontal part of the energy transfer term, S_h, as a function of depth inside a cold and warm eddy. Notice that the term fluctuates around zero with depth, indicating that at a specific location waves could act as an energy source or energy sink for balanced flows. On a similar note, using simplified energy transfer equations used by [57], Polzin [58] and Jing [177] found that NIWs extracted energy from slow balanced flows in a certain region, while Cusack *et al.* [178] found that NIWs fed slow balanced flows in a specific region. As mentioned above, given the sparsity of observational datasets, only point-wise statements on wave balance transfers can be obtained from realistic datasets. Computing the net energy transfer term T in (4.3) is infeasible at present due to limited datasets.

As hinted above, a spatial scale separation assumption for wave fields is commonly used for wave field calculations by oceanographers. The popularly implemented ray tracing or WKBJ method [179,180] assumes a scale separation such that a small spatial scale wave fluctuating

on a fast time scale propagates through a large-scale slow evolving mean flow. Despite the questionable nature of this assumption when applied to realistic wave fields in the ocean, the technique often results in findings that are qualitatively useful (see for example discussions in [121,181]). The ray tracing approach was used by Rainville & Pinkel [102] and Chavanne *et al*. [104] for low mode ITs propagating through balanced flows. Interestingly, both these studies found that low mode tides exchanged energy with the background eddy field. However, since a part of the energy loss of low mode tides could be to higher vertical mode waves as well, there is an inevitable uncertainty in precise estimates on energy transfers between waves and eddies.

A reader who parsed through the details presented in the previous two sections focusing on theoretical and numerical studies on wave-balance exchanges is likely to be surprised by the limited results we have based on oceanic observations, summarized in the above paragraphs. At present there is no conclusive observational evidence on net energy transfer between waves and balanced flows since this requires oceanic data with fine resolution in time and space, the latter needing to be obtained from a large number of moorings arranged at fine resolution in an oceanic region. The lack of such high-resolution data prevents any sort of comparison between the results we have from theory and numerics and observations. Consequently, as mentioned in the previous sections, present observational results are primarily useful to point out the relevant parameter regimes in the ocean, such as providing us with the relative energy levels of balance and waves, the dominant wave field, local Rossby number, etc. [44,117,144–146,182]. Furthermore, some features of wave-balance interactions, such as trapping of NIWs in anticyclonic mesoscale eddies [132,134–136] and scattering and loss of coherence of ITs by mesoscale eddies [98,99,102,104,106,107] can be obtained from observational datasets. Theoretical and numerical investigations can choose relevant regimes based on observations and can cross check their results qualitatively with some selected features observed in the ocean. Beyond these, a direct comparison between theory/numerics and observations, especially the predictions of net wave-balance energy exchanges is far fetched at present.

5. Comparison with interactions in $Ro \sim O(1)$ regime

Having focused on interactions in the $Ro \ll 1$ limit so far, in this section we will briefly look at $Ro \sim O(1)$ interactions. The last two decades have seen an explosive growth in research focusing on submesoscale flows with $O(1)$ Rossby numbers based on *in situ* and satellite altimeter datasets and ocean model simulations. Energetic submesoscale flows with $O(1)$ Rossby numbers are typically seen in connection with ocean fronts in high eddy energy regions such as Gulf stream and Kuroshio Current and in weakly sheared upper ocean flows [81,82,162,183–195]. Given that the QG equation (2.5) is derived assuming $Ro \ll 1$, it is hardly surprising that the characteristic features of the balanced flow, such as the inverse energy flux and large-scale coherent vortex formation, are affected as $Ro \to 1$. Internal gravity waves can interact with such $O(1)$ Rossby number flows and can result in rapid energy loss of both waves and eddy structures. Turbulent dynamics of $O(1)$ Rossby number flows have been extensively discussed in multiple review papers in the past [196–198]. Therefore, since the primary focus of this review is interactions in the $Ro \ll 1$ regime, here we will only look at some selected aspects of interactions in $O(1)$ Rossby number flows.

In the $Ro \sim O(1)$ regime, a significant amount of work has been dedicated to NIWs interacting with ocean fronts. Such interactions are reviewed in detail in [199] and the interaction can lead to instabilities and turbulence, resulting in dissipation of wave and the front. Interestingly, energy can go from wave to the front or vice versa in different scenarios. Barkan *et al*. [162] discuss a study in the $O(1)$ Rossby regime where NIWs can directly extract energy from low-frequency eddies and also facilitate an enhanced dissipation of around 30% of the power input into the low-frequency eddy field. Barkan *et al*. [200] further extended these results based on realistically forced regional ocean models to conclude that in $O(1)$ Rossby number regime wave-induced effects can result in depletion of low-frequency mesoscale balanced energy. Similar wave-eddy interaction experiments with idealized ocean models was undertaken by [201–203] leading to the

conclusion that waves can transfer energy to or extract energy from coherent eddies and fronts in the $O(1)$ Rossby number regime. Despite the lack of a specific energy transfer direction, energetic interactions in the $O(1)$ Rossby number regime generically result in dissipation of the wave and the coherent structure.

When compared to the $Ro \ll 1$ regime interaction results described in the previous sections, two key features stand out for interactions in the $O(1)$ Rossby number regime. First, fast–slow exchanges are much more intense and rapid in the $O(1)$ Rossby number regime, leading to the dissipation of fast and slow fields on a relatively short time scale, in comparison to small Rossby number regime interactions. To appreciate the second feature, recall that for $Ro \ll 1$, the balanced flow is severely hampered in cases where wave energy levels are much higher than balanced flow, identified as the SW regime in the previous sections. This is the regime where waves can prevent the formation of large-scale coherent vortices and can enhance small-scale dissipation of balanced energy. By contrast, as $Ro \to 1$, even a small amount of wave energy can lead to significant damage to balanced coherent structures and can dissipate them efficiently; this behaviour being observed in idealized studies and complex ocean model simulation results [186,194]. The significant drop in the amount of unbalanced or wave energy needed to dissipate balanced coherent structures as Rossby number increases is a striking feature that stands out as energetic interactions are contrasted for different Rossby number flows.

6. Summary and discussion

The oceanographic community in the past have considered O(10–100) km scales to be dominated by the balanced flow. However, the last two decades of satellite altimeter datasets, *in situ* measurements, and realistically forced ocean model outputs have challenged this age old paradigm. The datasets that have piled up from observations and global scale simulations have revealed the prominence of energetic internal gravity waves at O(10–100) km scales. At present, it is understood that waves and balanced flow can share similar spatial scales and in certain oceanographic regions internal waves can be significantly stronger than balanced flows.

The above-mentioned paradigm shift fueled a broad set of investigations focusing on the effect of internal waves on balanced flow. In the absence of waves, the balanced flow follows the QG turbulence phenomenology with an upscale transfer of energy. The role of waves in modifying QG turbulence phenomenology and the possibility of waves acting as an energy sink for balanced flow have been key questions that inspired in-depth wave-balance investigations in the oceanographic setting. This paper was written to provide readers an overview of the main findings in this direction, amalgamating major results obtained via theory, numerical simulations, and observations. For convenient reference, we have summarized the main results on wave-balance exchanges in the small Rossby number regime discussed in the previous sections in table 1.

(a) Global budget for energy sinks of balanced flow: role of internal waves

As is clear from discussions so far, internal waves lead to significant changes in balanced flow dynamics in the SW regime. However, this regime is relevant in oceanic regions where wave energy is significantly higher than balanced energy. As a result, the contribution of such regions to the global balanced energy would be relatively small. The CW regime on the other hand, where wave and balanced energy are comparable in strength, is applicable to much of the oceanic regions across the globe; this being particularly the case for NIWs: see for instance global maps of waves and balanced flow discussed in [44]. Furthermore, as pointed out earlier in §3b, NIWs can directly extract energy from balanced eddies in the CW regime and forced-dissipative numerical simulations in realistic configurations reveal that the energy extraction rate could be 5–30% of the power input into the low-frequency eddy field. If we conservatively set the global average extraction to be 10%, we have 0.08 TW out of the 0.8 TW power going to the geostrophic balanced flow being extracted by NIWs globally. Recall that this number corresponds to low Rossby

Table 1. Summary of wave-balance exchanges in the small Rossby limit. Interactions between balanced flow and low-mode internal tides (ITs), near-inertial waves (NIWs), and the internal wave continuum (IWC) are listed above.

no.	wave field	regime	effect of waves on balanced flow dynamics	effect of balanced flow on waves' dynamics
1	IT	CW	Balanced flow does not exchange energy with waves.	Waves are scattered and inhomogenized by the balanced flow. Less than 5% of first mode tidal energy is lost to higher modes over thousands of inertial periods.
2	IT	SW	Waves transfer energy to the balanced flow. As a result, balanced energy can increase by a factor of 2–10 over a few hundred eddy turn over time scales.	Same as above, although the scattering of waves and their inhomogenization decreases since balanced energy here is much less than in the CW regime.
3	NIW	CW	Balanced flow can lose energy to waves. Waves can extract 10–20% of balanced energy over a few hundred eddy turn over time scales, depending on the specific magnitudes of wave and balanced energy.	Balanced flow accelerates the vertical propagation and enhances forward flux and small-scale dissipation of waves. When compared to a case with no eddy field, waves' small-scale dissipation can increase by as much as a factor of 10 in the presence of an energetic eddy field.
4	NIW	SW	Balanced flow can gain energy from waves. Energy gain of balanced flow can vary from 5–15% over a few tens of eddy turn over time scales depending on the specific magnitudes of wave and balanced energy.	Same as above, although waves' vertical propagation, forward flux and small-scale dissipation is substantially weaker due to the eddy field being weaker here relative to the above regime.
5	IWC	CW	Balanced flow does not exchange energy with waves.	Waves' forward flux and small-scale dissipation is enhanced by the balanced flow. Compared to a case with no background eddy field, waves' forward flux and small-scale dissipation can increase by as much as a factor of 10 in the presence of an energetic balanced flow.
6	IWC	SW	Balanced flow exhibits a forward energy flux and small-scale dissipation. Balanced energy loss can vary from 30–50% over a few hundred eddy turn over time scales in this regime.	Same as above, although waves' forward flux and small-scale dissipation drops substantially here compared to the above case due to the lower balanced energy level.

number energy extraction alone. If we include O(1) Rossby number interactions detailed in §5 as well, we may increase the number to 0.1 TW of energy loss for balanced flow. This estimate of 0.1 TW might be thought of as a lower bound for internal wave-induced energy sink of balanced flow.

Based on estimates of other known energy sinks of balanced flow, we will now attempt to close the energy budget of balanced flow. If we attribute about 0.2 TW of balanced energy loss at western boundaries [50,52,53,204,205], 0.3 TW loss of balanced energy to bottom boundary layer drag [49,51], 0.15 TW to eddy energy loss due to interaction with rough topographic features and associated dissipation [54,55,206,207], and 0.05 TW to atmospheric winds damping mesoscale eddies [208–210], we have the net loss of balanced flow to be 0.1 TW (internal waves) + 0.2 TW (western boundaries) + 0.3 TW (bottom boundary layer drag) + 0.15 TW (rough topography) + 0.05 TW (atmospheric winds) = 0.8 TW. The different energy sinks can therefore add up to 0.8 TW, the net power input into the geostrophic energy field. Of course, the numbers we put in above are again approximate since there are significant uncertainties associated with them as noted in the references above. Nevertheless, we added up the approximate contribution of each energy sink above to point out that the different mechanisms can together dissipate and thereby equilibrate the geostrophic balanced flow in the ocean. Furthermore, the approximate estimates above indicate that internal waves can form a non-negligible global sink for balanced energy.

(b) Future

We will conclude this paper by pointing out some unresolved directions related to wave-balance exchanges that needs to be undertaken by future studies.

— Parametrization of wave-induced effects on balanced flow: given that waves can affect the balanced eddy field in the ocean, it is imperative that wave-induced effects are parametrized in large-scale ocean models. Parametrization of mesoscale eddy field is actively carried out in ocean models [211–215] and similar initiatives need to be undertaken to account for the effect of waves on mesoscales and submesoscale flows in the ocean. From table 1, it follows that depending on the relative energy level of wave and balanced flow and the kind of wave field, waves can extract energy from balanced flow or transfer energy to balanced flow. As a result, any attempt to parametrize the effect of waves in large-scale ocean or climate scale models must take into account the details of the wave field. In other words, *one size does not fit all*: parametrizations will need to adapt dynamically depending on the dominant wave field and its energy level.
— Waves and tracer dispersion: as mentioned in the Introduction, identifying the role of waves in modifying QG turbulence and the features of the balanced flow has been the driving force for much of the work reviewed in this paper. Waves affecting and modifying balanced flow dynamics leads to interesting secondary consequences. For example, studies investigating lateral dispersion of tracers in the ocean find that tracer stirring is more efficient than that expected based on QG turbulence scaling [216–222]. Although observational datasets, ocean model outputs and idealized studies have pointed out that waves can enhance lateral dispersion of tracers [223–226], the detailed role of different kinds of internal waves across broad parameter regimes in modifying lateral dispersion of tracers and turbulent tracer diffusivity remain unclear. An in-depth study in this direction would benefit the oceanographic community in developing diffusivity parametrizations at mesoscales and submesoscales for large-scale ocean models.
— Waves and balanced flow instabilities: given that internal waves and balanced flow coexist at similar spatial scales in the ocean, it would be useful to understand how the instabilities of the geostrophic flow get modified by the presence of waves. Baroclinic instability, for example, is usually studied in the context of the QG equation that discards waves. However, as (2.6) and (2.7) shows, internal waves can modify the QG equation and could therefore modify the balanced flow instabilities. The catalytic effects of waves on balanced flow have been noted in multiple idealized numerical simulations [227–230]. Although a broad set of geophysical flow instabilities have been identified, including

those in the limit of O(1) Rossby numbers [231–235], the role of internal waves in balanced flow instabilities remains unclear at present.

— Balance in $Ro \sim O(1)$ regime: in comparison to the matured level of results from the low Rossby number-based studies, much remains unknown in the $O(1)$ Rossby number regime. In this regard, one of the challenges with $O(1)$ Rossby number regime investigations is the lack of an unambiguous criterion for identifying balanced flow. A firm footing on balanced flow definitions that would hold equally well at low and O(1) Rossby number regimes would significantly benefit studies exploring geophysical turbulence in mesoscale and submesoscale regimes. Consequently, dedicated studies that underpin precise definitions of balanced flow that would hold in $O(1)$ Rossby number regime will be highly valuable for future investigations.

— Connecting different flow decompositions: multiple fast–slow and wave-balance decompositions are in use at present. These include the linear wave-balance decomposition that is based on the linear equations (2.3) (e.g. [172]), Eulerian fast–slow decomposition based on time filtering Eulerian fields (e.g. [162]), Lagrangian fast–slow decomposition based on time filtering of Lagrangian fields (e.g. [202]), and approximate asymptotic decompositions (e.g. [71]). The usage of these different kinds of flow decompositions makes it difficult to compare and contrast results from different studies. To highlight an example issue, although the Eulerian fast–slow decomposition would provide us with a slow field, the slow field is not necessarily in geostrophic balance. As a result, it is not straightforward to compare the results from Eulerian fast–slow decomposition with the linear wave-balance decomposition. Consequently, it will be beneficial to have a detailed understanding of how to connect the results of studies that use different kinds of flow decompositions.

In summary, the past decade of theoretical and numerical investigations, inspired by observations and high-resolution global scale ocean model outputs, have unraveled significant amounts of details on turbulent transfers between internal waves and balanced flows. These studies have clearly indicated that internal waves can modify QG turbulent dynamics in different parameter regimes and assert the need for introducing wave parametrizations in large-scale ocean models that are far from resolving internal waves. Development of wave-parametrizations for oceanic general circulation modes, identifying the secondary role of waves on flows, such as their effect on lateral dispersion of tracers, introducing new unambiguous definitions of balance across parameter regimes, and making closer connections between *in situ* and satellite measurements based results and theory/numerics are challenges that hopefully will be resolved in the near future.

Data accessibility. This article has no additional data.

Authors' contributions. J.T.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, writing—review and editing.

Conflict of interest declaration. I declare I have no competing interests.

Funding. This study was financially supported by the Science and Engineering Research Board (SERB) of India through the project SRG/2022/001071.

Acknowledgements. This paper primarily grew out of talks I gave at the *Geophysical Fluid Dynamics Summer Seminar Series* in 2021 at the Woods Hole Oceanographic Institution and the program *Multiscale Wave-Turbulence Dynamics in the Atmosphere and Ocean* in 2022 at Mathematisches Forschungsinstitut Oberwolfach. I am thankful to comments and suggestions on the manuscript by Carl Wunsch, Eric D'Asaro, Kurt Polzin, Shafer Smith, Eric Kunze, Jesse Cusack, Bo Qiu, and Jonas Nycander. I am also grateful to Carl Wunsch for providing data to genrate figure 1*a*, Ren-Chieh Lien for providing figure 1*b*, Edward Zaron for providing figure 2*e*, Clément Vic for providing figure 4*b*, and Zhaohui Chen for providing figure 7. Anugraha A. from the outreach team at ICTS-TIFR is specially thanked for helping in editing and arranging the figures in this paper. Finally, I thank Chris Garrett, who served as the editor for the paper, and anonymous reviewers for comments and suggestions that improved the presentation of material in this paper.

Appendix A. Boussinesq equations and non-dimensionalization

The Boussinesq equations on the f-plane in the hydrostatic limit, and in the absence of viscous and diffusive processes, are given by

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + f\hat{\boldsymbol{z}} \times \boldsymbol{v} + \nabla p = 0, \tag{A 1a}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} + N^2 w = 0, \tag{A 1b}$$

$$\frac{\partial p}{\partial z} = b, \tag{A 1c}$$

and

$$\nabla \cdot \boldsymbol{v} + \frac{\partial w}{\partial z} = 0, \tag{A 1d}$$

where $\boldsymbol{v} = u\hat{\boldsymbol{x}} + v\hat{\boldsymbol{y}}$ is the two-dimensional velocity vector, w is the vertical velocity, p is the pressure is the pressure, b is the buoyancy, $N(z)$ is the buoyancy frequency, f is the rotation rate, $\nabla = \hat{\boldsymbol{x}}\partial/\partial x + \hat{\boldsymbol{y}}\partial/\partial y$, and $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{v} \cdot \nabla + w\partial/\partial z$.

We non-dimensionalize the governing equations above as

$$\left.\begin{aligned} &t \to \frac{t}{f}, \quad \boldsymbol{x} \to L\boldsymbol{x}, \ \ z \to Hz, \ N \to N_0 N, \ \boldsymbol{v} \to U\boldsymbol{v} \\ \text{and} \quad &w \to (HU/L)w, p \to (fUL)p, \ b \to \left(\frac{fUL}{H}\right) b \end{aligned}\right\} \tag{A 2}$$

In the above scaling, the inertial frequency f was used to scale time while the vertical length scale was chosen as the depth of the domain, H. N_0 is a reference buoyancy frequency used to scale the dimensional buoyancy frequency $N(z)$. We chose the horizontal length scale to be the deformation scale, $L = N_0 H/f$. Horizontal velocity was scaled using an arbitrary velocity scale, U, and the scale for vertical velocity, HU/L, was chosen to satisfy continuity equation (A 1d). The scale for pressure was chosen such that horizontal pressure gradient (∇p) was of the same order as the Coriolis term ($\boldsymbol{f} \times \boldsymbol{v}$). Finally, the scale for buoyancy was obtained by setting the vertical pressure gradient $\partial p/\partial z$ to be of the same order as buoyancy, b. We further introduce the slow time $T = \epsilon t$ so that the time derivative term becomes $\partial/\partial t \to \partial/\partial t + \epsilon\partial/\partial T$. Using the above in (A 1) gives us the non-dimensional equations (2.1).

Equations (A 1) possess an exact material invariant: potential vorticity (see § 4.7.2 in [7])

$$\frac{\mathrm{D}}{\mathrm{D}t}[(\boldsymbol{\omega}_{3D} + f\hat{\boldsymbol{z}}) \cdot (\nabla_{3D} b + N^2\hat{\boldsymbol{z}})] = 0. \tag{A 3}$$

Above $\omega_{3D} = \hat{\boldsymbol{z}} \times \partial\boldsymbol{v}/\partial z + \zeta\hat{\boldsymbol{z}}$, with $\zeta = \partial v/\partial x - \partial u/\partial y$ and $\nabla_{3D} = \hat{\boldsymbol{x}}\partial/\partial x + \hat{\boldsymbol{y}}\partial/\partial y + \hat{\boldsymbol{z}}\partial/\partial z$.

A feature of the potential vorticity field is that in general the quantity is non-zero even for a fluid at rest. This can be seen from the expression given inside the square bracket of (A 3): notice that fN^2 is the quantity that remains when we set $\boldsymbol{v} = 0, b = 0$. To overcome this and to restrict potential vorticity to dynamics of the flow, Wagner & Young [71] introduced the concept of available potential vorticity (APV). Similar to the concept of available potential energy, APV is obtained by subtracting the potential vorticity contribution of a fluid at rest from the full potential vorticity expression. After manipulating the non-dimensionalized APV equation (see the procedure detailed in §3 of [71]), we get

$$\frac{\partial Q}{\partial t} + \boldsymbol{v} \cdot \nabla Q + w\frac{\partial Q}{\partial z} = 0, \ Q = \Pi + \epsilon\Gamma + \epsilon^2\Theta, \tag{A 4a}$$

where

$$\Pi = N^2\left[\zeta + \frac{\partial}{\partial z}\left(\frac{b}{N^2}\right)\right], \ \Gamma = \left(\hat{\boldsymbol{z}} \times \frac{\partial\boldsymbol{v}}{\partial z}\right) \cdot \nabla b + \zeta\frac{\partial b}{\partial z} - \frac{\Lambda''}{2N^2}b^2 \tag{A 4b}$$

$$\Theta = -\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\Lambda''}{N^2}\right)\frac{b^3}{6N^2} \ \text{ and } \ \Lambda'' = \frac{\mathrm{d}^2\Lambda}{\mathrm{d}z^2} \ \text{ with } \ \Lambda = \ln N^2 \tag{A 4c}$$

Appendix B. $O(\epsilon)$ equations and averaging

At $O(\epsilon)$ (2.1) gives us

$$\frac{\partial \boldsymbol{v}_1}{\partial t} + \hat{\boldsymbol{z}} \times \boldsymbol{v}_1 + \nabla p_1 + \frac{\partial \boldsymbol{v}_0}{\partial T} + \boldsymbol{v}_0 \cdot \nabla \boldsymbol{v}_0 + w_0 \frac{\partial \boldsymbol{v}_0}{\partial z} = 0, \tag{B 1a}$$

$$\frac{\partial b_1}{\partial t} + N^2 w_1 + \frac{\partial b_0}{\partial T} + \boldsymbol{v}_0 \cdot \nabla b_0 + w_0 \frac{\partial b_0}{\partial z} = 0, \tag{B 1b}$$

$$\frac{\partial p_1}{\partial z} = b_1 \tag{B 1c}$$

and

$$\nabla \cdot \boldsymbol{v}_1 + \frac{\partial w_1}{\partial z} = 0. \tag{B 1d}$$

which on fast time averaging gives us

$$\hat{\boldsymbol{z}} \times \overline{\boldsymbol{v}}_1 + \nabla \overline{p}_1 + \frac{\partial \overline{\boldsymbol{v}}_0}{\partial T} + \overline{\boldsymbol{v}}_0 \cdot \nabla \overline{\boldsymbol{v}}_0 + \overline{\boldsymbol{v}_0' \cdot \nabla \boldsymbol{v}_0'} + \overline{w_0' \frac{\partial \boldsymbol{v}_0'}{\partial z}} = 0, \tag{B 2a}$$

$$N^2 \overline{w}_1 + \frac{\partial \overline{b}_0}{\partial T} + \overline{\boldsymbol{v}}_0 \cdot \nabla \overline{b}_0 + \overline{\boldsymbol{v}_0' \cdot \nabla b_0'} + \overline{w_0' \frac{\partial b_0'}{\partial z}} = 0, \tag{B 2b}$$

$$\frac{\partial \overline{p}_1}{\partial z} = \overline{b}_1 \tag{B 2c}$$

and

$$\nabla \cdot \overline{\boldsymbol{v}}_1 + \frac{\partial \overline{w}_1}{\partial z} = 0. \tag{B 2d}$$

We now use the following identities in the above equations:

$$\overline{\boldsymbol{v}_0' \cdot \nabla \boldsymbol{v}_0'} + \overline{w_0' \frac{\partial \boldsymbol{v}_0}{\partial z}} = \hat{\boldsymbol{z}} \times \overline{\boldsymbol{v}}^S + \frac{1}{2} \nabla \overline{p}^S \tag{B 3a}$$

and

$$\overline{\boldsymbol{v}_0' \cdot \nabla b_0'} + \overline{w_0' \frac{\partial b_0'}{\partial z}} = N^2 \overline{w}^S, \tag{B 3b}$$

where the Stokes velocity ($\overline{\boldsymbol{v}}^S \hat{\boldsymbol{x}} + \overline{w}^S \hat{\boldsymbol{z}}$) and pressure field ($\overline{p}^S$) above are given by

$$(\overline{\boldsymbol{v}}^S, \overline{w}^S, \overline{p}^S) = \overline{\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z} \right) (\boldsymbol{v}_0', w_0', p_0')}. \tag{B 4}$$

The identities in (B 3) were derived in [71]. However, we have included the necessary steps in appendix D so that the reader can follow the calculations without a break in continuity.

Substituting (B 3*a*) and (B 3*b*) in (B 2*a*) and (B 2*b*) gives us

$$\overline{\boldsymbol{v}}_1 + \overline{\boldsymbol{v}}^S = \hat{\boldsymbol{z}} \times \nabla \left(\overline{p}_1 + \frac{\overline{p}^S}{2} \right) - \nabla \frac{\partial \overline{p}_0}{\partial T} - (\overline{\boldsymbol{v}}_0 \cdot \nabla) \nabla \overline{p}_0 \tag{B 5a}$$

and

$$\overline{w}_1 + \overline{w}^S = -\frac{1}{N^2} \left(\frac{\partial^2 \overline{p}_0}{\partial T \partial z} + \overline{\boldsymbol{v}}_0 \cdot \nabla \frac{\partial \overline{p}_0}{\partial z} \right). \tag{B 5b}$$

We now take the curl of (B 5*a*) to obtain

$$\overline{\zeta}_1 = \Delta \left(\overline{p}_1 + \frac{\overline{p}^S}{2} \right) - 2\partial \left[\frac{\partial \overline{p}_0}{\partial x}, \frac{\partial \overline{p}_0}{\partial y} \right] - \hat{\boldsymbol{z}} \cdot \nabla \times \overline{\boldsymbol{v}}^S, \tag{B 6}$$

where $\overline{\zeta}_1 = \hat{\boldsymbol{z}} \cdot \nabla \times \overline{\boldsymbol{v}}_1$ and $\partial[f,g] = \partial f/\partial x \partial g/\partial y - \partial f/\partial y \partial g/\partial x$ is the Jacobean. Thus, we obtain the slow part of APV at this order as

$$\begin{aligned}\overline{\Pi}_1 &= N^2 \left\{ \overline{\zeta}_1 + \frac{\partial}{\partial z}\left(\frac{\overline{b}_1}{N^2}\right)\right\} \\ &= N^2 \left\{ \Delta\left(\overline{p}_1 + \frac{\overline{p}^S}{2}\right) + \frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial \overline{p}_1}{\partial z}\right) - 2\partial\left[\frac{\partial \overline{p}_0}{\partial x}, \frac{\partial \overline{p}_0}{\partial y}\right] - \hat{\boldsymbol{z}} \cdot \nabla \times \overline{\boldsymbol{v}}^S \right\}, \end{aligned} \tag{B 7}$$

where we used (B 2*c*) to arrive at the last expression above.

The APV equation (2.2) at order $O(\epsilon)$ gives us

$$\frac{\partial}{\partial t}(\Pi_1 + \Gamma_0) + \frac{\partial \Pi_0}{\partial T} + \left(\boldsymbol{v}_0 \cdot \nabla + w_0 \frac{\partial}{\partial z}\right)\overline{\Pi}_0 = 0, \tag{B 8}$$

where

$$\Gamma_0 = \left(\hat{\boldsymbol{z}} \times \frac{\partial \boldsymbol{v}_0}{\partial z}\right) \cdot \nabla b_0 + \zeta \frac{\partial b_0}{\partial z} - \frac{\Lambda''}{2N^2} {b_0}^2. \tag{B 9}$$

Fast time averaging (B 8) to eliminate secular growth gives us

$$\frac{\partial \overline{\Pi}_0}{\partial T} + \overline{\boldsymbol{v}}_0 \cdot \nabla \overline{\Pi}_0 = 0. \tag{B 10}$$

Using the linear equations connecting balanced flow variables in (2.3) simplifies (B 10) to give us (2.5).

It is interesting to note that even though at the present asymptotic order of $O(\epsilon)$ fast waves project on slow variables, as can be seen from (B 2) and (B 5) for example, waves do not affect leading-order APV or project on the leading-order slow APV evolution equation (2.5). To capture the effect of waves on slow balanced dynamics, we need to proceed to the next asymptotic order. As mentioned earlier, the leading-order QG equation (2.5) is formally valid for $T \sim O(1)$ time scales. To capture the effect of waves on slow dynamics, we need to derive a higher-order approximation of the slow equation valid for longer time scales, $T \sim (1/\epsilon)$. With this goal, we set

$$\frac{\partial \overline{\Pi}_0}{\partial T} + \overline{\boldsymbol{v}}_0 \cdot \nabla \overline{\Pi}_0 = \epsilon \Phi, \tag{B 11}$$

where Φ is a correction term and will be obtained by eliminating secular growth at next order of asymptotics. This procedure of obtaining higher-order correction term is an asymptotic strategy that is described in [89] and was used in the same form as above in [70] to derive higher-order corrections to QG in rotating shallow water equations. We will determine the correction term Φ by eliminating resonant growth at $O(\epsilon^2)$ below.

To complete the asymptotic procedure at $O(\epsilon)$, we first subtract (B 10) from (B 8) and use the fast–slow decomposition $\Pi_1 = \overline{\Pi}_1 + \Pi_1'$ and $\Gamma_0 = \overline{\Gamma}_0 + \Gamma_0'$ in (B 8) to get

$$\frac{\partial}{\partial t}(\Pi_1' + \Gamma_0') + \left(\boldsymbol{v}_0' \cdot \nabla + w_0' \frac{\partial}{\partial z}\right)\overline{\Pi}_0 = 0, \tag{B 12}$$

which can be integrated in fast time t to obtain

$$\Pi_1' + \Gamma_0' = -\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z}\right)\overline{\Pi}_0. \tag{B 13}$$

We then time average (B 9) and use the linear equations (2.3) to get

$$\overline{\Gamma}_0 = \overline{\zeta}_0 \frac{\partial \overline{b}_0}{\partial z} + \overline{\left(\hat{\boldsymbol{z}} \times \frac{\partial \overline{\boldsymbol{v}}_0}{\partial z}\right) \cdot \nabla \overline{b}_0} - \frac{\Lambda''}{2N^2} \overline{{b}_0}^2 + \overline{\zeta_0' \frac{\partial b_0'}{\partial z}} + \overline{\left(\hat{\boldsymbol{z}} \times \frac{\partial \boldsymbol{v}_0'}{\partial z}\right) \cdot \nabla b_0'} - \frac{\Lambda''}{2N^2} \overline{{b_0'}^2} \tag{B 14}$$

$$= \frac{\partial^2 \overline{p}_0}{\partial z^2} \Delta \overline{p}_0 - \left(\nabla \frac{\partial \overline{p}_0}{\partial z}\right)^2 - \frac{\Lambda''}{2N^2}\left(\frac{\partial \overline{p}_0}{\partial z}\right)^2 + \overline{\zeta_0' \frac{\partial b_0'}{\partial z}} + \overline{\left(\hat{\boldsymbol{z}} \times \frac{\partial \boldsymbol{v}_0'}{\partial z}\right) \cdot \nabla b_0'} - \frac{\Lambda''}{2N^2} \overline{{b_0'}^2}. \tag{B 15}$$

At this point, we have all the variables required at $O(\epsilon)$ and we will therefore proceed to $O(\epsilon^2)$.

Appendix C. $O(\epsilon)^2$ equations and averaging

The APV equation (2.2) at $O(\epsilon^2)$ is

$$\frac{\partial}{\partial t}(\Pi_2+\Gamma_1+\Theta_0)+\frac{\partial}{\partial T}(\Pi_1+\Gamma_0)+\left(\boldsymbol{v}_0\cdot\nabla+w_0\frac{\partial}{\partial z}\right)(\Pi_1+\Gamma_0)+\left(\boldsymbol{v}_1\cdot\nabla+w_1\frac{\partial}{\partial z}\right)\overline{\Pi}_0+\Phi=0, \tag{C 1}$$

which on fast time averaging gives us

$$\frac{\partial}{\partial T}(\overline{\Pi}_1+\overline{\Gamma}_0)+\overline{\boldsymbol{v}}_0\cdot\nabla(\overline{\Pi}_1+\overline{\Gamma}_0)+\overline{\left(\boldsymbol{v}_0'\cdot\nabla+w_0'\frac{\partial}{\partial z}\right)(\Pi_1'+\Gamma_0')}+\left(\overline{\boldsymbol{v}}_1\cdot\nabla+\overline{w}_1\frac{\partial}{\partial z}\right)\overline{\Pi}_0+\Phi=0, \tag{C 2}$$

where we used $\overline{w}_0=0$ from (2.3).

We now use (B 13) to manipulate a term in (C 2) as

$$\begin{aligned}
&\overline{\left(\boldsymbol{v}_0'\cdot\nabla+w_0'\frac{\partial}{\partial z}\right)(\Pi_1'+\Gamma_0')}\\
&=-\overline{\left(\boldsymbol{v}_0'\cdot\nabla+w_0'\frac{\partial}{\partial z}\right)\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)}\overline{\Pi}_0=\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\left(\boldsymbol{v}_0'\cdot\nabla+w_0'\frac{\partial}{\partial z}\right)}\overline{\Pi}_0\\
&=\left\{\left[\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\boldsymbol{v}_0'}\right]\cdot\nabla+\left[\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)w_0'}\right]\frac{\partial}{\partial z}\right\}\overline{\Pi}_0\\
&\quad+\overline{((\boldsymbol{\xi}_0'+\eta_0'\hat{\boldsymbol{z}})(\boldsymbol{v}_0'+w_0'\hat{\boldsymbol{z}}))}\cdot\cdot\nabla\nabla\overline{\Pi}_0\\
&=\left(\overline{\boldsymbol{v}}^S\cdot\nabla+\overline{w}^S\frac{\partial}{\partial z}\right)\overline{\Pi}_0.
\end{aligned} \tag{C 3}$$

Above the double contraction term vanishes, as may be verified using the Einstein summation convention. Substituting (C 3) in (C 2) and then forming the sum (B 11) + ϵ(C 2) eliminates Φ and gives us

$$\frac{\partial}{\partial T}[\overline{\Pi}_0+\epsilon(\overline{\Pi}_1+\overline{\Gamma}_0)]+\overline{\boldsymbol{v}}_0\cdot\nabla[\overline{\Pi}_0+\epsilon(\overline{\Pi}_1+\overline{\Gamma}_0)]+\epsilon\left[(\overline{\boldsymbol{v}}_1+\overline{\boldsymbol{v}}^S)\cdot\nabla+(\overline{w}_1+\overline{w}^S)\frac{\partial}{\partial z}\right]\overline{\Pi}_0=O(\epsilon^2). \tag{C 4}$$

We now note the following two equations derived in appendix E:

$$(\overline{\boldsymbol{v}}_1+\overline{\boldsymbol{v}}^S)\cdot\nabla\overline{\Pi}_0=\left[\hat{\boldsymbol{z}}\times\nabla\left(\overline{p}_1-\frac{(\nabla\overline{p}_0)^2}{2}+\frac{\overline{p}^S}{2}\right)\right]\cdot\nabla\overline{\Pi}_0-\frac{\partial}{\partial T}[\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0]-(\overline{\boldsymbol{v}}_0\cdot\nabla)(\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0)+O(\epsilon) \tag{C 5}$$

and

$$(\overline{w}_1+\overline{w}^S)\frac{\partial\overline{\Pi}_0}{\partial z}=-\frac{1}{N^2}\frac{\partial}{\partial T}\left(\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)-\frac{1}{N^2}\left(\overline{\boldsymbol{v}}_0\cdot\nabla\left(\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)+\left[\hat{\boldsymbol{z}}\times\nabla\left(\frac{1}{2}\left(\frac{\partial\overline{p}_0}{\partial z}\right)^2\right)\right]\cdot\nabla\overline{\Pi}_0\right)+O(\epsilon). \tag{C 6}$$

Substituting (C 5) and (C 6) in (C 4) gives us

$$\frac{\partial}{\partial T}\left[\overline{\Pi}_0+\epsilon\left(\overline{\Pi}_1+\overline{\Gamma}_0-\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0-\frac{1}{N^2}\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)\right]$$
$$+\overline{\boldsymbol{v}}_0\cdot\nabla\left[\overline{\Pi}_0+\epsilon\left(\overline{\Pi}_1+\overline{\Gamma}_0-\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0-\frac{1}{N^2}\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)\right]$$
$$+\epsilon\left[\hat{\boldsymbol{z}}\times\nabla\left(\overline{p}_1+\frac{\overline{p}^S}{2}-\frac{(\nabla\overline{p}_0)^2}{2}-\frac{1}{2N^2}\left(\frac{\partial\overline{p}_0}{\partial z}\right)^2\right)\right]\cdot\nabla\overline{\Pi}_0=O(\epsilon^2). \tag{C 7}$$

To derive a consistent slow evolution equation, we combine slow fields at $O(1)$ and $O(\epsilon)$ as $\overline{\psi}=\overline{p}_0+\epsilon\overline{p}_1$ and drop the $O(\epsilon^2)$ higher-order error term on the right-hand side of (C 7) to get (2.6).

Appendix D. Stokes identities

For convenient manipulation of expressions that follow, we introduce particle displacement field, $(\boldsymbol{\xi}_0',\eta_0')$ so that $\boldsymbol{v}_0'=\partial\boldsymbol{\xi}_0'/\partial t$, $w_0'=\partial\eta_0'/\partial t$. Using this, we may integrate the buoyancy and continuity equations for the fast fields in (2.3) to obtain

$$b_0'=-N^2\eta_0',\quad \nabla\cdot\boldsymbol{\xi}_0'+\frac{\partial\eta_0'}{\partial z}=0 \tag{D 1}$$

$$\overline{\left(\boldsymbol{v}_0'\cdot\nabla+w_0'\frac{\partial}{\partial z}\right)\boldsymbol{v}_0'}=-\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\frac{\partial\boldsymbol{v}_0'}{\partial t}}$$
$$=\hat{\boldsymbol{z}}\times\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\boldsymbol{v}_0'}+\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\nabla p_0'}, \tag{D 2}$$

where we used $\partial\boldsymbol{v}_0'/\partial t=-\hat{\boldsymbol{z}}\times\boldsymbol{v}_0'-\nabla p_0'$ from the leading-order equations.

$$\overline{p}^S=\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)p_0'} \tag{D 3}$$

and

$$\frac{\partial\overline{p}^S}{\partial x}=\overline{\left(\boldsymbol{\xi}_0'\cdot\nabla+\eta_0'\frac{\partial}{\partial z}\right)\frac{\partial p_0'}{\partial x}}+\overline{\left(\frac{\partial\boldsymbol{\xi}_0'}{\partial x}\cdot\nabla+\frac{\partial\eta_0'}{\partial x}\frac{\partial}{\partial z}\right)p_0'} \tag{D 4}$$

Taking the dot product of linear momentum equation with $\boldsymbol{\xi}_0'$ and fast time averaging, we get:

$$-\overline{\boldsymbol{v}_0'^2}+\overline{\boldsymbol{\xi}_0'\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{v}_0')}+\overline{\boldsymbol{\xi}_0'\cdot\nabla p_0'}=0. \tag{D 5}$$

We further note that

$$\overline{\eta_0'\frac{\partial p_0'}{\partial z}}=\overline{\eta_0'b_0'}=-N^2\overline{\eta_0'^2} \tag{D 6}$$

Using (D 5), (D 6) and (D 3), we get

$$\overline{p}^S=\overline{\boldsymbol{v}_0'^2}+\overline{\boldsymbol{\xi}_0'\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{v}_0')}-N^2\overline{\eta_0'^2}, \tag{D 7}$$

$$\frac{\partial}{\partial x}\left[\overline{\boldsymbol{\xi}_0'\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{v}_0')}\right]=\overline{\frac{\partial\boldsymbol{\xi}_0'}{\partial x}\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{v}_0')}+\overline{\boldsymbol{\xi}_0'\cdot\left(\hat{\boldsymbol{z}}\times\frac{\partial\boldsymbol{v}_0'}{\partial x}\right)} \tag{D 8}$$

and

$$\overline{\frac{\partial\boldsymbol{\xi}_0'}{\partial x}\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{v}_0')}=\overline{\frac{\partial\boldsymbol{\xi}_0'}{\partial x}\cdot\left(\hat{\boldsymbol{z}}\times\frac{\partial\boldsymbol{\xi}_0'}{\partial t}\right)}=-\overline{\frac{\partial^2\boldsymbol{\xi}_0'}{\partial t\partial x}\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{\xi}_0')}$$
$$=-\overline{\frac{\partial\boldsymbol{v}_0'}{\partial x}\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{\xi}_0')}=\overline{\boldsymbol{\xi}_0'\cdot\left(\hat{\boldsymbol{z}}\times\frac{\partial\boldsymbol{v}_0'}{\partial x}\right)}. \tag{D 9}$$

Using (D 9) in (D 8), we get

$$\overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot (\hat{\boldsymbol{z}} \times \boldsymbol{v}_0')} = \frac{1}{2} \frac{\partial}{\partial x} \left[\overline{\boldsymbol{\xi}_0' \cdot (\hat{\boldsymbol{z}} \times \boldsymbol{v}_0')} \right]. \tag{D 10}$$

Taking the dot product of linear momentum equation with $\partial \boldsymbol{\xi}_0'/\partial x$ and fast time averaging, we get

$$\overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot \frac{\partial \boldsymbol{v}_0'}{\partial t}} + \overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot (\hat{\boldsymbol{z}} \times \boldsymbol{v}_0')} + \overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot \nabla p_0'} = 0. \tag{D 11}$$

The first term in the above equation can be modified as

$$\overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot \frac{\partial \boldsymbol{v}_0'}{\partial t}} = -\overline{\frac{\partial \boldsymbol{v}_0'}{\partial x} \cdot \boldsymbol{v}_0'} = -\frac{\partial}{\partial x} \left(\frac{\boldsymbol{v}_0'^2}{2} \right). \tag{D 12}$$

Further,

$$\overline{\frac{\partial \eta_0'}{\partial x} \frac{\partial p_0'}{\partial z}} = \overline{\frac{\partial \eta_0'}{\partial x} b_0'} = -N^2 \overline{\eta_0' \frac{\partial \eta_0'}{\partial x}} = -N^2 \frac{\partial}{\partial x} \left(\frac{\eta_0'^2}{2} \right). \tag{D 13}$$

Using (D 10) and (D 12) in (D 11), we get

$$\overline{\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot \nabla p_0'} = \frac{1}{2} \frac{\partial}{\partial x} \left(\frac{\boldsymbol{v}_0'^2}{2} - \overline{\boldsymbol{\xi}_0' \cdot (\hat{\boldsymbol{z}} \times \boldsymbol{v}_0')} \right). \tag{D 14}$$

Using (D 13), (D 14) and (D 7), we get

$$\overline{\left(\frac{\partial \boldsymbol{\xi}_0'}{\partial x} \cdot \nabla + \frac{\partial \eta_0'}{\partial x} \frac{\partial}{\partial z} \right) p_0'} = \frac{1}{2} \frac{\partial \overline{p}^S}{\partial x}. \tag{D 15}$$

Using the above equation in (D 4), we get

$$\overline{\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z} \right) \frac{\partial p_0'}{\partial x}} = \frac{1}{2} \frac{\partial \overline{p}^S}{\partial x}. \tag{D 16}$$

One can obtain a similar expression with y derivative replacing x derivative above. Combining these, we get

$$\overline{\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z} \right) \nabla p_0'} = \frac{1}{2} \nabla \overline{p}^S. \tag{D 17}$$

Using (D 17) and (B 4) in (D 2), we get (B 3*a*). Further,

$$\begin{aligned} \overline{\left(\boldsymbol{v}_0' \cdot \nabla + w_0' \frac{\partial}{\partial z} \right) b_0'} &= -\overline{\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z} \right) \frac{\partial b_0'}{\partial t}} \\ &= N^2 \overline{\left(\boldsymbol{\xi}_0' \cdot \nabla + \eta_0' \frac{\partial}{\partial z} \right) w_0'} + \frac{\partial (N^2)}{\partial z} \overline{\eta_0' w_0'}. \end{aligned} \tag{D 18}$$

The final term above vanishes since $w_0' = \partial \eta_0'/\partial t$ and thus we obtain (B 3*b*).

Appendix E. Derivation of equations (C 5) and (C 6)

We use (B 5*a*) and (B 5*b*) to obtain

$$\begin{aligned}
&(\overline{\boldsymbol{v}}_1+\overline{\boldsymbol{v}}^S)\cdot\nabla\overline{\Pi}_0\\
&=\left[\hat{\boldsymbol{z}}\times\nabla\left(\overline{p}_1+\frac{\overline{p}^S}{2}\right)\right]\cdot\nabla\overline{\Pi}_0-\nabla\frac{\partial\overline{p}_0}{\partial T}\cdot\nabla\overline{\Pi}_0-[(\overline{\boldsymbol{v}}_0\cdot\nabla)\nabla\overline{p}_0]\cdot\nabla\overline{\Pi}_0\\
&=\left[\hat{\boldsymbol{z}}\times\nabla\left(\overline{p}_1+\frac{\overline{p}^S}{2}\right)\right]\cdot\nabla\overline{\Pi}_0-[(\overline{\boldsymbol{v}}_0\cdot\nabla)\nabla\overline{p}_0]\cdot\nabla\overline{\Pi}_0-\frac{\partial}{\partial T}[\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0]+\nabla\overline{p}_0\cdot\nabla\frac{\partial\overline{\Pi}_0}{\partial T}\\
&=\left[\hat{\boldsymbol{z}}\times\nabla\left(\overline{p}_1+\frac{\overline{p}^S}{2}\right)\right]\cdot\nabla\overline{\Pi}_0-[(\overline{\boldsymbol{v}}_0\cdot\nabla)\nabla\overline{p}_0]\cdot\nabla\overline{\Pi}_0\\
&\quad-\frac{\partial}{\partial T}[\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0]-[\nabla\overline{p}_0\cdot\nabla](\overline{\boldsymbol{v}}_0\cdot\nabla\overline{\Pi}_0)+O(\epsilon),
\end{aligned}\tag{E 1a}$$

where we used (B 11). We now use the identity

$$\begin{aligned}
&(\nabla\overline{p}_0\cdot\nabla)(\overline{\boldsymbol{v}}_0\cdot\nabla\overline{\Pi}_0)+[(\overline{\boldsymbol{v}}_0\cdot\nabla)\nabla\overline{p}_0]\cdot\nabla\overline{\Pi}_0\\
&\quad=\left[\hat{\boldsymbol{z}}\times\nabla\left(\frac{(\nabla\overline{p}_0)^2}{2}\right)\right]\cdot\nabla\overline{\Pi}_0+(\overline{\boldsymbol{v}}_0\cdot\nabla)(\nabla\overline{p}_0\cdot\nabla\overline{\Pi}_0),
\end{aligned}\tag{E 2}$$

in (E 1*a*) to obtain (C 5). Similarly,

$$\begin{aligned}
(\overline{w}_1+\overline{w}^S)\frac{\partial\overline{\Pi}_0}{\partial z}&=-\frac{1}{N^2}\left(\frac{\partial^2\overline{p}_0}{\partial z\partial T}+\overline{\boldsymbol{v}}_0\cdot\nabla\frac{\partial\overline{p}_0}{\partial z}\right)\frac{\partial\overline{\Pi}_0}{\partial z}\\
&=-\frac{1}{N^2}\frac{\partial}{\partial T}\left(\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)\\
&\quad-\frac{1}{N^2}\left[\frac{\partial\overline{\Pi}_0}{\partial z}\overline{\boldsymbol{v}}_0\cdot\nabla\frac{\partial\overline{p}_0}{\partial z}+\frac{\partial\overline{p}_0}{\partial z}\frac{\partial}{\partial z}(\overline{\boldsymbol{v}}_0\cdot\nabla\overline{\Pi}_0)\right]+O(\epsilon)
\end{aligned}\tag{E 3}$$

We now use the identity:

$$\begin{aligned}
&\frac{\partial\overline{\Pi}_0}{\partial z}\overline{\boldsymbol{v}}_0\cdot\nabla\frac{\partial\overline{p}_0}{\partial z}+\frac{\partial\overline{p}_0}{\partial z}\frac{\partial}{\partial z}(\overline{\boldsymbol{v}}_0\cdot\nabla\overline{\Pi}_0)\\
&\quad=\overline{\boldsymbol{v}}_0\cdot\nabla\left(\frac{\partial\overline{p}_0}{\partial z}\frac{\partial\overline{\Pi}_0}{\partial z}\right)+\left[\hat{\boldsymbol{z}}\times\nabla\left(\frac{1}{2}\left(\frac{\partial\overline{p}_0}{\partial z}\right)^2\right)\right]\cdot\nabla\overline{\Pi}_0,
\end{aligned}\tag{E 4}$$

in (E 3) to get (C 6).

References

1. Warren BA, Wunsch C. 1981 Evolution of physical oceanography: scientific surveys in honor of Henry Stommel. *The MIT Press* **664**, 6–41.
2. Richardson PL. 2008 On the history of meridional overturning circulation schematic diagrams. *Prog. Oceanogr.* **76**, 466–486. (doi:10.1016/j.pocean.2008.01.005)
3. Ferrari R, Wunsch C. 2009 Ocean circulation kinetic energy: reservoirs, sources and sinks. *Annu. Rev. Fluid Mech.* **41**, 253–282. (doi:10.1146/annurev.fluid.40.111406.102139)
4. Alford M, MacKinnon J, Simmons H, Nash J. 2016 Near-inertial internal gravity waves in the ocean. *Annu. Rev. Marine Sci.* **8**, 95–123. (doi:10.1146/annurev-marine-010814-015746)
5. Garrett C, Kunze E. 2007 Internal tide generation in the deep ocean. *Annu. Rev. Fluid Mech.* **39**, 57–87. (doi:10.1146/annurev.fluid.39.050905.110227)
6. Stammer D. 1997 Global characteristics of ocean variability estimated from regional TOPEX/Poseidon altimeter measurements. *J. Phys. Oceanogr.* **27**, 1743–1769. (doi:10.1175/1520-0485(1997)027<1743:GCOOVE>2.0.CO;2)

7. Vallis GK. 2006 *Atmospheric and oceanic fluid dynamics*. Cambridge, UK: Cambridge University Press.
8. Scott RB, Wang FM. 2005 Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry. *J. Phys. Oceanogr.* **35**, 1650–1666. (doi:10.1175/JPO2771.1)
9. Chelton DB, Schlax MG, Samelson RM. 2011 Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* **91**, 167–216. (doi:10.1016/j.pocean.2011.01.002)
10. Richardson LF. 1922 *Weather prediction by numerical process*. Cambridge, UK: Cambridge University Press.
11. Lynch P. 2006 *The emergence of numerical weather prediction: Richardson's dream*. Cambridge, UK: Cambridge University Press.
12. Charney JG. 1948 On the scale of atmospheric motions. *Geofys. Publ. Oslo* **17**, 1–17.
13. Obukhov A. 1949 On the question of geostrophic wind. *Bull. USSR Acad. Sci. Geograph. Geophys. Series (in Russian)* **13**, 281–306.
14. Charney JG, Fjortoft R, von Neumann J. 1950 Numerical integration of the barotropic vorticity equation. *Tellus* **2**, 237–254. (doi:10.3402/tellusa.v2i4.8607)
15. Charney JG, Phillips N. 1953 Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. *J. Meteor.* **10**, 71–99. (doi:10.1175/1520-0469(1953)010<0071:NIOTQG>2.0.CO;2)
16. Machenhauer B. 1977 On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitr. Phys. Atmos.* **50**, 253–271.
17. Baer F, Tribbia JJ. 1977 On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.* **105**, 1536–1539. (doi:10.1175/1520-0493(1977)105<1536:OCFOGM>2.0.CO;2)
18. Leith C. 1980 Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.* **37**, 958–968. (doi:10.1175/1520-0469(1980)037<0958:NNMIAQ>2.0.CO;2)
19. Lorenz EN. 1980 Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.* **37**, 1685–1699. (doi:10.1175/1520-0469(1980)037<1685:ASAQGE>2.0.CO;2)
20. Lorenz EN, Krishnamurty V. 1987 On the non-existence of a slow manifold. *J. Atmos. Sci.* **44**, 2940–2950. (doi:10.1175/1520-0469(1987)044<2940:OTNOAS>2.0.CO;2)
21. Lorenz EN. 1986 On the existence of a slow manifold. *J. Atmos. Sci.* **43**, 1547–1557. (doi:10.1175/1520-0469(1986)043<1547:OTEOAS>2.0.CO;2)
22. Vautard R, Legras B. 1986 Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci.* **43**, 565–584. (doi:10.1175/1520-0469(1986)043<0565:IMQGAI>2.0.CO;2)
23. Vanneste J, Yavneh I. 2004 Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.* **61**, 211–223. (doi:10.1175/1520-0469(2004)061<0211:ESIWAT>2.0.CO;2)
24. Plougonven R, Zhang F. 2014 Internal gravity waves from atmospheric jets and fronts. *Rev. Geophys.* **52**, 33–76. (doi:10.1002/2012RG000419)
25. Williams P, Haine T, Read P. 2008 Inertia–gravity waves emitted from balanced flow: observations, properties, and consequences. *J. Atmos. Sci.* **65**, 3543–3556. (doi:10.1175/2008JAS2480.1)
26. Chouksey M, Eden C, Bruggemann N. 2018 Internal gravity wave emission in different dynamical regimes. *J. Phys. Oceanogr.* **48**, 1709–1730. (doi:10.1175/JPO-D-17-0158.1)
27. Eden C, Chouksey M, Olbers D. 2019 Gravity wave emission by shear instability. *J. Phys. Oceanogr.* **49**, 2393–2406. (doi:10.1175/JPO-D-19-0029.1)
28. Vanneste J. 2013 Balance and spontaneous wave generation in geophysical flows. *Annu. Rev. Fluid Mech.* **45**, 147–172. (doi:10.1146/annurev-fluid-011212-140730)
29. Pujol MI, Dibarboure G, Le Traon PY, Klein P. 2012 Using high-resolution altimetry to observe mesoscale signals. *J. Atmos. Oceanic Technol.* **29**, 1409–1416. (doi:10.1175/JTECH-D-12-00032.1)
30. Le Traon PY. 2013 From satellite altimetry to argo and operational oceanography: three revolutions in oceanography. *Ocean Sci.* **9**, 901–915. (doi:10.5194/os-9-901-2013)
31. Dufau C, Orsztynowicz M, Dibarboure G, Morrow R, Le Traon P. 2016 Mesoscale resolution capability of altimetry: present and future. *J. Geophys. Res. Oceans* **121**, 4910–4927. (doi:10.1002/2015JC010904)
32. Lumpkin R, Ozgokmen T, Centurioni L. 2017 Advances in the application of surface drifters. *Annu. Rev. Marine Sci.* **9**, 59–81. (doi:10.1146/annurev-marine-010816-060641)
33. Klein P, Lapeyre G, Siegelman L, Qiu B, Fu LL, Torres H, Su Z, Menemenlis D, Le Gentil S. 2019 Ocean–scale interactions from space. *Earth Space Sci.* **6**, 795–817. (doi:10.1029/2018EA000492)

34. Gregory CJ *et al.* 2022 Argo-two decades: global oceanography, revolutionized. *Annu. Rev. Marine Sci.* **14**, 379–403. (doi:10.1146/annurev-marine-022521-102008)
35. Richman JG, Arbic BK, Shriver JF, Metzger EJ, Wallcraft AJ. 2012 Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides. *J. Geophys. Res.* **117**, C12012. (doi:10.1029/2012JC008364)
36. Bühler O, Callies J, Ferrari R. 2014 Wave-vortex decomposition of one-dimensional ship-track data. *J. Fluid Mech.* **756**, 1007–1026. (doi:10.1017/jfm.2014.488)
37. Pinkel R. 2014 Vortical and internal wave shear and strain. *J. Phys. Oceanogr.* **44**, 2070–2092. (doi:10.1175/JPO-D-13-090.1)
38. Qiu B, Nakano T, Chen S, Klein P. 2017 Submesoscale transition from geostrophic flows to internal waves in the northwestern Pacific upper ocean. *Nat. Commun.* **8**, 14055. (doi:10.1038/ncomms14055)
39. Savage A *et al.* 2017 Spectral decomposition of internal gravity wave sea surface height in global models. *J. Geophys. Res. Oceans* **122**, 7803–7821. (doi:10.1002/2017JC013009)
40. Qiu B, Chen S, Klein P, Wang J, Torres H, Fu L, Menemenlis D. 2018 Seasonality in transition scale from balanced to unbalanced motions in the world ocean. *J. Phys. Oceanogr.* **48**, 591–605. (doi:10.1175/JPO-D-17-0169.1)
41. Tchilibou M, Gourdeau L, Morrow R, Serazin G, Djath B, Lyard F. 2018 Spectral signatures of the tropical Pacific dynamics from model and altimetry: a focus on the meso/submesoscale range. *Ocean Sci.* **14**, 1283–1301. (doi:10.5194/os-14-1283-2018)
42. Torres HS, Klein P, Menemenlis D, Qiu B, Su Z, Wang J, Chen S, Fu LL. 2018 Partitioning ocean motions into balanced motions and internal gravity waves: a modeling study in anticipation of future space missions. *J. Geophys. Res. Oceans* **123**, 8084–8105. (doi:10.1029/2018JC014438)
43. Lien RC, Sanford TB. 2019 Small-scale potential vorticity in the upper ocean thermocline. *J. Phys. Oceanogr.* **49**, 1845–1872. (doi:10.1175/JPO-D-18-0052.1)
44. Yu X, Ponte AL, Elipot S, Menemenlis D, Zaron ED, Abernathey R. 2019 Surface kinetic energy distributions in the global oceans from a high-resolution numerical model and surface drifter observations. *Geophys. Res. Lett.* **46**, 9757–9766. (doi:10.1029/2019GL083074)
45. Qiu B, Nakano T, Chen S, Klein P. 2022 Bi-directional energy cascades in the pacific ocean from equator to subarctic gyre. *Geophys. Res. Lett.* **49**, e2022GL097713. (doi:10.1029/2022GL097713)
46. Fu LL, Keffer T, Niiler P, Wunsch C. 1982 Observations of mesoscale variability in the western North Atlantic: a comparative study. *J. Mar. Res.* **40**, 809–848.
47. Wunsch C, Stammer D. 1998 Satellite altimetry, the marine geoid and the oceanic general circulation. *Annu. Rev. Earth Planet. Sci.* **26**, 219–254. (doi:10.1146/annurev.earth.26.1.219)
48. Ferrari R, Wunsch C. 2010 The distribution of eddy kinetic and potential energies in the global ocean. *Tellus. Ser. A* **60**, 92–108. (doi:10.3402/tellusa.v62i2.15680)
49. Sen A, Scott RB, Arbic BK. 2013 Global energy dissipation rate of deep-ocean low-frequency flows by quadratic bottom boundary layer drag: comparisons from current-meter data. *Geophys. Res. Lett.* **35**, L09606.
50. Dewar WK, Hogg AM. 2010 Topographic inviscid dissipation of balanced flow. *Ocean Modell.* **32**, 1–13. (doi:10.1016/j.ocemod.2009.03.007)
51. Arbic B *et al.* 2009 Estimates of bottom flows and bottom boundary layer dissipation of the oceanic general circulation from global high-resolution models. *J. Geophys. Res.* **114**, C02024. (doi:10.1029/2008JC005072)
52. Zhai X, Johnson HL, Marshall DP. 2010 Significant sink of ocean-eddy energy near western boundaries. *Nat. Geosci.* **3**, 608–612. (doi:10.1038/ngeo943)
53. Dewar WK, Berloff P, Hogg A. 2011 Submesoscale generation by boundaries. *J. Mar. Res.* **69**, 501–522. (doi:10.1357/002224011799849345)
54. Nikurashin M, Ferrari R. 2011 Global energy conversion from geostrophic flow into internal lee waves in the deep ocean. *Geophys. Res. Lett.* **38**, L08610. (doi:10.1029/2011GL046576)
55. Nikurashin M, Vallis GK, Adcroft A. 2013 Routes to energy dissipation for geostrophic flows in the Southern Ocean. *Nat. Geosci.* **6**, 48–51. (doi:10.1038/ngeo1657)
56. Gula J, Molemaker MJ, McWilliams JC. 2016 Topographic generation of submesoscale centrifugal instability and energy dissipation. *Nat. Commun.* **7**, 12811. (doi:10.1038/ncomms12811)
57. Muller P. 1976 On the diffusion of momentum and mass by internal gravity waves. *J. Fluid Mech.* **77**, 789–823. (doi:10.1017/S0022112076002899)

58. Polzin KL. 2010 Mesoscale eddy-internal wave coupling. Part II: energetics and results from PolyMode. *J. Phys. Oceanogr.* **40**, 789–801. (doi:10.1175/2009JPO4039.1)
59. Munk W, Wunsch C. 1998 Abyssal recipes II: energetics of tidal and wind mixing. *Deep-Sea Res. I* **45**, 1977–2010. (doi:10.1016/S0967-0637(98)00070-3)
60. Gregg MC, Sanford TB, Winkel DP. 2003 Reduced mixing from the breaking of internal waves in equatorial waters. *Nature* **422**, 513–515. (doi:10.1038/nature01507)
61. Hu X, Liao G, Fan Z, Liu X. 2023 Impacts of inertial internal wave breaking mixing parameterization on a global ocean model. *Progr. Oceanogr.* **212**, 102997. (doi:10.1016/j.pocean.2023.102997)
62. Jochum M, Briegleb BP, Danabasoglu G, Large WG, Jayne SR, Alford MH, Bryan FO. 2013 The impact of oceanic near-inertial waves on climate. *J. Climate* **26**, 2833–2844. (doi:10.1175/JCLI-D-12-00181.1)
63. Melet A, Legg S, Hallberg R. 2016 Climatic impacts of parameterized local and remote tidal mixing. *J. Climate* **29**, 3473–3500. (doi:10.1175/JCLI-D-15-0153.1)
64. MacKinnon JA *et al.* 2017 Climate process team on internal-wave driven ocean mixing. *Bull. Am. Meteorol. Soc.* **98**, 2429–2454. (doi:10.1175/BAMS-D-16-0030.1)
65. Whalen CB, De Lavergne C, Naveira Garabato AC, Klymak JM, MacKinnon JA, Sheen KL. 2020 Internal wave-driven mixing: governing processes and consequences for climate. *Nat. Rev. Earth Environ.* **1**, 606–621. (doi:10.1038/s43017-020-0097-z)
66. Dewar WK, Killworth PD. 1995 Do fast gravity waves interact with geostrophic motions? *Deep Sea Res. Part I* **42**, 1063–1081. (doi:10.1016/0967-0637(95)00040-D)
67. Majda A, Embid P. 1998 Averaging over fast gravity waves for geophysical flows with unbalanced initial data. *Theoret. Comput. Fluid Dyn.* **11**, 155–169. (doi:10.1007/s001620050086)
68. Reznik GM, Zeitlin V, Ben J. 2001 Nonlinear theory of geostrophic adjustment. Part 1. Rotating shallow-water model. *J. Fluid Mech.* **445**, 93–120. (doi:10.1017/S002211200100550X)
69. Zeitlin V, Reznik GM, Ben J. 2003 Nonlinear theory of geostrophic adjustment. Part 2. Two-layer and continuously stratified primitive equations. *J. Fluid Mech.* **491**, 207–228. (doi:10.1017/S0022112003005457)
70. Thomas J. 2016 Resonant fast-slow interactions and breakdown of quasi-geostrophy in rotating shallow water. *J. Fluid Mech.* **788**, 492–520. (doi:10.1017/jfm.2015.706)
71. Wagner GL, Young WR. 2015 Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**, 401–424. (doi:10.1017/jfm.2015.626)
72. Mcwilliams JC. 1984 The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.* **146**, 21–43. (doi:10.1017/S0022112084001750)
73. McWilliams JC. 1989 Statistical properties of decaying geostrophic turbulence. *J. Fluid Mech.* **198**, 199–230. (doi:10.1017/S0022112089000108)
74. Smith KS, Vallis GK. 2001 The scales and equilibration of midocean eddies: freely evolving flow. *J. Phys. Oceanogr.* **31**, 554–571. (doi:10.1175/1520-0485(2001)031<0554:TSAEOM>2.0.CO;2)
75. Smith KS, Vallis GK. 2002 The scales and equilibration of midocean eddies: forced-dissipative flow. *J. Phys. Oceanogr.* **32**, 1699–1721. (doi:10.1175/1520-0485(2002)032<1699:TSAEOM>2.0.CO;2)
76. Nadiga BT. 2014 Nonlinear evolution of a baroclinic wave and imbalanced dissipation. *J. Fluid Mech.* **756**, 965–1006. (doi:10.1017/jfm.2014.464)
77. Muraki D, Snyder C, Rotunno R. 1999 The next-order corrections to quasigeostrophic theory. *J. Atmos. Sci.* **56**, 1547–1560. (doi:10.1175/1520-0469(1999)056<1547:TNOCTQ>2.0.CO;2)
78. Rotunno R, Muraki DM, Snyder C. 2000 Unstable baroclinic waves beyond quasi-geostrophic theory. *J. Atmos. Sci.* **57**, 3285–3295. (doi:10.1175/1520-0469(2000)057<3285:UBWBQT>2.0.CO;2)
79. Hakim G, Snyder C, Muraki D. 2002 A new surface model for cyclone-anticyclone asymmetry. *J. Atmos. Sci.* **59**, 2405–2420. (doi:10.1175/1520-0469(2002)059<2405:ANSMFC>2.0.CO;2)
80. Rudnick DL. 2001 On the skewness of vorticity in the upper ocean. *Geophys. Res. Lett.* **28**, 2045–2048. (doi:10.1029/2000GL012265)
81. Shcherbina AY, D'Asaro EA, Lee C, Klymak JM, Molemaker MJ, McWilliams JC. 2013 Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. *Geophys. Res. Lett.* **40**, 4706–4711. (doi:10.1002/grl.50919)
82. Buckingham CE. 2016 Seasonality of submesoscale flows in the ocean surface boundary layer. *Geophys. Res. Lett.* **43**, 2118–2126. (doi:10.1002/2016GL068009)

83. Salmon R. 2016 Variational treatment of inertia-gravity waves interacting with a quasi-geostrophic mean flow. *J. Fluid Mech.* **809**, 502–529. (doi:10.1017/jfm.2016.693)
84. Craik ADD, Leibovich S. 1976 A rational model for Langmuir circulations. *J. Fluid Mech.* **73**, 401–426. (doi:10.1017/S0022112076001420)
85. Leibovich S. 1980 On wave-current interaction theories of Langmuir circulations. *J. Fluid Mech.* **99**, 715–724. (doi:10.1017/S0022112080000857)
86. Bühler O, McIntyre ME. 1998 On non-dissipative wave mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**, 301–343. (doi:10.1017/S002211209700774X)
87. Thomas J. 2017 *Wave-vortex interactions in rotating, stratified, and compressible flows*. PhD thesis, New York University.
88. Craik A. 1985 *Wave interactions and fluid flows*. Cambridge, UK: Cambridge University Press.
89. Ablowitz M. 2011 *Nonlinear dispersive waves-asymptotic analysis and solitons*. Cambridge, UK: Cambridge University Press.
90. Thomas J, Bühler O, Smith KS. 2018 Wave-induced mean flows in rotating shallow water with uniform potential vorticity. *J. Fluid Mech.* **839**, 408–429. (doi:10.1017/jfm.2018.22)
91. Kantha L, Tierney C. 1997 Global baroclinic tides. *Prog. Oceanogr.* **40**, 163–178. (doi:10.1016/S0079-6611(97)00028-1)
92. Taylor GI. 1919 Tidal friction in the Irish Sea. *Phil. Trans. R. Soc. Lond. A* **220**, 1–33. (doi:10.1098/rsta.1920.0001)
93. Jeffreys H. 1920 Tidal friction in shallow seas. *Phil. Trans. R. Soc. Lond. A* **221**, 239. (doi:10.1098/rsta.1921.0008)
94. Egbert GD, Ray RD. 2000 Significant tidal dissipation in the deep ocean inferred from satelite altimeter data. *Nature* **405**, 775–778. (doi:10.1038/35015531)
95. Egbert GD, Ray RD. 2001 Estimates of M2 tidal energy dissipation from TOPEX/Poseidon altimeter data. *J. Geophys. Res.* **106**, 22 475–22 502. (doi:10.1029/2000JC000699)
96. Egbert GD, Ray RD. 2003 Semi-diurnal and diurnal tidal dissipation from TOPEX/POSEIDON altimetry. *Geophys. Res. Lett.* **30**, 1907. (doi:10.1029/2003GL017676)
97. Alford MH. 2003 Energy available for ocean mixing redistributed through long-range propagation of internal waves. *Nature* **423**, 159–163. (doi:10.1038/nature01628)
98. Ray RD, Zaron ED. 2016 M2 internal tides and their observed wavenumber spectra from satellite altimetry. *J. Phys. Oceanogr.* **46**, 3–22. (doi:10.1175/JPO-D-15-0065.1)
99. Zhao Z, Alford MH, Girton J, Rainville L, Simmons H. 2016 Global observations of open-ocean mode-1 M_2 internal tides. *J. Phys. Oceanogr.* **46**, 1657–1684. (doi:10.1175/JPO-D-15-0105.1)
100. Fu L, Flierl GR. 1980 Nonlinear energy and enstrophy transfers in a realistically stratified ocean. *Dyn. Atmos. Oceans* **4**, 219–246. (doi:10.1016/0377-0265(80)90029-9)
101. Wunsch C. 1997 The vertical partition of oceanic horizontal kinetic energy and the spectrum of global variability. *J. Phys. Oceanogr.* **27**, 1770–1794. (doi:10.1175/1520-0485(1997)027<1770:TVPOOH>2.0.CO;2)
102. Rainville L, Pinkel R. 2006 Propagation of low-mode internal waves through the ocean. *J. Phys. Oceanogr.* **36**, 1220–1237. (doi:10.1175/JPO2889.1)
103. Zhao Z, Alford MH. 2009 New altimetric estimates of mode-1 M2 internal tides in the central North Pacific Ocean. *J. Phys. Oceanogr.* **39**, 1669–1684. (doi:10.1175/2009JPO3922.1)
104. Chavanne C, Gurgel KW. 2010 The surface expression of semidiurnal internal tides near a strong source at Hawaii. Part II: interactions with mesoscale currents. *J. Phys. Oceano.* **40**, 1180–1200. (doi:10.1175/2010JPO4223.1)
105. Zhao Z, Alford MH, Girton JB. 2012 Mapping low-mode internal tides from multisatellite altimetry. *Oceanography* **25**, 42–51. (doi:10.5670/oceanog.2012.40)
106. Huang X, Wang Z, Zhang Z, Yang Y, Zhou C, Tian J. 2018 Role of mesoscale eddies in modulating the semidiurnal internal tide: observation results in the northern South China Sea. *J. Phys. Oceano.* **48**, 1749–1770. (doi:10.1175/JPO-D-17-0209.1)
107. Lob J, Kohler J, Mertens C, Walter M, Li Z, von Storch JS, Zhao Z, Rhein M. 2020 Observations of the low-mode internal tide and its interaction with mesoscale flow south of the Azores. *J. Geophys. Res. Oceans* **125**, e2019JC015879. (doi:10.1029/2019JC015879)
108. Dunphy M, Lamb KG. 2014 Focusing and vertical mode scattering of the first mode internal tide by mesoscale eddy interaction. *J. Geophys. Res.* **119**, 523–536. (doi:10.1002/2013JC009293)
109. Ponte AL, Klein P. 2015 Incoherent signature of internal tides on sea level in idealized numerical simulations. *Geophys. Res. Lett.* **42**, 1520–1526. (doi:10.1002/2014GL062583)

110. Lelong MP, Riley JJ. 1991 Internal wave-vortical mode interactions in strongly stratified flows. *J. Fluid Mech.* **232**, 1–19. (doi:10.1017/S0022112091003609)
111. Majda A. 2002 *Introduction to partial differential equations and waves for the atmosphere and ocean-courant lecture notes, Bd. 9*. Providence, RI: American Mathematical Society.
112. Ward ML, Dewar W. 2010 Scattering of gravity waves by potential vorticity in a shallow-water fluid. *J. Fluid Mech.* **663**, 478–506. (doi:10.1017/S0022112010003721)
113. Savva MAC, Vanneste J. 2018 Scattering of internal tides by barotropic quasi-geostrophic flows. *J. Fluid Mech.* **856**, 504–530. (doi:10.1017/jfm.2018.694)
114. Wagner GL, Ferrando G, Young WR. 2017 An asymptotic model for the propagation of oceanic internal tides through quasi-geostrophic flow. *J. Fluid Mech.* **828**, 779–811. (doi:10.1017/jfm.2017.509)
115. Zaron ED. 2019 Baroclinic tidal sea level from exact-repeat mission altimetry. *J. Phys. Oceanogr.* **49**, 193–210. (doi:10.1175/JPO-D-18-0127.1)
116. Thomas J, Yamada R. 2019 Geophysical turbulence dominated by inertia-gravity waves. *J. Fluid Mech.* **875**, 71–100. (doi:10.1017/jfm.2019.465)
117. Alford MH. 2003 Improved global maps and 54-year history of wind-work on ocean inertial motions. *Geophys. Res. Lett.* **30**, 1424. (doi:10.1029/2002GL016614)
118. Watanabe M, Hibiya T. 2002 Global estimates of the wind-induced energy flux to inertial motions in the surface mixed layer. *Geophys. Res. Lett.* **29**, 80. (doi:10.1029/2001GL014422)
119. Jiang J, Lu Y, Perrie W. 2005 Estimating the energy flux from the wind to ocean inertial motions: the sensitivity to surface wind fields. *Geophys. Res. Lett.* **32**, L15610. (doi:10.1029/2005GL023289)
120. Garrett C. 2001 What is the 'near-inertial' band and why is it different from the rest of the internal wave spectrum?. *J. Phys. Oceanogr.* **31**, 962–971. (doi:10.1175/1520-0485(2001)031<0962:WITNIB>2.0.CO;2)
121. Young WR, Ben Jelloul M. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55**, 735–766. (doi:10.1357/0022240973224283)
122. Balmforth N, Llewellyn Smith SG, Young WR. 1998 Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. *J. Mar. Res.* **56**, 1–40. (doi:10.1357/002224098321836091)
123. Balmforth N, Young WR. 1999 Radiative damping of near-inertial oscillations. *J. Mar. Res.* **57**, 561–584. (doi:10.1357/002224099321549594)
124. Klein P, Llewellyn Smith SG. 2001 Horizontal dispersion of near-inertial oscillations in a turbulent mesoscale eddy field. *J. Mar. Res.* **59**, 697–723. (doi:10.1357/002224001762674908)
125. Klein P, Llewellyn Smith SG, Lapeyre G. 2004 Organization of near-inertial energy by an eddy field. *Q.J.R. Meteorol. Soc.* **130**, 1153–1166. (doi:10.1256/qj.02.231)
126. Danioux E, Vanneste J, Bühler O. 2015 On the concentration of near-inertial waves in anticyclones. *J. Fluid Mech.* **773**, R2. (doi:10.1017/jfm.2015.252)
127. Danioux E, Vanneste J. 2016 Near-inertial wave scattering by random flows. *Phys. Rev. Fluids* **1**, 033701. (doi:10.1103/PhysRevFluids.1.033701)
128. Thomas J, Smith K, Bühler O. 2017 Near-inertial wave dispersion by geostrophic flows. *J. Fluid Mech.* **817**, 406–438. (doi:10.1017/jfm.2017.124)
129. Asselin O, Thomas LN, Young WR, Rainville L. 2020 Refraction and straining of near-inertial waves by barotropic eddies. *J. Phys. Oceanogr.* **50**, 3439–3454. (doi:10.1175/JPO-D-20-0109.1)
130. Asselin O, Young WR. 2020 Penetration of wind-generated near-inertial waves into a turbulent ocean. *J. Phys. Oceanogr.* **50**, 1699–1716. (doi:10.1175/JPO-D-19-0319.1)
131. Yu X, Naveira Garabato AC, Vic C, Gula J, Savage AC, Wang J, Waterhouse AF, MacKinnon JA. 2022 Observed equatorward propagation and chimney effect of near-inertial waves in the midlatitude ocean. *Geophys. Res. Lett.* **49**, e2022GL098522. (doi:10.1002/essoar.10510719.1)
132. Vic C, Ferron B, Thierry V, Mercier H, Lherminier P. 2021 Tidal and near-inertial internal waves over the Reykjanes Ridge. *J. Phys. Oceanogr.* **51**, 419–437. (doi:10.1175/JPO-D-20-0097.1)
133. Lee DK, Niiler PP. 1998 The inertial chimney: the near-inertial energy drainage from the ocean surface to the deep layer. *J. Geophys. Res.* **103**, 7579–7591. (doi:10.1029/97JC03200)
134. Elipot S, Lumpkin R, Prieto G. 2010 Modification of inertial oscillations by the mesoscale eddy field. *J. Geophys. Res.* **115**, C09010. (doi:10.1029/2009JC005679)
135. Joyce T, Toole J, Klein P, Thomas L. 2013 A near-inertial mode observed within a Gulf Stream warm-core ring. *J. Geophys. Res.* **118**, 1797–1806. (doi:10.1002/jgrc.20141)

136. Sanford TB, Ma BB, Alford MH. 2021 Stalling and dissipation of a near-inertial wave (NIW) in an anticyclonic ocean eddy: direct determination of group velocity and comparison with theory. *JGR. Oce.* **126**, e2020JC016742. (doi:10.1002/essoar.10505701.1)
137. Xie J, Vanneste J. 2015 A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143–169. (doi:10.1017/jfm.2015.251)
138. Wagner GL, Young WR. 2016 A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic. *J. Fluid Mech.* **802**, 806–837. (doi:10.1017/jfm.2016.487)
139. Rocha CB, Wagner GL, Young WR. 2018 Stimulated generation-extraction of energy from balanced flow by near-inertial waves. *J. Fluid Mech.* **847**, 417–451. (doi:10.1017/jfm.2018.308)
140. Thomas J, Arun S. 2020 Near-inertial waves and geostrophic turbulence. *Phys. Rev. Fluids* **5**, 014801. (doi:10.1103/PhysRevFluids.5.014801)
141. Gertz A, Straub DN. 2009 Near-inertial oscillations and the damping of midlatitude gyres: a modelling study. *J. Phys. Oceanogr.* **39**, 2338–2350. (doi:10.1175/2009JPO4058.1)
142. Thomas J, Daniel D. 2020 Turbulent exchanges between near-inertial waves and balanced flows. *J. Fluid Mech.* **902**, A7. (doi:10.1017/jfm.2020.510)
143. Alford MH. 2001 Internal swell generation: the spatial distribution of energy flux from the wind to mixed layer near-inertial motions. *J. Phys. Oceanogr.* **31**, 2359–2368. (doi:10.1175/1520-0485(2001)031<2359:ISGTSD>2.0.CO;2)
144. Alford MH, Whitmont M. 2007 Seasonal and spatial variability of near-inertial kinetic energy from historical moored velocity records. *J. Phys. Oceanogr.* **37**, 2022–2037. (doi:10.1175/JPO3106.1)
145. Chaigneau A, Pizarro O, Rojas W. 2008 Global climatology of near-inertial current characteristics from Lagrangian observations. *Geophys. Res. Lett.* **35**, L13603. (doi:10.1029/2008GL034060)
146. Silverthorne KE, Toole JM. 2009 Seasonal kinetic energy variability of near-inertial motions. *J. Phys. Oceanogr.* **39**, 1035–1049. (doi:10.1175/2008JPO3920.1)
147. D'Asaro E. 1985 Upper ocean temperature structure, inertial currents, and richardson numbers observed during strong meteorological forcing. *J. Phys. Oceanogr.* **15**, 943–962. (doi:10.1175/1520-0485(1985)015<0943:UOTSIC>2.0.CO;2)
148. Whalen CA, MacKinnon JA, Talley LD. 2018 Large-scale impacts of the mesoscale environment on mixing from wind-driven internal waves. *Nat. Geo. Sci.* **11**, 842–847. (doi:10.1038/s41561-018-0213-6)
149. Taylor S, Straub D. 2016 Forced near-inertial motion and dissipation of low-frequency kinetic energy in a wind-driven channel flow. *J. Phys. Oceanogr.* **46**, 79–93. (doi:10.1175/JPO-D-15-0060.1)
150. Taylor S, Straub D. 2020 Effects of adding forced near-inertial motion to a wind-driven channel flow. *J. Phys. Oceanogr.* **50**, 2983–2996. (doi:10.1175/JPO-D-19-0299.1)
151. Garrett CJR, Munk WH. 1972 Space-timescales of internal waves. *Geophys. Fluid Dyn.* **2**, 225–264. (doi:10.1080/03091927208236082)
152. Garrett CJR, Munk WH. 1975 Space-timescales of internal waves. A progress report. *J. Geophys. Res.* **80**, 291–297. (doi:10.1029/JC080i003p00291)
153. Garrett C, Munk W. 1979 Internal waves in the ocean. *Annu. Rev. Fluid Mech.* **11**, 339–369. (doi:10.1146/annurev.fl.11.010179.002011)
154. Muller P, Olbers DJ. 1975 On the dynamics of internal waves in the deep ocean. *J. Geophys. Res.* **80**, 3848–3860. (doi:10.1029/JC080i027p03848)
155. Olbers DJ. 1976 Nonlinear energy transfer and the energy balance of the internal wave field in the deep ocean. *J. Fluid Mech.* **74**, 375–399. (doi:10.1017/S0022112076001857)
156. McComas CH, Bretherton FP. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **82**, 1397–1412. (doi:10.1029/JC082i009p01397)
157. Muller P, Holloway G, Henyey F, Pomphrey N. 1986 Nonlinear interactions among internal gravity waves. *Rev. Geophys.* **24**, 493–536. (doi:10.1029/RG024i003p00493)
158. Winters KB, D'Asaro EA. 1997 Direct simulation of internal wave energy transfer. *J. Phys. Oceanogr.* **27**, 1937–1945. (doi:10.1175/1520-0485(1997)027<1937:DSOIWE>2.0.CO;2)
159. Lvov Y, Tabak E. 2001 Hamiltonian formalism and the Garrett-Munk spectrum of internal waves in the ocean. *Phys. Rev. Lett.* **87**, 168501. (doi:10.1103/PhysRevLett.87.168501)
160. Lvov Y, Polzin K, Tabak E. 2004 Energy spectra of the ocean's internal wave field: theory and observations. *Phys. Rev. Lett.* **92**, 128501. (doi:10.1103/PhysRevLett.92.128501)

161. Lvov YV, Polzin KL, Yokoyama N. 2012 Resonant and near-resonant internal wave interactions. *J. Phys. Oceanogr.* **42**, 669–691. (doi:10.1175/2011JPO4129.1)
162. Barkan R, Winters KB, McWilliams JC. 2017 Stimulated imbalance and the enhancement of eddy kinetic energy dissipation by internal waves. *J. Phys. Oceanogr.* **47**, 181–198. (doi:10.1175/JPO-D-16-0117.1)
163. Eden C, Olbers D. 2014 An energy compartment model for propagation, nonlinear interaction, and dissipation of internal gravity waves. *J. Phys. Oceanogr.* **44**, 2093–2106. (doi:10.1175/JPO-D-13-0224.1)
164. Eden C, Pollmann F, Olbers D. 2019 Numerical evaluation of energy transfers in internal gravity wave spectra of the ocean. *J. Phys. Oceanogr.* **49**, 737–749. (doi:10.1175/JPO-D-18-0075.1)
165. Eden C, Pollmann F, Olbers D. 2020 Towards a global spectral energy budget for internal gravity waves in the ocean. *J. Phys. Oceanogr.* **50**, 935–944. (doi:10.1175/JPO-D-19-0022.1)
166. Pan Y, Arbic BK, Nelson AD, Menemenlis D, Peltier WR, Xu W, Li Y. 2020 Numerical investigation of mechanisms underlying oceanic internal gravity wave power-law spectra. *J. Phys. Oceanogr.* **50**, 2713–2733. (doi:10.1175/JPO-D-20-0039.1)
167. Polzin KL, Lvov Y. 2011 Toward regional characterizations of the oceanic internal wavefield. *Rev. Geophys.* **49**, RG4003. (doi:10.1029/2010RG000329)
168. Pollmann F. 2020 Global characterization of the ocean's internal wave spectrum. *J. Phys. Oceanogr.* **50**, 1871–1891. (doi:10.1175/JPO-D-19-0185.1)
169. Le Boyer A, Alford MH. 2021 Variability and sources of the internal wave continuum examined from global moored velocity records. *J. Phys. Oceanogr.* **51**, 2807–2823. (doi:10.1175/JPO-D-20-0155.1)
170. Muller M, Arbic BK, Richman JG, Shriver JF, Kunze EL, Scott RB, Wallcraft AJ, Zamudio L. 2015 Toward an internal gravity wave spectrum in global ocean models. *Geophys. Res. Lett.* **42**, 3474–3481. (doi:10.1002/2015GL063365)
171. Rocha CB, Chereskin TK, Gille ST, Menemenlis D. 2016 Mesoscale to submesoscale wavenumber spectra in drake passage. *J. Phys. Oceanogr.* **46**, 601–620. (doi:10.1175/JPO-D-15-0087.1)
172. Thomas J, Daniel D. 2021 Forward flux and enhanced dissipation of geostrophic balanced energy. *J. Fluid Mech.* **911**, A60. (doi:10.1017/jfm.2020.1026)
173. Wunsch C. 2015 *Modern observational physical oceanography: understanding the global ocean*. Princeton, NJ: Princeton University Press.
174. Batchelor GK. 1982 *The theory of homogeneous turbulence*. Cambridge, UK: Cambridge University Press.
175. Brown ED, Owens WB. 1981 Observations of the horizontal interactions between the internal wave field and the mesoscale flow. *J. Phys. Oceanogr.* **11**, 1474–1480. (doi:10.1175/1520-0485(1981)011<1474:OOTHIB>2.0.CO;2)
176. Qi L, Zhaohui C, Shoude G, Haiyuan Y, Zhao J, Yongzheng L, Bingrong S, Lixin W. 2022 Enhanced near-inertial waves and turbulent diapycnal mixing observed in a cold- and warm-core eddy in the kuroshio extension region. *J. Phys. Oceanogr.* **31**, 962–971.
177. Jing Z, Chang P, DiMarco SF, Wu L. 2018 Observed energy exchange between low-frequency flows and internal waves in the Gulf of Mexico. *J. Phys. Oceano.* **48**, 995–1008. (doi:10.1175/JPO-D-17-0263.1)
178. Cusack JM, Brearley J, Naveira Garabato AC, Smeed DA, Polzin KL, Velzeboer N, Shakespeare CJ. 2020 Observed eddy-internal wave interactions in the Southern Ocean. *J. Phys. Oceanogr.* **50**, 3043–3062. (doi:10.1175/JPO-D-20-0001.1)
179. Garrett CJR. 1968 On the interaction between internal gravity waves and a shear flow. *J. Fluid Mech.* **34**, 711. (doi:10.1017/S0022112068002181)
180. Jones WL. 1969 Ray tracing for internal gravity waves. *J. Geophys. Res.* **74**, 2028–2033. (doi:10.1029/JB074i008p02028)
181. Kunze E. 1985 Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.* **15**, 544–565. (doi:10.1175/1520-0485(1985)015<0544:NIWPIG>2.0.CO;2)
182. D'Asaro E, Eriksen C, Levine M, Niiler P, Paulson C, van Meurs P. 1995 Upper ocean inertial currents forced by a strong storm. Part I: data and comparisons with linear theory. *J. Phys. Oceanogr.* **25**, 2909–2936. (doi:10.1175/1520-0485(1995)025<2909:UOICFB>2.0.CO;2)
183. Lumpkin R, Elipot S. 2010 Surface drifter pair spreading in the North Atlantic. *J. Geophys. Res.* **115**, C1207. (doi:10.1029/2010JC006338)

184. Xu Y, Fu LL. 2011 Global variability of the wavenumber spectrum of oceanic mesoscale turbulence. *J. Phys. Oceanogr.* **41**, 802–809. (doi:10.1175/2010JPO4558.1)
185. Capet X, McWilliams JC, Molemaker MJ, Shchepetkin AF. 2008 Mesoscale to submesoscale transition in the California Current system: flow structure, eddy flux, and observational tests. *J. Phys. Oceanogr.* **38**, 29–43. (doi:10.1175/2007JPO3671.1)
186. Capet X, McWilliams JC, Molemaker MJ, Shchepetkin AF. 2008 Mesoscale to submesoscale transition in the California Current system: energy balance and flux. *J. Phys. Oceanogr.* **38**, 2256–2269. (doi:10.1175/2008JPO3810.1)
187. Qiu B, Chen S, Klein P, Sasaki H, Sasai Y. 2014 Seasonal mesoscale and submesoscale eddy variability along the north pacific subtropical countercurrent. *J. Phys. Oceanogr.* **44**, 3079–3098. (doi:10.1175/JPO-D-14-0071.1)
188. Brannigan L, Marshall DP, Naveira-Garabato A, Nurser AG. 2015 The seasonal cycle of submesoscale flows. *Ocean Model* **92**, 69–84. (doi:10.1016/j.ocemod.2015.05.002)
189. Thompson AF, Lazar A, Buckingham CE, Naveira Garabato AC, Damerell GM, Heywood KJ. 2016 Open-ocean submesoscale motions: a full seasonal cycle of mixed layer instabilities from gliders. *J. Phys. Oceanogr.* **46**, 1285–1307. (doi:10.1175/JPO-D-15-0170.1)
190. Poje AC, Ozgokmen TM, Bogucki DJ, Kirwan AD. 2017 Evidence of a forward energy cascade and Kolmogorov self-similarity in submesoscale ocean surface drifter observations. *Phys Fluids* **29**, 020701. (doi:10.1063/1.4974331)
191. Yu X, Naveira Garabato AC, Martin AP, Buckingham CE, Brannigan L, Su Z. 2019 An annual cycle of submesoscale vertical flow and restratification in the upper ocean. *J. Phys. Oceanogr.* **46**, 1439–1461. (doi:10.1175/JPO-D-18-0253.1)
192. Erickson ZK, Thompson AF, Callies J, Yu X, Naveira Garabato A, Klein P. 2020 The vertical structure of open-ocean submesoscale variability during a full seasonal cycle. *J. Phys. Oceanogr.* **50**, 145–160. (doi:10.1175/JPO-D-19-0030.1)
193. Callies J, Ferrari R, Klymak J, Gula J. 2015 Seasonality in submesoscale turbulence. *Nat. Commun.* **6**, 6862. (doi:10.1038/ncomms7862)
194. Thomas J, Vishnu R. 2022 Turbulent transition of a flow from small to O(1) Rossby numbers. *J. Phys. Oceanogr.* **52**, 2609–2625. (doi:10.1175/JPO-D-21-0270.1)
195. Balwada D, Xie J, Marino R, Feraco F. 2022 Direct observational evidence of an oceanic dual kinetic energy cascade and its seasonality. *Sci. Adv.* **8**, eabq2566. (doi:10.1126/sciadv.abq2566)
196. Thomas LN, Tandon A, Mahadevan A. 2008 Submesoscale processes and dynamics. Ocean modeling in an eddying regime. *Geophys. Monogr.* **Ser.177**, 17–38. (doi:10.1029/177GM04)
197. McWilliams JC. 2016 Submesoscale currents in the ocean. *Proc. R. Soc. A* **472**, 20160177. (doi:10.1098/rspa.2016.0117)
198. Taylor JR, Thompson AF. 2023 Submesoscale dynamics in the upper ocean. *Annu. Rev. Fluid Mech.* **55**, 103–127. (doi:10.1146/annurev-fluid-031422-095147)
199. Thomas LN. 2017 On the modifications of near-inertial waves at fronts: implications for energy transfer across scales. *Ocean Dyn.* **67**, 1335–1350. (doi:10.1007/s10236-017-1088-6)
200. Barkan R, Srinivasan K, Yang L, McWilliams JC, Gula J, Vic C. 2021 Oceanic mesoscale eddy depletion catalyzed by internal waves. *Geophys. Res. Lett.* **48**, 18. (doi:10.1029/2021GL094376)
201. Nagai TA, Tandon A, Kunze E, Mahadevan A. 2015 Spontaneous generation of near-inertial waves by the Kuroshio front. *J. Phys. Oceanogr.* **45**, 2381–2406. (doi:10.1175/JPO-D-14-0086.1)
202. Shakespeare CJ, Hogg AM. 2017 Spontaneous surface generation and interior amplification of internal waves in a regional-scale ocean model. *J. Phys. Oceanogr.* **47**, 811–826. (doi:10.1175/JPO-D-16-0188.1)
203. Shakespeare CJ, Hogg AM. 2018 The life cycle of spontaneously generated internal waves. *J. Phys. Oceanogr.* **48**, 343–359. (doi:10.1175/JPO-D-17-0153.1)
204. Evans DG, Frajka-Williams E, Naveira Garabato AC, Polzin KL, Forryan A. 2020 Mesoscale eddy dissipation by a 'Zoo' of submesoscale processes at a western boundary. *J. Geophys. Res. Oceans* **125**, e2020JC016246. (doi:10.1029/2020JC016246)
205. Hogg A, Dewar WK, Berloff P, Ward ML. 2011 Kelvin wave hydraulic control induced by interactions between vortices and topography. *J. Fluid Mech.* **687**, 194–208. (doi:10.1017/jfm.2011.344)
206. Nikurashin M, Ferrari R. 2010 Radiation and dissipation of internal waves generated by geostrophic motions impinging on small-scale topography: theory. *J. Phys. Ocean.* **40**, 1055–1074. (doi:10.1175/2009JPO4199.1)
207. Wu Y, Kunze E, Tandon A, Mahadevan A. 2009 Reabsorption of Lee-Eave energy in bottom-intensified currents. *J. Phys. Oceanogr.* **53**, 477–491. (doi:10.1175/JPO-D-22-0058.1)

208. Rai S, Hecht M, Maltrud M, Aluie H. 2021 Scale of oceanic eddy killing by wind from global satellite observations. *Sci. Adv.* **7**, eabf4920. (doi:10.1126/sciadv.abf4920)
209. Xu C, Zhai X, Shang X. 2016 Work done by atmospheric winds on mesoscale ocean eddies. *Geophys. Res. Lett.* **43**, 174–180. (doi:10.1002/2016GL071275)
210. Duhaut TH, Straub DN. 2006 Wind stress dependence on ocean surface velocity: implications for mechanical energy input to ocean circulation. *J. Phys. Oceanogr.* **36**, 202–211. (doi:10.1175/JPO2842.1)
211. Jansen M, Adcroft A, Khani S, Kong H. 2019 Towards an energetically consistent, resolution aware parameterization of ocean mesoscale eddies. *J. Adv. Model. Earth* **11**, 2844–2860. (doi:10.1029/2019MS001750)
212. Porta Mana PGL, Zanna L. 2014 Toward a stochastic parameterization of ocean mesoscale eddies. *J. Phys. Oceanogr.* **79**, 1–20. (doi:10.1016/j.ocemod.2014.04.002)
213. Tandon A, Garrett C. 1996 On a recent parameterization of mesoscale eddies. *J. Phys. Oceanogr.* **26**, 406–411. (doi:10.1175/1520-0485(1996)026<0406:OARPOM>2.0.CO;2)
214. Gent PR, Willebrand J, McDougall TJ, McWilliams JC. 1995 Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.* **25**, 463–474. (doi:10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2)
215. Gent PR, McWilliams JC. 1990 Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.* **20**, 150–155. (doi:10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2)
216. Spiro Jaeger G, MacKinnon JA, Lucas AJ, Shroyer E, Nash J, Tandon A, Farrar JT, Mahadevan A. 2020 How spice is stirred in the Bay of Bengal. *J. Phys. Oceanogr.* **50**, 2669–2688.
217. Kunze E, Klymak JM, Lien RC, Ferrari R, Lee CM, Sundermeyer MA, Goodman L. 2015 Submesoscale water-mass spectra in the sargasso sea. *J. Phys. Oceanogr.* **45**, 1325–1338. (doi:10.1175/JPO-D-14-0108.1)
218. Klymak JM, Crawford W, Alford MH, MacKinnon JA, Pinkel R. 2015 Along-isopycnal variability of spice in the North Pacific. *J. Geophys. Res. Oceans* **120**, 2287–2307. (doi:10.1002/2013JC009421)
219. Callies J, Ferrari R. 2013 Interpreting energy and tracer spectra of upper-ocean turbulence in the submesoscale range (1–200 km). *J. Phys. Oceanogr.* **43**, 2456–2474. (doi:10.1175/JPO-D-13-063.1)
220. Cole S, Rudnick D. 2012 The spatial distribution and annual cycle of upper ocean thermohaline structure. *J. Geophys. Res.* **117**, C02027. (doi:10.1029/2011JC007033)
221. Cole S, Rudnick D, Colosi J. 2010 Seasonal evolution of upper-ocean horizontal structure and the remnant mixed layer. *J. Geophys. Res.* **115**, C04012. (doi:10.1029/2009JC005654)
222. Ferrari R, Rudnick D. 2000 Thermohaline variability in the upper ocean. *J. Geophys. Res.* **105**, 16 857–16 884. (doi:10.1029/2000JC900057)
223. Suanda SH, Feddersen F, Spydell MS, Kumar N. 2018 The effect of barotropic and baroclinic tides on three-dimensional coastal dispersion. *Geophy. Res. Lett.* **45**, 11 235–11 246. (doi:10.1029/2018GL079884)
224. Meyerjurgens J, Ricker M, Schakau V, Badewien TH, Stanev EV. 2020 Relative dispersion of surface drifters in the North Sea: the effect of tides on mesoscale diffusivity. *J. Geophys. Res.-Ocean* **124**, e2019JC015925.
225. Thomas J, Gupta A. 2022 Wave-enhanced tracer dispersion. *JGR Oceans* **127**, e2020JC017005. (doi:10.1029/2020JC017005)
226. Hernandez-Duenas G, Lelong MP, Smith LM. 2021 Impact of wave-vortical interactions on oceanic submesoscale lateral dispersion. *J. Phys. Oceanogr.* **51**, 3495–3511. (doi:10.1175/JPO-D-20-0299.1)
227. Remmel M, Smith L. 2009 New intermediate models for rotating shallow water and an investigation of the preference for anticyclones. *J. Fluid Mech.* **635**, 321–359. (doi:10.1017/S0022112009007897)
228. Brunner-Suzuki AMEG, Sundermeyer M, Lelong MP. 2012 Vortex stability in a large-scale internal wave shear. *J. Phys. Oceanogr.* **42**, 1668–1683. (doi:10.1175/JPO-D-11-0137.1)
229. Brunner-Suzuki AMEG, Sundermeyer M, Lelong MP. 2014 Upscale energy transfer by the vortical mode and internal waves. *J. Phys. Oceanogr.* **44**, 2446–2469. (doi:10.1175/JPO-D-12-0149.1)
230. Hernandez-Duenas G, Smith L, Stechmann S. 2014 Investigation of Boussinesq dynamics using intermediate models based on wave–vortical interactions. *J. Fluid Mech.* **747**, 247–287. (doi:10.1017/jfm.2014.138)

231. Haine T, Marshall J. 1998 Gravitational, symmetric and baroclinic instability of the ocean mixed layer. *J. Phys. Oceanogr.* **28**, 634–658. (doi:10.1175/1520-0485(1998)028<0634:GSABIO>2.0.CO;2)
232. Stone PH. 1966 On non-geostrophic baroclinic stability. *J. Atmos. Sci.* **23**, 390–400. (doi:10.1175/1520-0469(1966)023<0390:ONGBS>2.0.CO;2)
233. Boccaletti G, Ferrari R, Fox-Kemper B. 2007 Mixed layer instabilities and restratification. *J. Phys. Oceanogr.* **37**, 2228–2250. (doi:10.1175/JPO3101.1)
234. Stamper MA, Taylor JR. 2016 The transition from symmetric to baroclinic instability in the Eady model. *Ocean Dyn.* **67**, 65–80. (doi:10.1007/s10236-016-1011-6)
235. Sakai S. 1989 Rossby-Kelvin instability: a new type of ageostrophic instability caused by a resonance between Rossby waves and gravity waves. *J. Fluid Mech.* **202**, 149–176. (doi:10.1017/S0022112089001138)