The distribution of eddy kinetic and potential energies in the global ocean

By RAFFAELE FERRARI and CARL WUNSCH*, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

(Manuscript received 21 November 2009; in final form 17 December 2009)

ABSTRACT
Understanding of the major sources, sinks, and reservoirs of energy in the ocean is briefly updated in a diagram. The nature of the dominant kinetic energy reservoir, that of the balanced variability, is then found to be indistinguishable in the observations from a sum of barotropic and first baroclinic ordinary quasi-geostrophic modes. Little supporting evidence is available to partition the spectra among forced motions and turbulent cascades, along with significant energy more consistent with weakly non-linear wave dynamics. Linear-response wind-forced motions appear to dominate the high frequency (but subinertial) mooring frequency spectra. Turbulent cascades appear to fill the high wavenumber spectra in altimetric data and numerical simulations. Progress on these issues is hindered by the difficulty in connecting the comparatively easily available frequency spectra with the variety of theoretically predicted wavenumber spectra.

1. Introduction
The quantitative description of the sources, sinks and reservoirs of energy in the ocean has emerged in recent years as a focus of attention both because it constitutes a piece of fundamental understanding, and because it has consequences for determining how the system can change. Description and physical understanding are dependent upon both observations and theories of oceanic motions; given the complexity and huge range of time and space scales involved in oceanic flows, the challenge of a full description is a formidable one. The energetics of the circulation have traditionally been given little attention in physical oceanography presumably because, as in much of fluid dynamics, a knowledge of the flow field is diagnostic of the energy, whereas the energy is not diagnostic of the fluid flow except in a gross sense. Kinetic energy is here given a special status because it is intimately connected to the movement of water, and hence directly to the general circulation. The energetics of the circulation have traditionally been given little attention in physical oceanography presumably because, as in much of fluid dynamics, a knowledge of the flow field is diagnostic of the energy, whereas the energy is not diagnostic of the fluid flow except in a gross sense. Kinetic energy is here given a special status because it is intimately connected to the movement of water, and hence directly to the general circulation.

This paper begins with a recapitulation and update of the overall energy budget described by Wunsch and Ferrari (2004) and Ferrari and Wunsch (2009, supplemental material). As will be seen, much remains obscure, especially in the transfer of energy to and from the various reservoirs, but one of the more troubling problems concerns the largest kinetic energy reservoir—that of the geostrophically balanced flows. Thus in the remainder of the paper, we attempt a discussion of the theory and observations pertaining to its understanding.

2. The energy reservoirs
As discussed in textbooks (e.g. Gill, 1982; Vallis, 2006) fluids contain three distinct energy types: kinetic (KE), potential (PE), and internal (IE). Each represents a different set of problems in making an appropriate estimate of its value, and in the interpretation of that value. Dynamic coupling means that a change in the value of any one of them generally implies an eventual change in one or both of the others.

For many purposes, the kinetic energy becomes the focus of attention because it represents most directly the fluid flow, its transport properties and the mixing rates of the system (see Ferrari and Wunsch, 2009). Although there is some ambiguity of definition (in terms of separation of mean and anomaly contributions), interpretation of KE is comparatively straightforward. Both PE and IE involve a considerable degree of arbitrariness—the defining geopotential for the former, and reference temperature for the latter, and changes in their values are physically more useful than are their absolute values, however defined. To further complicate their discussion, much published attention is focussed on how much potential and internal energies are available in the sense that they can be converted through specifically defined processes, into kinetic energy, and how much is necessarily passive.

Figure 1 shows an updated schematic of oceanic energy reservoirs, sources, sinks and transfer routes. The ocean is divided into a near-surface energetic mixed-layer, an upper ocean (roughly
Fig. 1. A representation of the major energy reservoirs, their sources, and interchanges. An earlier version of this diagram appeared in the supplemental material of Ferrari and Wunsch (2009). See the text for discussion. Thin dashed arrows all imply Joule heating. Heavy dashed arrow indicates the reservoir whose structure is discussed in this paper.

100–1500 m) and an abyssal ocean below about 1500 m to the bottom. Internal-inertial waves, including internal tides, are displayed separately although they are part of the overall oceanic energy reservoirs. Green denotes primarily ageostrophic motions, blue those that are dominantly in geostrophic balance. Transfers are in terawatts (TW = 10¹²W), and reservoir values are in exajoules (EJ = 10¹⁸J) except for the time-mean general circulation which is in yottajoules (YJ = 10²⁴J). The latter reservoir is almost entirely potential energy and the value is not very meaningful except that it is very large compared to anything else. Its value does become an issue when discussing qualitative, climate-scale, general circulation shifts that would increase or decrease it (e.g. in homogenizing the whole water column). The boxed area on lower left is a separate recapitulation of the general circulation between time mean and variability. As far as practical, values have been inferred from direct observations rather than from models—for whose energy cycles many other questions remain.

Figure 1 is necessarily incomplete and somewhat inconsistent. The near-surface-mixed layer variability includes a geostrophic component as well as a strongly ageostrophic one, and the latter is present everywhere as well. Loss of balance in the abyssal ocean is thought to be another source of internal waves, but the details and rates remain obscure. Ultimately, all dissipation is in
the form of Joule heating. Where no value is shown, no plausible estimate was found. A value of 0 implies something with magnitude of less than about 0.01TW (some may be negative). With the exception of the total tidal dissipation, all these numbers probably have uncertainties of a factor of 2, and in some cases factor of 10 errors are conceivable. True numerical balances have not been attempted, and most numbers are mainly invitations to the reader to provide better ones.

Letters (a), (b), . . . , (A), (B), . . . denote the reference list in the box. References must be consulted for what is a tangled and uncertain tale. Thorpe (2005, appendix 4) shows comparable values for some of the components here as well as some regional values.

References for power numbers (TW) in Fig. 1: (a) Lunisolar tidal dissipation in the ocean. Dickey et al. (1994); Munk (1997). (b) Conversion to internal tides and shallow water dissipation; primarily semi-diurnal; Egbert and Ray (2003). (c) Wind generation of inertial waves and transfer to upper ocean. Alford and Whitmont (2007), based on an approximate mean of 0.5 × 10^{-3} W m^{-2} and is all kinetic energy; cf. Furuichi et al. (2008). (d) Wind work rate on sea surface Wang and Alford (2004); Rascle et al. (2008). (e) and (f) Work rate from evaporation, precipitation, total heat flux. Huang (2004). This component of possible energy transfer has given rise to a contentious and increasingly obscure literature; see Winters and Young (2009) for entry. (g) Pressure work rate. Ponte (2009) from ECCO-GODAE state estimate v2.216. Wang et al. (2006) estimate is 10 times larger. (h) Work on the Ekman layer. Wang and Huang (2004). (j) Rate of work on geostrophic flow. Wunsch (1998); von Storch et al. (2007) value reduced by 25% according to Hughes and Wilson (2008). Compare to Scott and Xu (2008). von Storch et al. (2007) suggest the sum of the Ekman and geostrophic powers are 3.8TW, but when reduced by 25% is indistinguishable from values here. (k) Power input to surface waves; Rascle et al. (2008). (m) Flux onto beaches. R. Flick (personal communication, 2007), based on the assumption of a m average significant wave height and an exposed coastline of 600,000 km. See also Beyene and Wilson (2006), as the issue is important for deriving energy from waves. (n) Transfer from mixed layer to ocean below. von Storch et al. (2007) reduced by 25% for the eddy-effect and so fortuitously the same as Wunsch (1998) value. (o) Biomixing. Dewar et al. (2006); see also Gregg and Horne (2009), Katija and Dabiri (2009). (p) Geothermal energy input. Huang (2004). (q) Sen et al. (2008). Estimated from numbers of Hughes and Wilson (2008). (s) Baroclinic instability. WF2004; Huang (2004) estimate is 1.1TW; D. Ferreira (personal communication, 2008) 0.3TW. (t) Marshall and Naveira Garabato (2008), Southern Ocean alone, with their μ ≤ 10^{-3} m^{-2} s^{-1} over 1000 m abyssal depth, an area of 6 × 10^{13} m^{2} and mixing efficiency, Γ = 0.2. Watson (1985) also suggests values near 1TW but from a different mechanism and numbers were not added here. (u) Wind work against eddies extrapolated from results of Duhaut and Straub (2006) and others as described by Ferrari and Wunsch (2009). (v) Western boundaries in particular are intended (D. Marshall, private communication, 2009). (w) Williams et al. (2008) suggest as much as 1.5TW could be involved.

Reservoir values (EJ or YJ): (A) Internal waves. Munk (1981). How much of this includes inertial waves is not clear. (B) Inertial waves assigned 50% of the internal wave energy based on North Atlantic current metre records. (C) Model results; B. Arbic (personal communication, 2008). (D) Time mean general circulation. Oort et al. (1989). Essentially all potential energy, and not especially meaningful. Determining how much is “available” is a complex, somewhat ambiguous, undertaking not carried out here, but they estimate about 200EJ are available potential energy; see Winters and Young (2009). (E) Geostrophic variability. Zang and Wunsch (2001); D. Menemenlis (personal communication, 2007, from ECCO2 model); Wunsch (1998). (F) Very uncertain. See (P. 211 Thorpe, 2005). Vortical modes, exceptionally, are thought to represent at least in part a transfer of internal wave energy into geostrophic motions.

3. Balanced kinetic energy

Despite many remaining quantitative uncertainties, a zero-order understanding exists of most of the major energy reservoirs in the ocean, including generation and dissipation of internal waves, the generation of large-scale potential energy by winds and buoyancy, etc. Some progress has been made in understanding the major kinetic energy reservoir—that of the balanced flows (geostrophic eddies; see Ferrari and Wunsch, 2009), but its maintenance and dissipation mechanisms and rates remain poorly understood. The problem of the description of the motions contained there is now discussed.

Over the past 20 yr, satellite altimetry has emerged as the central data set for describing and understanding the distribution and controls on oceanic kinetic energy: the data are global in scope and, almost uniquely, provide measurements of the spatial structures (e.g. Wunsch and Stammer, 1998; Fu and Cazenave, 2000). The broad spatial resolution is crucial, because theories of energy redistribution operate in wavenumber space.

One major issue (among others) confronts anyone using altimetric data for the study of oceanic kinetic energy distributions: The measurement represents surface pressure distributions, and these arise from a large number of differing physical processes. As with much physics generally, separation of different processes is commonly most straightforwardly done by time scale. Thus a surface disturbance with a spatial scale of 200 km can arise from balanced motions having time scales of weeks to months and longer, or from internal waves having time scales
of hours to a day or two. But altimeters, with their orbital constraints producing repeat times of days to weeks, are ill-suited to producing data with the high frequency sampling necessary to distinguish these two radically different physics and to avoid aliasing.

An additional problem arises because the signal to noise ratio in altimetric data is a strong function of wavenumber, apparently falling sharply at scales shorter than about 200 km (see Stammer, 1997, Fig. 8) and despite strenuous efforts to extract the signal from the noise (Scott and Wang, 2005), the results remain not completely convincing, or at best spatially and temporally sporadic during intervals and regions of larger signal-to-noise ratios.

That internal waves are visible in the altimeter data is made most concrete by the widespread observations (e.g. Egbert and Ray, 2000) of first mode internal tides. (Because of their fixed narrow-band spectral properties, the aliasing of the 12.42 hr M2 tide into an apparent period near 60 d renders the sampling problem much simpler than with the broad-band processes characterizing other types of internal waves. Ferrari and Wunsch (2009) display some representative frequency spectra including the internal wave band.) Wavenumber spectra at spatial scales of order 200 km have generally been interpreted (e.g. Katz, 1975) as being wholly attributable to internal waves as described by the Garrett and Munk spectrum (Munk, 1981). In contrast, Scott and Wang (2005) have interpreted the motions seen in altimeters on these scales as being wholly provided by balanced motions. What seems reasonable is that both motions are present, although in what proportions remains unknown. Any major reduction in the variance attributed to the GM spectrum would, however, raise apparent conflicts with a multitude of other data types; some discussion of the spectral shapes is therefore taken up below.

Motions seen in altimeters on scales (wavelengths) greater than 200–300 km, are unlikely to be dominated either by internal waves or their low wavenumber aliases. The usual interpretation has relied on results from moorings (Wunsch, 1997) that showed, crudely speaking, 40% of the kinetic energy not in the internal wave band was barotropic in nature, with about another 40% lying in the first baroclinic mode (with ‘barotropic’, ‘baroclinic’ and ‘mode’ being used as in the flat-bottom, linear, otherwise resting, ocean theory; see Gill, 1982). Because the buoyancy frequency, \( N(z) \), in the ocean is surface intensified, the kinetic energy contribution of the first baroclinic mode is also intensified there, leading to the conclusion that, to a useful first approximation, kinetic energy inferred from altimeter data is primarily (but not wholly) in the first baroclinic mode. This result appeared consistent with theory: Fu and Flierl (1980) and Smith and Vallis (2001) showed in simulations of quasi-geostrophic turbulence that the energy is rapidly transferred to deep vertical modes and concentrates in the barotropic and first baroclinic modes for stratification with a thermocline. Scott and Wang (2005) found evidence for a transfer of energy to the deep modes with a spectral analysis of the kinetic energy fluxes from altimetry. (When interpreting altimetric results, it is important to keep in mind that the barotropic mode contribution to the surface elevation, \( \eta \), is at least comparable to that of the baroclinic mode(s), and generally exceeds it—it is only the long spatial scales of the barotropic motions which reduce its velocity and hence its relative kinetic energy.)

More recently, it has been suggested that the near-surface, balanced motions, are instead described by so-called surface geostrophic solutions. As the difference between the two descriptions is, at least superficially, quite profound with implications for the interpretation of altimeter data generally, we now turn to a discussion of the oceanic balanced motions.

### 4. Balanced motions: the partition of interior modes and surface solutions

#### 4.1. Theory

Klein et al. (2009) and Lapeyre (2009) pointed out that Wunsch (1997), Smith and Vallis (2001) and other authors since did not consider the presence of surface density anomalies in their characterization of geostrophic motions. These anomalies drive surface intensified eddy motions, in addition to the interior baroclinic and barotropic modes. Tulloch and Smith (2009) confirmed that in simulations of quasi-geostrophic (QG) turbulence where surface density anomalies are allowed, a large fraction of the surface energy is associated with the surface quasi-geostrophic (SQG) solutions, using the meteorological nomenclature (Held et al., 1995; LaCasce and Mahadevan, 2006, and others). Lapeyre (2009) further speculated that the SQG solutions dominate the signal observed by the altimeter. In the remainder of this paper, the observational evidence for SQG motions and their importance to the estimates of eddy kinetic energy is explored.

QG scalings appear to accurately describe the range of scales that characterize the oceanic eddy field between about 10 and 500 km (Charney and Flierl, 1981). Within the QG approximation, the dynamics are fully described by the distribution of the QG potential vorticity, \( q \), given by,

\[
q = f_{0} + \beta y + \nabla^{2} \psi + \frac{\partial}{\partial z} \left( \frac{f_{0}^{2}}{N^{2}} \frac{\partial \psi}{\partial z} \right),
\]

\[
\nabla^{2} = \frac{\partial^{2}}{\partial x^{2}} + \frac{\partial^{2}}{\partial y^{2}}, \quad -H < z < 0,
\]

(1)

and buoyancy at the top and bottom boundaries,

\[
b = f_{0} \left. \frac{\partial \psi}{\partial z} \right|_{z=0,-H},
\]

(2)

where \( b = -g \rho / \rho_{0} \) is the buoyancy anomaly, \( \rho \) is the density, \( \rho_{0} \) is a reference density, \( \psi \) is the geostrophic streamfunction, \( f_{0} \) is the Coriolis parameter at the latitude considered, \( \beta \) is the planetary vorticity gradient, and \( H \) is the ocean depth. An ‘invertibility
principle" follows from eqs. (1) and (2); given the distributions of \( q \) and \( b \), one can solve for the geostrophic streamfunction, \( \psi \), and hence the eddy motions. In other words, knowledge of \( q \) and \( b \) is sufficient to diagnostically reconstruct the eddy dynamics—the elliptic problem for \( \psi \) has a unique solution for known boundary conditions.

Although \( \psi \) and the full spectrum of motions can be inferred from the distributions of \( q \) and \( b \) (the horizontal velocity field is given by gradients of the geostrophic streamfunction, \( u = -\partial_y \psi \) and \( v = \partial_x \psi \)), one needs prognostic equations to represent the evolution in time of the fields. In the QG approximation, the prognostic equation is given by the statement that the QG potential vorticity is stirred by the geostrophic velocity field, but is otherwise conserved,

\[
\frac{\partial q}{\partial t} + J(\psi, q) = 0, \quad -H < z < 0, \tag{3}
\]

where \( J \) is the Jacobian. Surface buoyancy, \( b_s \), is both advected by the surface geostrophic streamfunction, \( \psi \), and it is forced by vertical advection through the Ekman velocity \( (w_E) \) and by the divergence of surface fluxes, \( B \),

\[
\frac{\partial b_s}{\partial t} + J(\psi, b_s) = -w_E N^2 + B, \quad z = 0. \tag{4}
\]

At the bottom boundary, the forcing is primarily due to the vertical velocity induced by topographic variations and boundary layer drag,

\[
\frac{\partial b_b}{\partial t} + J(\psi, b_b) = -N^2 \left[ J(\psi, h) + \sqrt{\frac{\nu}{2 f_0}} \nabla^2 \psi \right], \quad z = -H, \tag{5}
\]

where \( h \) are the departures of bottom topography from \( z = -H \) and \( \nu \) is the viscosity. These forcings result in weak density fluctuations, because the abyssal ocean is less stratified than the upper, that is, \( N^2 \) is small and \( b_b \approx 0 \). For simplicity, topographic variations are ignored and the bottom boundary conditions is of vanishing buoyancy fluctuations, \( b_b = 0 \). Note that the bottom boundary condition (5) holds under the QG assumption that the mean ocean depth \( H \) is much larger than any topographic variation \( h \). To the extent that this condition is violated, the vertical mode decomposition presented below is not very accurate.

Following Charney (1971), Hoskins et al. (1985) and Bishop and Thorpe (1994), the principle of potential vorticity inversion can be used to decompose \( \psi \) into the so-called interior modes and the surface solutions. \( \psi \) is obtained from (1) and (2) by splitting it into two parts, \( \psi_{\text{int}} \) and \( \psi_{\text{surf}} \) (see Lapeyre and Klein, 2006, for more details),

\[
\nabla^2 \psi_{\text{int}} + \frac{\partial}{\partial z} \left( \frac{f^2}{N^2} \frac{\partial \psi_{\text{int}}}{\partial z} \right) = q - f_0 - \beta y, \tag{6}
\]

\[
f \left. \frac{\partial \psi_{\text{int}}}{\partial z} \right|_{z=0} = 0, \tag{7}
\]

\[
f \left. \frac{\partial \psi_{\text{int}}}{\partial z} \right|_{z=-H} = 0 \tag{8}
\]

and

\[
\frac{f^2}{N^2} \frac{\partial \psi_{\text{surf}}}{\partial z} = q - f_0 - \beta y, \tag{9}
\]

\[
f \left. \frac{\partial \psi_{\text{surf}}}{\partial z} \right|_{z=0} = b_s, \tag{10}
\]

\[
f \left. \frac{\partial \psi_{\text{surf}}}{\partial z} \right|_{z=-H} = 0. \tag{11}
\]

Because the problem is formally linear, this separation is just the well-known classical one of a problem with interior sources, but a homogeneous boundary condition; and of one a homogeneous interior problem, with an inhomogeneous boundary condition (see e.g. Jackson, 1975, and the Appendix here). These two elliptic problems produce two different solution sets: \( \psi_{\text{int}} \) satisfies a homogeneous surface boundary condition and is governed only by the interior \( q \) distribution, giving rise to the Sturm-Liouville modes described below. In contrast, \( \psi_{\text{surf}} \) is associated with surface buoyancy anomalies with a zero interior \( q \). It defines the so-called SQG solutions, \( E \) (Lapeyre, 2009).

The interior solution \( \psi_{\text{int}} \) can be projected onto vertical eigenfunctions \( F_j(z) \) and Fourier modes in the horizontal and in time,

\[
\psi_{\text{int}} = \frac{1}{(2\pi)^3} \frac{1}{H} \int \int \int \psi_{\text{int}}(k_x, k_y, \omega) \times F_j(z) e^{i k_x x + i k_y y - i \omega t} \, dx \, dy \, d\omega, \tag{12}
\]

with

\[
(k^2 + \lambda_j^2) \tilde{\psi}_{\text{int}}(k_x, k_y, \omega) = -\int \int \int_{-H}^{0} [q(x, y, z, t) - f_0 - \beta y] \times F_j(z) e^{-i k_x x - i k_y y + i \omega t} \, dx \, dz \, dt, \tag{13}
\]

where \( H \) is the ocean depth and \( k^2 = \kappa^2 + \lambda_j^2 \). The \( F_j(z) \) are the eigenfunctions solutions of the Sturm-Liouville vertical operator,

\[
\frac{d}{dz} \left( \frac{f^2}{N^2} \frac{d F_j}{dz} \right) = -\lambda_j^2 F_j, \quad \frac{d F_j}{dz} = 0, \quad z = 0, -H, \tag{14}
\]

with the eigenvalues \( \lambda_j \), which are the inverse deformation radii. (Mode \( j = 0 \) is the barotropic one, \( j = 1, 2, \ldots \), are the first, second, etc., baroclinic ones.) A question arises as to whether the \( q - f_0 - \beta y \) term can be expanded into periodic Fourier modes. While this is a concern for climatological potential vorticity distributions, here the focus is on potential vorticity fluctuations associated with geostrophic eddies on scales smaller that

\[\text{Tellus 62A (2010), 2}\]
O(100 km) which do not have large-scale linear gradients and can be expanded efficiently into Fourier modes.

The Fourier transform of the surface streamfunction \( \psi_{\text{surf}} \) can be written in the form,

\[
\hat{\psi}_{\text{surf}} = \int_0^1 \hat{b}_j(\kappa_x, \kappa_y, \omega) E(\kappa, z),
\]

(14)

where \( \hat{b}_j \) is the amplitude of the Fourier transform in time and space of buoyancy. For each wavenumber the ‘surface solution’, \( E(\kappa, z) \), satisfies,

\[
\frac{\partial}{\partial z} \left( \frac{f^2}{N^2} \frac{\partial E}{\partial z} \right) = \kappa^2 E, \quad \frac{\partial E}{\partial z} = 1 \quad \text{at} \quad z = 0,
\]

\[
\frac{\partial E}{\partial z} = 0 \quad \text{at} \quad z = -H. \quad \text{(15)}
\]

Unlike the \( F_j \) modes, the vertical structure of \( E \) depends upon the horizontal wavenumber, and the system is forced.

Finally, the Fourier transform of the geostrophic streamfunction is given by the sum of interior modes and surface solutions,

\[
\hat{\Psi}(\kappa_x, \kappa_y, z, \omega) = \sum_j \hat{\psi}^{\text{int}}_j(\kappa_x, \kappa_y, \omega) F_j(z) + \frac{\hat{b}_j(\kappa_x, \kappa_y, \omega)}{f_0} E(\kappa, z). \quad \text{(16)}
\]

The vertical structure of the Fourier modes is given by the \( F_j \) modes for motions associated with interior potential vorticity anomalies, and by the \( E \) functions for the motions driven by the surface buoyancy anomalies.

It is useful to compare the decomposition, eq. (16), with the more traditional one of just the linear modes of the QG equations. The former decomposition relies on the separation of motions driven by interior potential vorticity anomalies and those arising from surface buoyancy anomalies. The latter are based on the normal modes that arise from the QG equations linearized about a fixed mean state on a flat bottom. Only in special limits do the linear modes correspond to interior modes or surface solutions.

If the mean state includes an interior potential vorticity gradient on a flat bottom, but no surface buoyancy contribution, then the linear modes are the free modes discussed, for example, by Gill (1982) and used by Wunsch (1997). These modes have the same vertical structure and are identical to the interior modes \( F_j(z) \), because they both satisfy homogeneous boundary conditions. Alternatively, the mean state can be chosen to have only a surface buoyancy contribution, but no interior potential vorticity variations. Classical examples are the Eady (1949) problem on the \( f \)-plane or the modified one on the beta plane, where the planetary vorticity gradient is cancelled by a vertically sheared velocity (Lindzen, 1994). In both examples the linear solutions are surface trapped and correspond to the \( E(\kappa, z) \).

Philander (1978) and Frankignoul and Müller (1979a,b) considered the linear modes with an interior PV gradient that arise if specific forcing is applied at the surface. Whether the ‘forced’ linear modes project onto surface solutions or interior modes, depends on the mean state considered. A more detailed analysis of the forced modes is deferred to Section 4.4, where the ocean response to a time dependent wind forcing is discussed.

### 4.2. Projection Onto the interior vertical modes

Following (hereafter W97 Wunsch, 1997), the full geostrophic streamfunction, composed of interior modes and surface contributions, can be projected onto the interior modes, \( F_j(z) \), because they represent a mathematically complete basis for \( u, v \); see the Appendix. The total streamfunction can be expressed as a function of horizontal wavenumber and modes \( F_j \) alone,

\[
\hat{\Psi}_j = \hat{\psi}^{\text{int}}_j(\kappa_x, \kappa_y, \omega) + \frac{f_0 F_j(0)}{H N^2(0)} \frac{\hat{b}_j(\kappa_x, \kappa_y, \omega)}{\kappa^2 + \lambda_j^2}, \quad \text{(17)}
\]

and \( \hat{\Psi}_j \) is the total amplitude of the mode. Note that the WKBJ approximation of the normal modes shows \( F_j(0) = [2HN(0)]^{1/2} \int_{-H}^0 N(z') \, dz' \), that is, \( F_j(0) \) is independent of mode number for \( j \geq 1 \) (the barotropic mode has unit amplitude). The projection of surface solutions onto the interior ones decreases with mode number and horizontal wavenumber.

Because the \( F_j \) are complete in representing \( u, v \), and the \( E \) project onto them, the distinction between \( E \) and the ordinary linear \( F_j \), is primarily one of efficiency. To the degree that motions are dominated by \( E \), the projection onto the \( F_j \) would produce a decomposition in which the latter were phase-locked in time in such a way as to maintain the near-surface amplification. If, however, the \( F_j \) are randomly phased, the \( E \) solutions cannot dominate. Phase-locking is necessary, but not sufficient, to imply the presence of \( E \)—as it can be produced by other physics (e.g. surface wind forcing or bottom topography).

Anticipating that oceanic motions are well described by a superposition of the barotropic and first baroclinic modes (W97 and Section 4.7 of this paper), the phase relationship between these two modes can then be used to test whether the motions are consistent with SQG theory predictions. Altimetric data suggest that the energy-containing eddies have scales close to the first deformation radius. If those motions were associated with surface buoyancy anomalies, they would decay exponentially from the surface with an e-folding scale close to that of \( F_1 \), as can be seen by substituting \( \kappa = \lambda_1 \) in the hyperbolic problem for \( E \) in (15). The barotropic and first baroclinic modes can sum up to reproduce such a vertical profile, if they are phase-locked to reinforce at the surface and cancel at depth. Lacking such a phase-locking, the SQG hypothesis is not tenable.

Potential vorticity and streamfunction are less accessible from observations than are KE and PE. In a QG system, the total energy takes the form,

\[
\mathcal{E} = \frac{1}{2} \iiint |\nabla \psi|^2 \, dx \, dy \, dz \, dt + \frac{1}{2} \iiint \frac{f^2}{N^2} |\partial_z \psi|^2 \, dx \, dy \, dz \, dt.
\]
The first term is the KE and the second the PE, and the two contributions to energy can be expressed in terms of the interior modes and surface solutions,

\[
\mathcal{E}_k = \frac{1}{2} \int \int \int |\nabla \psi|^2 \, dx \, dy \, dz \, dt \\
= \frac{1}{2} \left( \frac{1}{(2\pi)^2} \right) \frac{1}{H} \\
\times \sum_j \int \int \int \kappa^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \, d\kappa_x \, d\kappa_y \, d\omega,
\]

(18)

\[
\mathcal{E}_p = \frac{1}{2} \int \int \int \frac{f_L^2}{N^2} |\partial_t \psi|^2 \, dx \, dy \, dz \, dt \\
= \frac{1}{2} \left( \frac{1}{(2\pi)^2} \right) \frac{1}{H} \\
\times \sum_j \int \int \int \lambda_j^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \, d\kappa_x \, d\kappa_y \, d\omega.
\]

(19)

In Section 4.7 it is shown that, consistent with the results of W97, the bulk of the KE measured by current meters mounted on moorings is typically concentrated in the barotropic \( (j = 0) \) and first baroclinic \( (j = 1) \) modes. However the goal of this paper is not to analyse the partitioning of energy across \( F_j \) modes, but rather to investigate its partitioning between surface and interior solutions. The first step is then to convert the results by integrating over the horizontal wavenumbers to obtain frequency spectra. Expressions for the frequency spectra in the various grating over the horizontal wavenumbers to obtain frequency solutions. The first step is then to convert the results by integrating over all frequencies of the expressions for KE and PE rather to investigate its partitioning between surface and interior integrals over all frequencies of the expressions for KE and PE in (18) and (19),

\[
S_{ij}(\kappa_x, \kappa_y) = \int \left( \kappa^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \right) \, d\omega,
\]

(20)

\[
S_{ij}(\kappa_x, \kappa_y) = \int \left( \lambda_j^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \right) \, d\omega,
\]

(21)

where the angle brackets indicate an average over many realizations of the turbulent field. At scales smaller than the deformation radius, theory predicts that the wavenumber spectra are isotropic and it is useful to work in terms of the isotropic spectra, that is, spectra integrated along circles in wavenumber space, defined as,

\[
S_{ij}^{i}(\kappa_x, \kappa_y, \phi) = \frac{1}{2\pi} \int_{0}^{2\pi} \kappa S_{ij}(\kappa_x, \kappa_y, \phi) \, d\phi,
\]

(22)

\[
S_{ij}^{p}(\kappa_x, \kappa_y, \phi) = \frac{1}{2\pi} \int_{0}^{2\pi} \kappa S_{ij}(\kappa_x, \kappa_y, \phi) \, d\phi,
\]

(23)

where \((\kappa, \phi)\) are polar coordinates in wavenumber space.

Unfortunately, few oceanic data provide simultaneous measurements of the horizontal and vertical distributions of eddy energy. Current meters mounted on moorings do provide measurements of eddy energy as a function of depth and time. Eddy energy spectra as a function of frequency and mode number can be constructed,

\[
S_{ij}(\omega) = \int \int \int \kappa^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \, d\kappa_x \, d\kappa_y,
\]

(24)

\[
S_{ij}(\omega) = \int \int \int \lambda_j^2 \left( \hat{\psi}_j^{uu} + \frac{f_L F_j(0)}{H^2(0)} \frac{\dot{b}_j}{\kappa^2 + \lambda_j^2} \right)^2 \, d\kappa_x \, d\kappa_y,
\]

(25)

The question then is how to relate wavenumber to frequency spectra—because we have theories for the former, but data for the latter.

The velocity and buoyancy fluctuations at a fixed point in a turbulent fluid are due to the sweeping of small eddies by the energy-containing eddies. That is, a probe measuring at a fixed point in space would map the spatial frequency of the turbulence into the time-frequency by the ‘Taylor hypothesis’,

\[
\omega = U\kappa,
\]

(26)

where \( U \) is the RMS velocity of the energy containing scales (e.g. Vallis, 2006). Such a relationship holds only if the KE at a wavenumber \( \kappa \) is dominated by the larger energy containing eddies. This hypothesis requires that the KE spectrum rolls off faster than \( \kappa^{-1} \) in wavenumber (Vallis, 2006), consistent with the turbulence generated in QG models. The relationship also requires that the energy-containing eddies are more persistent than small-scale ones, so that \( U \) is approximately steady on
many turnover times of the smaller scale eddies. In general, one would anticipate that $U$ would necessarily have a stochastic character. The argument further assumes that no waves are present in the turbulent field, so that the frequency variability is entirely generated by stirring and not by intrinsic oscillatory motions.

Alternatively, the relationship between frequencies and wavenumbers can be diagnosed from the oceanic energy spectrum as measured by altimeters. Fig. 2 shows an estimate of the zonal wavenumber–frequency power density spectrum of sea level (from Wunsch, 2009). Dashed lines indicate the linear dispersion curves for the barotropic and first baroclinic mode basic theory Rossby waves of a flat bottom ocean. Although significant energy is indistinguishable from the first baroclinic mode at the very lowest frequencies and wavenumbers, the great bulk of the energy lies far from the dispersion curves of linear theory. As already discussed above, altimetric observations do not separate the energy in wavenumbers having radically different frequency content, and the signal to noise ratio at high wave numbers is poor. The amplification of the energy spectrum along a straight line (the ‘non-dispersive line’) confirms that for motions lying close to that region, a linear relationship between $\kappa_x$ and $\omega$ is reasonable and would set a bound, over 15 yr, on variations in $U$, if the Taylor hypothesis is to be invoked. Similar results are found at other latitudes far from strong currents. In major currents the zonal wavenumber–frequency is still linear but Doppler-shifted by the mean flow as shown in Fig. 3, a frequency–zonal wavenumber spectral estimate for 55°S in the Southern Ocean. Even though such a relationship seems to vindicate Taylor’s hypothesis, it is unlikely to be the result of advection of small eddies by larger ones. Because the larger scale eddies are here part of the turbulence, $U$ itself is a time-varying stochastic variable, and the latitude dependence, and sharpness of the non-dispersive line over 15 yr, is not easy to reconcile with its necessary temporal variation. Chelton et al. (2007) show that the linear relationship is likely the result of waves and eddies propagating zonally at a wave speed $c$, so that
$$\omega = c\kappa_x. \quad (27)$$

Consistent with this interpretation, the energy distribution in frequency–meridional wavenumber shows no preferential phase speed (not shown). Regardless, based on Fig. 2, one can assume a linear relationship between frequency and zonal wavenumber. This key result can then be used to make predictions about oceanic frequency spectra. If the wavenumber spectrum is isotropic and is proportional to $\kappa^{-\alpha}$, then the one-dimensional zonal wavenumber spectrum is also proportional to $\kappa^{-\alpha}$. The
relationships (26 or 27) then imply that the frequency spectrum is proportional to \( \omega^{-3} \).

The characteristics of QG turbulence depend on whether it is generated primarily by stirring of large-scale interior potential vorticity anomalies or whether it is mostly due to stirring of large-scale surface buoyancy gradients. In the former case, turbulence is dominated by the interior modes, while surface buoyancy contributions to KE and PE can be ignored. Charney (1971) and, more recently, Smith and Ferrari (2009), show that in this case the spectrum of \(| \Psi_j^{\mu} (k_x, k_y) |^2 \) is confined to the barotropic and first baroclinic modes, is nearly horizontally isotropic in wavenumber space, and scales as \( \kappa^{-3} \) for horizontal scales shorter than the first deformation radius. Using eq. (22), \( S_j (\kappa) \sim \kappa^{-3} \) with contributions for the barotropic and first baroclinic modes. The barotropic mode does not project on \( \Psi_j \) and (23) implies that \( S_j (\kappa) \) is dominated by the first baroclinic mode and rolls off as \( \kappa^{-5} \). Using the relationships in (26 or 27), one obtains scalings for the Eulerian frequency spectra of the \( F_j \),

\[
S_{ij} (\omega) \propto \omega^{-3}, \quad \text{for } j = 0, 1 \tag{28}
\]

\[
S_{ij} (\omega) \propto \omega^{-5}, \quad \text{for } j = 1. \tag{29}
\]

Tulloch and Smith (2009) show that the buoyancy anomalies in SQG turbulence simulations have a 2-D spectrum \( | \hat{b}_j |^2 \propto \kappa^{-8/3} \) (corresponding to an isotropic buoyancy spectrum rolling off as \( \kappa^{-5/3} \)) for scales smaller than \( O(100) \) km. If the surface solutions dominate the overall energy, the energy spectra be proportional to,

\[
S_{ij} (\omega) \propto |F_j (0)|^2 \omega^{1/3}, \quad \text{for } j = 0, 1 \tag{30}
\]

\[
S_{ij} (\omega) \propto |F_j (0)|^2 \omega^{-5/3}, \quad \text{for } j = 1. \tag{31}
\]

Once again, the wavenumber spectra have been converted into Eulerian frequency spectra using (26). For wavenumbers/frequencies smaller than the corresponding deformation radius, \( S_{ij} \propto \omega^{1/3} \), while for higher wavenumbers/frequencies \( S_{ij} \propto \omega^{-11/3} \). The same scaling is obtained using relationship (27).

Data (discussed below) suggest a power law close to---2 and are not consistent with either the interior potential vorticity or surface buoyancy turbulence prediction. Turbulence theories are not irrelevant to the ocean, but apparently other processes dominate the excitation of energy at high (but subinertial) frequencies.

4.4. Forced Eulerian frequency spectra

The turbulent hypothesis considered above is that all eddy variance at high frequencies is generated through a turbulent cascade of potential vorticity and buoyancy variance from large to small scales. An alternative view is that eddy energy at high frequencies is directly forced by winds and buoyancy fluxes at the ocean surface, and dominates the observations. In this scenario, the linearized potential vorticity and buoyancy equations in the presence of external forcing are,

\[
\partial_t q + \beta \partial_z \psi = 0, \quad -H < z < 0, \tag{32}
\]

\[
\partial_t b = Q \quad \text{at } z = 0, \quad b = 0 \quad \text{at } z = -H. \tag{33}
\]

where \( Q = -u E N^2 + B \) (recall eq. 4) is the sum of wind forcing and buoyancy fluxes. Note that the wind forcing is being written as an equivalent buoyancy forcing from Ekman layer divergences. Apart from external forcing, the equations are linearized around a uniform background state (i.e., no mean currents). It is worth interpreting this system in light of the surface/interior modes described above. Eq. (33) shows that external forcing generates perturbations, \( b \), in the QG approximation. Eq. (32) then shows that the geostrophic streamfunction stirs the planetary potential vorticity gradient in the ocean interior and generates \( q \) perturbations. In summary, the forced problem generates both \( b \) and \( q \) perturbations and therefore excites a superposition of \( F_j \) and \( E \).

Solutions to this system of equations are given in Frankignoul and Müller (1979a),

\[
\hat{\Psi} (k_x, k_y, \omega, z) = \sum_j f_j \hat{Q} (k_x, k_y, \omega) F_j (z) \tag{34}
\]

where \( \hat{Q} (k_x, k_y, \omega) \) is the Fourier transform of the forcing, \( Q \), and \( \omega_j = \beta k_x / (\kappa^2 + \lambda_j^2) \) are the intrinsic frequencies of Rossby waves.² Absent other effects, the amplitude of the normal modes \( \hat{\Psi}_j \) becomes infinite when the frequency of the forcing matches the dispersion relation of Rossby waves for some wavenumber.

The oceanic response can therefore be off-resonant or resonant. For frequencies larger than the maximum frequency, \( \omega_j^{\max} \), of Rossby waves, no resonance can occur and the oceanic response can be described in terms of finite power spectra. In the frequency range, \( \omega \leq \omega_j^{\max} \), there always exists a wavenumber for which the oceanic response is resonant and formally infinite. The singularities could be removed by introducing dissipation mechanisms. Because, however, the nature of dissipation

² Flierl (1978) shows that one can define a single parameter \( \lambda^2 = -k^2 - \beta k_x / \omega \), which expresses the ‘equivalent depth’ of the forcing (Lindzen, 1967; Philander, 1978). One can then write the solution for a forced problem as,

\[
\hat{\Psi}_j = \frac{f_j \hat{Q} (k_x, k_y, \omega)}{\tau_0 H N^2 (0)} \sum_j F_j (0) F_j (z) \tag{41}
\]

The parameter \( \lambda^2 \) can be either positive or negative, corresponding to positive or negative equivalent depths \( f_\beta \lambda^2 / g \). Negative equivalent depths \( (\lambda^2 < 0) \) give a response with a pressure signal (the geostrophic streamfunction) decreasing rapidly with depth. Setting \( \beta = 0 \), prevents the generation of interior potential vorticity anomalies, \( \lambda^2 \) becomes negative, and the forced linear modes project only onto the surface modes.
is still controversial, as discussed in Section 2, the resonant response cannot be described easily. Frankignoul and Müller (1979a) show that the response at resonance could be used to compute the rate of energy input at those frequencies. This energy would then flux to other frequencies through wave–wave interactions and dissipation.

Data, however, suggest that resonant responses in the ocean are unlikely. For barotropic motions, Luther (1982), and Woodworth et al. (1995) show evidence for a weak excess energy near 5-d periods. Frequency–wavenumber spectra (Wunsch, 2009) from altimetry show almost all of the energy at mid-latitudes measurably removed from any linear dispersion relationship—a requirement for resonance (see Fig. 2). Absorbing and scattering topographic features and ocean boundaries, as well as strong non-linear interactions, probably preclude anything approaching true basin resonances or forced travelling wave ones. Even the excess energy observed at the vanishing zonal group velocities for equatorially trapped gravity wave modes (Wunsch and Gill, 1976), is not apparent at the analogous frequency and wavenumber for Rossby wave modes.

The off-resonant response is proportional to the atmospheric forcing function \( Q \). Following Frankignoul and Müller (1979a), the \( Q \) is assumed to be a realization of a statistically stationary and homogeneous process with zero mean and power spectrum \( S_Q(\kappa_x, \kappa_y, \omega) \) defined by,

\[
\langle \hat{Q}(\kappa_x, \kappa_y, \omega) \hat{Q}^*(\kappa'_x, \kappa'_y, \omega') \rangle = S_Q(\kappa_x, \kappa_y, \omega) \delta(\kappa_x - \kappa'_x) \delta(\kappa_y - \kappa'_y) \delta(\omega - \omega'),
\]

where the angle brackets denote ensemble averages and the asterisks complex conjugates. Reality, stationarity, and homogeneity imply \( S_Q(\kappa_x, \kappa_y, \omega) = S_Q(-\kappa_x, -\kappa_y, -\omega) \). Note that definition eq. (35) implies random phases, thereby excluding standing (basin) modes.

In mid-latitudes, the dominant time scale of the atmospheric fields is a few days and most of the air-sea fluxes are associated with the eastward travelling frontal cyclones and anticyclones with wavelengths from 3000 to 7000 km. However, the forcing of the ocean is not confined to the energetic weather system band, because atmospheric forcing has significant energy at lower frequencies and wavenumbers. Atmospheric spectra are approximately white in frequency for periods longer than 10–20 d, expect for the annual peak (Hasselmann, 1976). Frankignoul and Müller (1979b) show that the wind curl stress dominates the atmospheric forcing of the ocean and has a white isotropic spectrum in wavenumber for scales shorter than about 3000 km. Furthermore the frequency and wavenumber components of the spectrum appear to be separable (Willson, 1975), so that a reasonable approximation of the atmospheric forcing spectrum is,

\[
S_Q(\kappa_x, \kappa_y, \omega) = S^0_Q(\kappa) S_Q(\omega),
\]

where \( S^0_Q(\omega) \) is the white frequency component of the spectrum and \( S_Q(\kappa) \) is the wavenumber component of the spectrum, which is zero at large-scales and becomes white for scales shorter than \( \sim 3000 \) km. For present purposes, the wavenumber dependence of the spectrum is not very relevant, because we are interested in the frequency response of the ocean obtained by integrating over all wavenumbers.

The off-resonant oceanic spectrum generated by atmospheric forcing is obtained by substituting the spectrum (36) into the solution for \( \hat{\Psi} \), eq. (35). For frequencies much larger than those of Rossby waves, the oceanic spectra are predicted to be red and behave like \( \omega^{-2} \). This feature is a general one of the response of a much simplified long time scale ‘climate’ system to short time scale variability of the ‘weather’ (Hasselmann, 1976).

Interestingly, in the off-resonant limit, the interior PV gradient \( \beta \) does not enter at leading order in the solution and the wind-driven response can be understood in terms of forced SQG dynamics—and which is then the same as the linear forced solutions. The wind-driven Ekman pumping generates surface buoyancy fluctuations, which drive interior motions. The resulting streamfunction spectrum takes the form,

\[
|\hat{\Psi}|^2 \approx \left[ f_z^2 S^0_Q(0) E^2(\kappa, z) \right] \left[ S^0_Q(0) \omega^{-2} \right],
\]

that is, it is separable into its \( z \)-wavenumber dependence and its frequency dependence. The \( z \)-wavenumber dependence is in the form of the SQG \( E \) surface solution with a magnitude set by the winds amplitudes through \( S^0_Q(0) \). The frequency dependence is instead proportional to \( \omega^{-2} \) at all levels.

4.5. Altimetric data

Why might one anticipate the existence of the \( E \) solutions? Le Traon et al. (2008) find that sea level wavenumber spectra are significantly different from a \( \kappa^{-5} \) law as expected for turbulence generated by interior potential vorticity anomalies (the sea level spectrum is equal to the eddy kinetic energy spectrum divided by \( \kappa^2 \)). Instead the spectra have a roll-off not inconsistent with a \( \kappa^{-5/3} \) slope on scales between 100 and 300 km. Furthermore, they find that the variance of SST estimated from satellite microwave radiometers also display a \( \kappa^{-5/3} \) slope at scales below 300 km. The \(-5/3 \) slope in eddy kinetic energy and SST variance would indicate that the SQG is a better dynamical framework than the QG turbulence theory to describe the ocean surface dynamics. But the analysis is based on high eddy energy regions like the Gulf Stream, Kuroshio and Agulhas regions to maximize the signal to noise ratio at small scales, and questions remain as to the generality of the results. This interpretation, moreover relies on the assumption that all motions at these scales are generated by turbulent stirring with little contribution from wind forcing.

SST spectra with a \( \kappa^{-2} \) rolloff are also reported from SeaSoar data (e.g. Ferrari and Rudnick, 2000, and references therein). A \(-2 \) rolloff is observationally indistinguishable from \(-5/3 \). The traditional explanation is that the \( \kappa^{-2} \) is a reflection of frontogenesis in the upper ocean (fronts are step discontinuities
Fig. 4. Positions of the Southern Ocean moorings analysed here superimposed upon depth contours in metres. A, B, C denote the positions of three moorings discussed in the text.

Fig. 5. Temperature, salinity and buoyancy frequency, \( N(z) \), profiles (left- to right-hand side) from the Gouretski and Koltermann (2004) climatology at the position of mooring A. The deep stratification is nearly uniform and weaker than in the strong near-surface thermocline, but distinctly non-zero. Baroclinic modes will be strongly amplified near surface by the increase in \( N(z) \).

4.6. Mooring data

A crucial issue is the vertical structure of the motions. In the compilation by W97, the then available data base was found marginal at best for drawing conclusions about the vertical structure, and the result might be best summarized as showing that there is no conflict with the inference, already noted above, that roughly 40% of the water column kinetic energy at periods beyond one day lies in the barotropic mode (a bit more in the North Atlantic), about 40% in the first baroclinic mode, and the rest either in higher modes or observational noise. Although much detail is ignored, this crude summary does capture the basic result. At the time of that calculation, the possibility of surface-trapped motions was set aside, primarily on the grounds that almost no moorings existed with sufficient near-surface instrument positions to demonstrate a failure of the flat-bottom, free-mode description. It is important to recall that a single mooring has no horizontal scale discrimination power, although it is a powerful means for separating the internal wave field from lower frequency motions, something not generally possible with the altimeter. To demonstrate the nature of the remaining issues, we here briefly analyse the results from three moorings from the Southern Ocean, not available for the earlier study, and one mid-latitude mooring. The fundamental decomposition is in the form,

\[
[u(z, t), v(z, t)] = \sum_{n=0}^{M} [\alpha_n(t), \alpha_n(t)] F_j(z) + \text{residual}. \quad (38)
\]

with a \( \kappa^{-2} \) spectrum). The shallow SQG spectra at the ocean surface are the QG approximation of frontogenesis and hence the two interpretations are consistent.

Mooring data provide an alternative view of oceanic spectra, because they resolve high frequencies. The frequency–wavenumber relationships (26 or 27) are now used to interpret oceanic frequency spectra.

4.6.1. Near 60°S (Mooring A). A Southern Ocean mooring at 57.5°S, 4.05°E, Fig. 4, produced records of about 200 d from about 4500 m of water with five usable current meter and temperature records (from E. Fahrbach). The climatological temperature (in situ), salinity and buoyancy frequency profiles from the Gouretski and Koltermann (2004) climatology are shown in Fig. 5. Stratification is nearly uniform below about 400 m, with a very sharp near-surface thermocline, which can be expected to amplify the \( F_j(z) \) near the surface, \( j \geq 1 \).
Fig. 6. Instrument depths (mooring A) superimposed upon the horizontal velocity/pressure modes, \( F_i \), \( i = 0,1 \), (left-hand side) and vertical displacement modes, \( G_i \), \( i = 0,1,2 \) for a flat-bottom, resting ocean, subject to a rigid lid surface boundary condition. The sharp increase in the amplitude of the \( F_1(z) \) mode near the surface, \( z = 0 \), is an important characteristic, but is poorly defined by the available instrumentation. The barotropic vertical displacement mode is plotted as zero amplitude.

Fig. 7. Kinetic energy profile at mooring A. Note the linear scales. Increase towards the bottom is assumed to be owing to the existence of a bottom trapped mode. The KE increase towards the surface is roughly consistent with the amplification there of \( F_1(z) \).

The modal shapes and instrument depths can be seen in Fig. 6. The vertical displacement modes, \( G_i(z) = f^2 N^{-2} \partial F_i / \partial z \), vanish at \( z = 0 \), producing almost no buoyancy signature, whereas the horizontal velocity or pressure modes, \( F_j(z) \), are finite there. The kinetic energy profile from the mooring data is in Fig. 7. The instrument closest to the surface at 180 m (Fig. 7) shows an increase in kinetic energy relative to the instruments at mid-water depths, not inconsistent with the very sharp near-surface thermocline and the corresponding amplification of \( F_1(0) \), which increases by about a factor of six, from 500 m to the surface. The shallowest instrument lies below the region of strongest increase in velocity amplitude towards the surface expected from the \( F \) modes. Data duration is only 214 d, and inadequate temporal coverage is a pervasive problem.

Following the methodology of W97, a modal fit for \( u(t) \) was done. When averaged over the whole record, for \( u \), 71% of the variance is barotropic, 23% first baroclinic and 5% second baroclinic. For \( v \), the corresponding numbers are 72, 19 and 7%. For temperature, 77% is in the first baroclinic mode, 21% in the second baroclinic and 2% in the third baroclinic mode. In summary, this mooring, south of the Antarctic Circumpolar Current, shows a strong predominance of barotropic kinetic energy in periods shorter than about 200 d, with a comparatively modest contribution from the first baroclinic mode. If the SQG solution contains primarily motions near the Rossby-radius, its vertical structure is almost indistinguishable from that of the near-surface behaviour of the first baroclinic mode.

Spectral densities as a function of temporal frequency of the modal coefficients (Fig. 8) are strongly red, with no obvious tendency to flatten at low frequencies. These were computed using a Daniell window on a periodogram, as the usually more desirable multitaper method introduces a low frequency negative bias into the spectra. Power laws are roughly \(-2\) for the barotropic and first baroclinic modes. KE spectral estimates from individual instruments as a function of depth (not shown) do not display any obvious change of slope as one approaches the surface.

The coherence between the barotropic and first baroclinic modal amplitudes for the zonal velocity (Fig. 9) is weak, but significant (values around 0.6) at 180° between the two modes. This phase is such (see Fig. 6) to amplify the surface kinetic energy by phase locking the two modes as either \( E \) solutions or forced modes would require.

4.6.2. Agulhas Retroflection (Mooring B). In contrast is a mooring from an Agulhas Retroflection Experiment (WHOI 835, Luyten et al., 1990) at 40°S, 16.5°E, which ran for nearly a year in 4847 m of water. This region is an unusual one, and so in the interests of brevity, we only summarize the results. Note too that the pressure sensor shows that the instrument nominally at 408 m reached a pressure of almost 1100 m during one event. Such excursions render the interpretations of the fits doubtful during energetic periods. Modal coefficient spectra tend towards white noise at low frequencies, but with nearly equal kinetic energies in modes zero and one, more typical of mid-latitudes than the mooring south of the Circumpolar Current. Power law behaviour is again near \(-2\) in modes 0 and 1 at periods shorter than about 50 d. The coherence coupling of the lowest two modes is very strong down to 4 d so as to amplify the surface velocity. Agulhas eddies are, however, unlikely typical of the world ocean.

4.6.3. South of Tasmania (Mooring C). A mooring South of Tasmania (51°S, 143°E; See Phillips and Rintoul, 2000 and the
Fig. 8. For a mooring near 60°S in the Southern Ocean, spectra by mode number (0–3) for kinetic energy (left-hand panel) and (1–3) for vertical displacement (right-hand panel). Straight line segment is an $s^{-2}$ power law ($s = \omega / 2 \pi$).

Fig. 9. Coherence (upper panel) of the coefficients of the barotropic and first baroclinic modes at mooring A in the $u$-component. Horizontal dashed line is an approximate level-of-no-significance at 95% confidence. Low frequency phases (lower panel) are in the sense of tending to magnify the surface velocity.

Fig. 10. (Upper panel). $\alpha_u(0)$ (solid line) and $\alpha_u(1)$ (dashed) showing the temporal variability of the coefficients of the zonal component of flow south of Tasmania. (Lower panel) The same as in the upper panel except $\alpha_v(0)$, $\alpha_v(1)$.

position map, Fig. 4), produces barotropic and first baroclinic modes that are strongly coupled, again being additive near surface and subtractive at depth (Fig. 10). The modal partition is 59% in the barotropic mode with 37% in the first baroclinic mode for $u$, and with the partition being 67 and 27%, respectively for $v$. As noted above, the phase locking is consistent with the presence of the SQG surface solution, but also with other explanations. Spectra of the modal kinetic energies and of the vertical displacement power are shown in Fig. 11. Kinetic energy spectra display an approximate 2 power law.

3 Instrument depths provided in the data files obtained from the WOCE Current Meter Archive at Oregon State University are the instrument nominal depths. Actual pressure depths used here are taken from Table 1 of Phillips and Rintoul (2000), which appears to have an error in labelling the 3320 m instrument as having failed, rather than the one at 1150 m.
4.6.4. Nares Abyssal Plane. To contrast with the Southern Ocean moorings, the example of one on the Nares Abyssal Plane in the North Atlantic at 23° N, 64° W is recapitulated (it was included in W97). Spectral densities (Fig. 12) show more high mode contribution than in the Southern Ocean, but with a continued barotropic KE dominance in the lowest frequency band. Some marginal coherence at about 0.4 exists in a narrow band between about 15 and 20 d periods (not shown), but there is none detectable in the more energetic lower frequencies and the modes are not significantly phase-locked. This mooring is, as one expects absent strong modal coherence, one that was found to give very different values of surface KE values when independent and phase-locked modes were assumed.

4.6.5. Implications of the mooring data. Together with the results of W97, the mooring data lead to the conclusion that the $F_j(z)$ prove adequate to represent the horizontal velocities in essentially all of the available data. (Vertical displacement analyses, with some spectra shown here, have been de-emphasized because of the noisiness in the calculations arising from the need to use the time-varying temperature profiles.) Kinetic energy frequency spectra are generally close to $\omega^{-2}$ over much of the range of accessible time scales.

The existence of SQG surface solutions, $E$, requires an analysis of the phase coupling of the coefficients, $\alpha_r(t)$. In the Southern Ocean moorings here, there is some indication of barotropic/baroclinic mode coherence leading to near-surface amplification which would be consistent with the presence of a forced $E$. However, in the compilation of W97, and as tabulated there in the column marked ‘ratio’ in his Table 1, and as in the Nares Abyssal Plane mooring, as in many places, there is no sign of modal coupling. A very important added complication, not discussed here, is the expected presence of strong ageostrophic motions near the sea surface and for which there are almost no useful observations.

5. Discussion

A zero-order, but nonetheless only semi-quantitative, picture exists of the sources of energy sustaining the oceanic general circulation against dissipative losses (Fig. 1). Kinetic energy of the circulation system is strongly dominated by the geostrophic (balanced) motions whose spatial structure (wavenumber distribution) can be controlled by a number of competing processes including direct atmospheric forcing, up and downscale turbulent energy cascades, topographic interactions, etc. Testing various ideas against data about the energy flow within the balanced motions is very difficult for a number of reasons: at high wavenumbers, altimetric data are very noisy and the temporal sampling is infrequent. Mooring data are inadequate to determine horizontal spatial scales, and theory says almost nothing about the structure of the readily observed frequency spectra.

Geostrophic turbulence theories produce frequency spectra that are inconsistent with mooring observations—the frequency spectra generally being closer to an $\omega^{-2}$ power law at all depths, and thus flatter than predicted by turbulence theories. A linear theory of forced response, with no resonant modes, is more
consistent with what is observed. An important implication is that high frequency (but subinertial) variability in the ocean is directly forced by winds. Turbulent eddy–eddy interactions instead appear to shape the wavenumber spectrum at small scales as seen in altimetric and sea surface temperature observations. It appears that mooring and altimetric data are both necessary to fully describe the oceanic frequency–wavenumber spectrum, because they provide complementary views of ocean variability.

Despite the overwhelming evidence that oceanic motions are turbulent, in the sense that non-linear interactions rapidly redistribute energy across wavenumbers, classical theories of turbulence cannot account for many aspects of the observations. First, turbulence theories assume that eddies decorrelate on timescales shorter than a wave period, so that energy can spread away from the linear wave dispersion relationship \( \omega = \omega(\kappa_x, \kappa_z) \). However the spectra shown in Fig. 2 show that the bulk of the mid-latitude ocean KE is confined to a narrow strip in \( \kappa_z, \omega \) space suggesting that (non-linear) wave dynamics remains relevant despite any truly turbulent interactions. Second, turbulent theories assume that there is a wide range of scales where dynamics are controlled by internal interactions and not by external forcing. Observations suggest that wind forcing remains important on all scales.

The presence of structures in the frequency–wavenumber spectra corresponding to wave-like motions in the presence of a broad-background energy characteristic of turbulence, means that direct measurements of such spectra over the entire range of scales are required to definitively understand the nature of the motions. Because of orbit restrictions, altimeters do not directly sample motions with periods shorter than about 20 d, so that internal wave motions are not separable from balanced motions in the data (recall the prominence of the internal tides near a 60-d period). Isolated moorings almost never have durations exceeding two years and most are far-shorter; upper-ocean sampling on moorings is very ‘thin’, and they provide no horizontal spatial structure, so that the observed frequency spectrum is some poorly determined summation over all wavenumbers and their physics. To advance beyond using untestable features of numerical models, some combination of altimetry with properly instrumented, multiyear mooring deployments will be required, probably usefully supplemented with such techniques as towed sensors, gliders, and possibly seismic oceanography from ships.

6. Acknowledgments

Written in honor of G. Walin on the occasion of his 70th birthday. Current meter data were obtained from the WOCE Current Meter Archive at Oregon State University. Helpful comments were made by Patrice Klein, Ross Tulloch, Shafer Smith, Joe Pedlosky and Joe Lacasce. Glenn Flierl pointed out that the surface solution ought to project onto the interior modes. Supported in part by the National Ocean Partnership Program (ECCO), the National Aeronautics and Space Administration and by the National Science Foundation award OCE-0849233.

7. Appendix

Bretherton (1966) shows that the elliptic problem for \( \psi \) with non-homogeneous boundary conditions in (1 and (2) is equivalent to the elliptic problem with homogeneous boundary conditions,

\[
q = \frac{f_0}{N^2} \beta (z) = f_0 + \beta y + \nabla^2 \psi + \frac{\partial}{\partial z} \left( \frac{f_0}{N^2} \frac{\partial \psi}{\partial z} \right),
\]

\[
\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \quad \partial_z \psi = 0, \quad z = 0, -H,
\]

(39)

where the delta function represents the effect of surface buoyancy. The modes \( F_j \) are a complete basis for this problem, because they satisfy the same boundary conditions. However, a complete basis is defined as one that can represent functions possibly with a finite number of discontinuities, but finite over the domain of interest. PV is not such a function because of the delta function contribution at the boundary. The functions \( \psi, u, v \) are instead such functions and can be expressed as a linear combination of \( F_j \) modes. This is not to say that surface trapped solutions do not contribute to the total streamfunction; both interior modes and surface solutions project onto the \( F_j \) basis, and the projection of the surface solution is

\[
\frac{1}{H} \int_{-H}^{0} F_j E dz = \frac{f_0}{H N^2 (0)} \frac{F_j (0)}{\kappa^2 + \lambda_j^2}.
\]

(40)

The projection of the surface solution on the barotropic mode is the special case with \( F_j (0) = 1 \) and \( \lambda_j = 0 \).

References


