Large-Scale Propagating Disturbances: Approximation by Vertical Normal Modes

PETER D. KILLWORTH AND JEFFREY R. BLUNDELL

Southampton Oceanography Centre, Southampton, United Kingdom

ABSTRACT

Propagating features and waves occur everywhere in the ocean. This paper derives a concise description of how such small-amplitude, large-scale oceanic internal disturbances propagate dynamically against a slowly varying background mean flow and stratification, computed using oceanic data. For a flat-bottomed ocean, assumed here, the linear internal modes, computed using the local stratification, form a useful basis for expanding the oceanic shear modes of propagation. Remarkably, the shear modal structure is largely independent of orientation of the flow. The resulting advective velocities, which are termed *pseudovelocities,* comprise background flow decomposed onto normal modes, and westward planetary wave propagation velocities. The diagonal entries of the matrix of pseudovelocities prove to be reasonably accurate descriptors of the speed and direction of propagation of the shear modes, which thus respond as if simply advected by this diagonal-entry velocity field. The complicated three-dimensional propagation problem has thus been systematically reduced to this simple rule.

The first shear mode is dominated by westward propagation, and possesses a midlatitude speed-up over the undisturbed linear first-mode planetary wave. The pseudovelocity for the second shear mode, in contrast, while still dominated by westward propagation at lower latitudes, shows a gyrelike structure at latitudes above 30°. This resembles in both shape and direction the geostrophic baroclinic flow between about 500- and 1000-m depth, but are much slower than the flow at these depths. Features may thus be able to propagate some distance around a subtropical or subpolar gyre, but not, in general, at the speed of the circulation.

---

1. Introduction

Propagating features—of which waves are a subset—are ubiquitous in the ocean. They play an important climatic role, permitting information about changes in one region of the ocean to be transmitted to another. A variety of data sources now exists to demonstrate the existence of propagating features: historical hydrographic and XBT records, and remotely sensed sea surface temperature and sea surface height. Hydrographic data, while inadequate for many purposes, have proven ideal for analysis of large-scale changes (White 1977 is an early example). They have been used in the North Atlantic to show interdecadal changes (Levitus 1989), and in the North Pacific (Jacobson and Spiesberger 1998) to show planetary waves. Deser et al. (1996) and Zhang and Levitus (1997) used EOF analyses on the new Levitus data to examine subsurface propagation in the North Pacific, seeing cold pulses moving downward from the surface (where they may well be initiated by atmospheric forcing) into the main thermocline, on interdecadal timescales. At depths of 250 m, the move-

ment appeared clockwise around the subtropical gyre plus a westward phase propagation, similar to suggestions of Latif and Barnett (1994).

Sea surface temperature shows the movement of anomalies eastward across the North Atlantic, at much slower speeds than can be explained by advection, and has attracted recent attention. Sutton and Allen (1997) used the Comprehensive Ocean–Atmosphere Data Set and filtered the data. C. Gordon (2000, personal communication), using a different interpolation of the same data, finds a complicated regime of propagating features, some of which are traveling westward, and R. E. Houseago-Stokes and P. G. Challenor (2001, personal communication) find a faster signal (which would have been removed by the filter used by Sutton and Allen) in satellite data. The role of the atmosphere is unclear in these signals: the features may be coupled or ocean-only. An example of a feature which, apart from its formation, appears to be confined to the ocean is the Great Salinity Anomaly (Dickson et al. 1988), which is believed to have taken 10 yr to propagate around the North Atlantic subpolar gyre. The Antarctic Circumpolar Wave (White and Peterson 1996), whether a wave or a collection of events, is by contrast clearly coupled to the atmosphere.

Altimetry is the main source of data for propagating disturbances. Starting with Tokmakian and Challenor (1993) a number of authors have found planetary waves

Corresponding author address: Dr. Peter Killworth, Southampton Oceanography Centre, Empress Dock, Southampton, SO14 3ZH, United Kingdom.
E-mail: p.killworth@soc.soton.uk

in altimeter data. Chelton and Schlax (1996) produced the first global analysis and showed that the zonal speed was distinctly higher than that predicted by linear theory, though Kessler (1990) had found suggestions of this earlier in XBT data. [Zang and Wunsch (1999) have debated whether waves are in fact faster than linear theory.] However, planetary waves are not only visible in altimeter data. Cipollini et al. (1997) showed that in the Azores region they were clearly visible in infrared measurements of surface temperature, and Hill et al. (2000) have extended these observations to the global ocean. Cipollini et al. (2001) have made similar observations with ocean color.

The Chelton and Schlax (1996) observations triggered much theoretical work. Killworth et al. (1997, henceforth KCdS) showed that the speed of the first baroclinic mode of propagation is mainly increased by the presence of (baroclinic) mean flow, through changes in the background potential vorticity, and gave the first proof of the noninteraction theorem; this was confirmed by Dewar (1998) and Liu (1999b). These latter papers, together with de Szoeke and Chelton (1999), gave arguments for why the increase in speed was found predominantly in mid and high latitudes. Dewar and Morris (2000) have examined baroclinic wave propagation in an eddy-resolving flat-bottomed quasigeostrophic model, finding modifications to wave speeds consistent with the above theories. Killworth and Blundell (1999) showed that the effects of basinscale topography on propagation were locally large, but globally small. The combination of the mean flow and topography has yet to be investigated.

Interest has moved recently toward higher modes, when they exist (e.g., Cipollini et al. 1997). Liu and Zhang (1999), using ray tracing in a 2.5-layer model of the North Pacific, showed that the first baroclinic planetary wave mode moved predominantly westward, but the second mode was advected by the mean flow. Ezer (1999) modeled decadal variability of the upper North Atlantic in the subtropics, finding cold temperature anomalies propagating eastward and downward, possibly forced by the Gulf Stream.

We need to both understand the process by which features and waves propagate, and obtain good numerical predictions for that behavior. The situation for passive tracers is simple: they are advected along mean streamlines. We therefore limit attention to *dynamical* features henceforth. These are not advected along mean streamlines under almost any circumstances, since if q is some conserved quantity [e.g., potential vorticity, as used by Zhang and Liu (1999), or potential temperature, etc.] with mean value $\overline{q}$ and perturbation q', then q' satisfies

$$q'_t + \mathbf{u}' \cdot \nabla \overline{q} + \overline{\mathbf{u}} \cdot \nabla q' = 0 \qquad (1.1)$$

so that anomalies are not simply advected by $\overline{\mathbf{u}}$ (which would lie along $\overline{q} = \text{const}$). Any description of the anomaly propagation must solve (some version of) the full three-dimensional problem. There are implicit differences in the roles of the horizontal and vertical in this problem, due to the existence of wave propagation dependant upon the mean vertical structure.

Understanding the behavior of (1.1) has not proven easy. It is fully three-dimensional, so that use of a simplified numerical model (e.g., frictional planetary geostrophy, which can handle both mean flow and topography) or a full general circulation model, while simulating propagation of disturbances, does not yield easy *insight* into the dynamics of the disturbances. Such insight involves some simplification, particularly of the vertical aspect of the problem, without losing accuracy [though rays can be traced three-dimensionally, e.g., Yang (2000)]. The quasigeostrophic approximation requires a spatially uniform stratification, which is not useful if our aim, as here, is to describe propagation on basin scales. Simplifications such as 1.5- and 2-layer models give understanding but will not be reliable for calculations in the real ocean. Indeed, Flierl (1978) argues cogently that layered models need calibration before they yield useful results, and recommends the use of modal decomposition instead. This approach—but using vertical normal modes appropriate to the local stratification, and that vary slowly in the horizontal—will be employed here. The concept is basically WKBJ: the background circulation and stratification vary on the basin scale, while the (geostrophic) waves vary on scales much shorter than the background. The sole disadvantage of casting onto vertical normal modes is that the correct bottom boundary condition for varying topography cannot be handled with any accuracy using normal modes. In this paper we therefore use the flat-bottom boundary condition [cf. Killworth and Blundell (1999) for a proper treatment in the absence of mean flow].

This paper, then, attempts to produce a simple descriptor of the propagation of internal, baroclinic, large-scale disturbances against a background ocean in steady motion (using a flat-bottom condition but including basinscale depth variation). Section 2 details the normal mode decomposition, in which small perturbations are made to a system with a background mean baroclinic flow (any barotropic flow simply gives a Doppler shift). We show that the vector of normal modes is advected horizontally by a matrix of *pseudovelocities,* which are discussed in section 3. The normal-mode vector can be considered as a collection of (nonorthogonal) eigenvectors of the *shear modes.* The properties of the advective system are treated in section 4, which shows that the diagonal pseudovelocities give reasonable fits to the direction and speed with which each shear mode propagates, resulting in a great simplification of the problem. Each shear mode is shown, to a good degree of approximation, to propagate with a well-defined horizontal velocity which varies with horizontal position. Estimates of the first and second shear mode velocities are given in section 5. The first shear mode pseudoveloc-

ity remains stubbornly westward in all locations, but at latitudes above 30°, the second shear mode propagates in directions resembling, but not identical to, the mid-depth (500–1000 m) baroclinic background flow, and at somewhat slower rates.

2. Normal-mode decomposition

The equations of motion are now cast onto a single variable M, using Welander's (1959) notation. This uses the planetary geostrophic approximation, which permits variation of background stratification (unlike the quasigeostrophic approximation), though only for geostrophic flow. For simplicity only, the analysis is presented on a Cartesian plane. All data analysis and computations in this paper are, however, performed using full spherical geometry. In terms of M, we have

$$M_z = \frac{p}{\rho_0}; \qquad M_{zz} = -\frac{g\rho}{\rho_0}; \qquad u = -\frac{M_{zy}}{f};$$

$$v = \frac{M_{zx}}{f}; \qquad w = \frac{\beta}{f^2}M_x,$$

where the symbols have their usual meaning, and $\beta = f_y$ is not necessarily constant. When substituted into the conservation of density ($\rho_t + \mathbf{u} \cdot \nabla\rho = 0$) we obtain

$$M_{zzt} - \frac{M_{zy}M_{zzx}}{f} + \frac{M_{zx}M_{zzy}}{f} + \frac{\beta}{f^2}M_xM_{zzz} = 0. \quad (2.1)$$

As explained earlier, we can assume $M(-H) = 0$, which is the flat-bottom boundary condition corresponding to zero vertical velocity. (Recall that varying bottom depth cannot be attacked easily with normal modes.) We now pose a small perturbation M' to the mean flow $\overline{M}$. The latter is assumed a function of all three spatial coordinates; Pedlosky (1987) discusses the effects this may have on the small perturbations from the aspect of the instability problem. In what follows, both $\overline{M}$ and M' are assumed to vanish at the surface $z = 0$. The former will not be true in general, but the additional barotropic component can be added in later, as will be shown. The vanishing of M' at the surface corresponds to unforced propagation or free waves. The linearized system becomes

$$M'_{zzt} - \frac{1}{f}(\overline{M}_{zy}M'_{zzx} + M'_{zy}\overline{M}_{zzx}) + \frac{1}{f}(\overline{M}_{zx}M'_{zzy} + M'_{zx}\overline{M}_{zzy})$$

$$+ \frac{\beta}{f^2}(\overline{M}_xM'_{zzz} + M'_x\overline{M}_{zzz}) = 0. \quad (2.2)$$

We now make the fundamental approximation for this paper, standard for the WKBJ theory to be used later, that

$$L_{\text{basin}} \gg L_{\text{pert}} \gg a; \quad \text{or} \quad \frac{L_{\text{pert}}}{L_{\text{basin}}} = \varepsilon \ll 1, \quad (2.3)$$

where L_{pert}, L_{basin} are length scales for the perturbation

and the mean flow, respectively, and a is the deformation radius. All aspects of the mean flow (including the stratification) vary with length L_{basin}. The perturbation scale L_{pert} is, however, not small enough for the assumption of geostrophy to fail, by the second half of (2.3). The approximation (2.3) permits rather more progress than is usual within WKBJ theory. In practice—as with almost any scale separation theory—there are areas in the World Ocean in which the background flow changes sufficiently rapidly to invalidate (2.3), more usually in the north–south direction. However, experience with WKBJ theory suggests that such theory continues to give useful results even when (2.3) is not well satisfied. In the wave propagation cases considered here, the dependence of east–west phase speeds on latitude means that short north–south scales develop in the solution, helping to mitigate the more rapid north–south variation of the mean flow.

In (2.2), then, the first term in the last parentheses is formally smaller than the second by an amount $O(\varepsilon)$. Henceforth, this term is omitted; its inclusion makes negligible changes to the results.[1] Another, more useful, way to write (2.2) is

$$M'_{zzt} + \overline{\mathbf{u}}_H \cdot \nabla M'_{zz} - \overline{\mathbf{u}}_{Hz} \cdot \nabla M'_z + \frac{\beta N^2}{f^2}M'_x = 0, \quad (2.4)$$

where the suffix H denotes the horizontal component, and we have written $\overline{M}_{zzz} = N^2(x, y, z)$. Due to geostrophy, (2.4) is equivalent to the long-wave assumption.

We now seek to simplify the fully stratified problem above. This is frequently done using a layered decomposition. However, Flierl (1978) pointed out that casting onto normal modes gave a considerably more accurate representation of dynamics, at least in the quasigeostrophic regime he considered. Accordingly, we cast (2.4) onto the normal modes for a background density stratification given by a buoyancy frequency $N^2(x, y, z)$. The normal modes are

$$\hat{h}_j(x, y, z): \quad \hat{h}_{jzz} + \frac{N^2}{C_j^2}\hat{h}_j = 0; \qquad \hat{h} = 0,$$

$$z = 0, \; -H(x, y). \quad (2.5)$$

Here C_j is the internal wave speed of the jth eigenmode. The equivalent basis function for horizontal velocity is $\hat{u}_i(x, y, z) = \hat{h}_{iz}$. The $\hat{h}$, $\hat{u}$ are normalized at all points by

$$\int_{-H}^0 N^2\hat{h}_i\hat{h}_j \, dz = HC_i^2\delta_{ij}; \qquad \int_{-H}^0 \hat{u}_i\hat{u}_j \, dz = H\delta_{ij} \quad (2.6)$$

so that $\hat{h}$ has units of length and $\hat{u}$ is dimensionless. The

---

[1] The neglected term is in fact much smaller than the simple scaling argument given would suggest; Dewar (1998) provides rationale for this.

sign convention is such that all the $\hat{u}$ are positive at the surface. Progress will be possible because the $\hat{h}$, $\hat{u}$ vary only *slowly* across the basin with scale L_{basin}, while the perturbation fields vary on the more rapid scale L_{pert}.

We expand the background (now only horizontal) mean velocity in the $\hat{u}$ and the M' in the $\hat{h}$:

$$\overline{\mathbf{u}}_H = \sum_{k=1}^{\infty} \mathbf{u}_k(x, y)\hat{u}_k(x, y, z);$$

$$M' = \sum_{k=1}^{\infty} M'_k(x, y, t)\hat{h}_k(x, y, z) \quad (2.7)$$

(recall that we can add in the mean barotropic component later, so the first sum starts from 1).

The mean quantities can easily be calculated from data. Assuming $\overline{\mathbf{u}}_H$ is known from density data and the use of the thermal wind relationship, the orthogonality condition implies

$$\mathbf{u}_k = \frac{1}{H} \int_{-H}^{0} \overline{\mathbf{u}}_H \hat{u}_k \, dz. \quad (2.8)$$

We now note that all horizontal and time derivatives in (2.4) act on the perturbation quantity M'. From the second part of (2.7), we have, for example,

$$M'_x = \sum_{k=1}^{\infty} \{M'_{kx}\hat{h}_k + M'_k \hat{h}_{kx}\}.$$

In this sum, the first terms involve derivatives on the rapid (L_{pert}) scale, while the second involve derivatives on the slower (L_{basin}) scale. Accordingly, the derivatives of M' *dominate* over those involving horizontal derivatives of the basis functions $\hat{h}$, so that the latter may be neglected. In other words, casting onto *local* modes and neglecting their slow lateral variation is legitimate within the scale assumption (2.3).

Substituting (2.7) into (2.4) gives, neglecting terms as just demonstrated,

$$-\sum_j M'_{jt}\frac{N^2}{C_j^2}\hat{h}_j + \sum_i \sum_j \left\{-u_i\hat{h}_{iz}M'_{jx}\frac{N^2}{C_j^2}\hat{h}_j - v_i\hat{h}_{iz}M'_{jy}\frac{N^2}{C_j^2}\hat{h}_j\right\}$$

$$+ \sum_i \sum_j \left\{u_j\frac{N^2}{C_j^2}\hat{h}_j M'_{ix}\hat{h}_{iz} + v_j\frac{N^2}{C_j^2}\hat{h}_j M'_{iy}\hat{h}_{iz}\right\}$$

$$+ \frac{\beta}{f^2}\sum_j M'_{jx}\hat{h}_j N^2 = 0, \quad (2.9)$$

where (2.5) has been used for second derivatives of $\hat{h}$. Multiplying by $\hat{h}_k$ and integrating top to bottom, we obtain, using orthogonality and multiplying by $-1/H$,

$$M'_{kt} + \sum_i \sum_j \frac{\alpha_{ijk}}{C_j^2}\{u_i M'_{jx} + v_i M'_{jy} - u_j M'_{ix} - v_j M'_{iy}\}$$

$$- \frac{\beta}{f^2}C_k^2 M'_{kx} = 0. \quad (2.10)$$

Here the triple interaction coefficient

$$\alpha_{ijk} = \frac{1}{H} \int_{-H}^{0} N^2 \hat{h}_{iz}\hat{h}_j\hat{h}_k \, dz \quad (2.11)$$

is a dimensional equivalent to that defined in KCdS. Note that α is symmetric in its last two suffixes. Flierl (1978) used similar coefficients in the quasigeostrophic regime.

The last term in (2.10) is simply westward advection of mode k by the kth internal planetary wave speed $\beta C_k^2/f^2$. The remainder of the equation can be further simplified by defining the dimensionless interaction coefficient

$$\gamma_{ijk} = \frac{\alpha_{jki}}{C_k^2} - \frac{\alpha_{kji}}{C_j^2}, \quad (2.12)$$

which is antisymmetric in its last two suffixes. Then we write the sum

$$\mathbf{u}_{kn}^{\text{eff}} = \sum_m \mathbf{u}_m \gamma_{kmn} \quad (2.13)$$

and substitute into (2.10), obtaining

$$M'_{kt} + \sum_n \{u_{kn}^{\text{eff}}M'_{nx} + v_{kn}^{\text{eff}}M'_{ny}\} - \frac{\beta}{f^2}C_k^2 M'_{kx} = 0. \quad (2.14)$$

By defining

$$u_{kn}^{\text{net}} = u_{kn}^{\text{eff}} - \frac{\beta}{f^2}C_k^2\delta_{kn}; \qquad v_{kn}^{\text{net}} = v_{kn}^{\text{eff}} \quad (2.15)$$

(2.14) becomes

$$M'_{kt} + \sum_n \{u_{kn}^{\text{net}}M'_{nx} + v_{kn}^{\text{net}}M'_{ny}\} = 0, \quad (2.16)$$

which is a coupled set of advective equations for the components M'_k.

Thus, the vector of modal contributions M'_k is advected by a set of pseudovelocities, obeying the simple (horizontal) advection rule

$$\mathbf{M}_t + \mathbf{U}^{\text{net}} \cdot \boldsymbol{\nabla}\mathbf{M} = 0, \quad (2.17)$$

where $\mathbf{M}$ is the column vector M'_k and

$$\mathbf{U}^{\text{net}} = [u_{kn}^{\text{net}}, v_{kn}^{\text{net}}]. \quad (2.18)$$

It is of interest to examine a single mode, R, say. From (2.16), M'_R is advected by all internal modes. In particular it is self-advected through $\mathbf{u}_{RR}^{\text{eff}}$. From (2.13), this is $\Sigma_i \mathbf{u}_i \gamma_{RiR}$. The antisymmetry in the last two components of γ shows that the only background velocity mode that does *not* contribute to this sum is the Rth mode itself. This is another demonstration of the non-self-interaction theorem for geostrophic normal modes, first shown for continuous stratification by KCdS; see also Liu (1999b) for a discussion in the finite layered system.

Addition of the background barotropic mode, if required, is now straightforward. Since it has no density signature, it merely adds an extra term $\mathbf{u}_B\delta_{kn}$ to the $\mathbf{u}_{kn}^{\text{eff}}$ term, where $\mathbf{u}_B$ is the depth-independent velocity, obtained by some other approach, for example, from a global model.
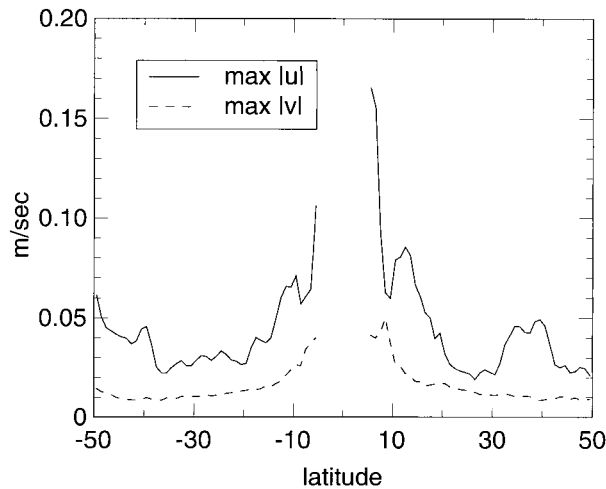
FIG. 1. The zonal average of the vertical maximum of $|u|$ and $|v|$ as functions of longitude. At all locations the maximum of $|u|$ is much larger than that for $|v|$. Values within 5° of the equator are not shown.

The sums above are formally infinite. In practice, there will be truncation. This can happen in two independent ways. First, the number of modes K included in the sum (2.13) to obtain $\mathbf{u}^{\text{eff}}$ is chosen. Second, the number of terms in the sum (2.17) for the evolution equation, L, say, can be chosen such that $L \leq K$; for example, one could imagine computing $\mathbf{u}_{11}^{\text{eff}}$ using many modes but only retaining that single term in (2.17) as an approximant.

3. Computation of the pseudovelocities

The pseudovelocities $\mathbf{u}_{kn}^{\text{net}}$, $\mathbf{u}_{kn}^{\text{eff}}$ were computed from annual-mean temperature–salinity data for the World Ocean (Levitus and Boyer 1994; Levitus et al. 1994) on a 1° horizontal grid. These were analyzed to produce local values of N^2, u_z, and v_z as functions of depth at the standard depths. The baroclinic u and v are then obtained by integration plus the requirement of no net vertically integrated horizontal flow. The precise location of the ocean floor was taken from an averaging of the 5-minute gridded Earth topography data set, ETOPO5 (National Geophysical Data Center 1988), onto the same grid.

The no-mean-flow eigenproblems are solved as matrix eigenvalue problems, ordered by westward phase speed. In most cases, at least the first four eigenvalues and vectors are retained in what follows.

a. The background velocity modal loadings $\mathbf{u}_i$

These computations, en route, demonstrated a fact that to our knowledge has not been noted previously. At each horizontal location, the minimum and maximum of u and v were computed within the fluid column. Because the flow is defined to be baroclinic, the mini-

mum (maximum) is negative (positive). For each flow direction, the minimum and maximum were combined to give the maximum modulus of the flow in that direction. The average of these moduli around a latitude band is plotted in Fig. 1. Throughout the paper, we concentrate on the latitude range ±50°, since turning latitudes for planetary waves of annual frequency are of this order. The zonally averaged speeds are a reasonable proxy for the full fields over most of the deep ocean.

The north–south flow is seen to be much weaker than the east–west flow on average at all latitudes. It is also smaller in 89% of the 1° × 1° ocean squares between ±5 and ±50°, and indeed the maximum modulus north–south flow is less than half the east–west equivalent in two-thirds of these squares. (The north–south flow is larger within a few degrees of eastern and western boundaries in most oceans, and in a few isolated midlatitude locations in the North Atlantic and Southern Oceans. This ignores the main western boundary currents, since these are sited over shelf-slope regions. We have excluded such regions, using a cutoff depth of 1000 m, from this essentially deep ocean theory because the assumptions underlying a modal treatment become inappropriate.)

The reason for mainly weaker north–south flows is the relative width of most ocean basins compared with the north–south distance over which surface density changes owing to atmospheric forcing. The atmosphere itself is, of course, an example in which neglect of x variation can be a reasonable approximation. While this cannot be exactly true in the ocean, theories have been worked out for zonally invariant basins (e.g., Welander and Liu 1976).

If v is in some sense "small" over most of the ocean compared with u, then the majority of advective-influenced flows will tend to be oriented east–west, so that propagation of disturbances around a gyre is made more difficult. However, though v is mainly small, the extended width of typical midlatitude basins means that even a small north–south velocity can be associated with significant latitudinal migrations under suitable circumstances.

The relative smallness of v is also indicated in Fig. 2, which shows the longitudinal average of the modal decomposition of the background u and v fields for the first four modes. The values are clearly dominated by u_1 (recall that this does not yet include the planetary wave β term), with values of order 1 cm s^{-1} in most regions, becoming largest in the Antarctic Circumpolar Current and near the equator. The values of u_2 and u_3 are much smaller than u_1; recall that KCdS showed that it was the second internal background east–west mode—u_2 in this notation—that was responsible for much of the increase in planetary wave speed observed, through its interaction with the propagating mode 1. Although this point will be discussed in more detail below, it should be noted that the *sign* of u_1 in higher midlatitudes
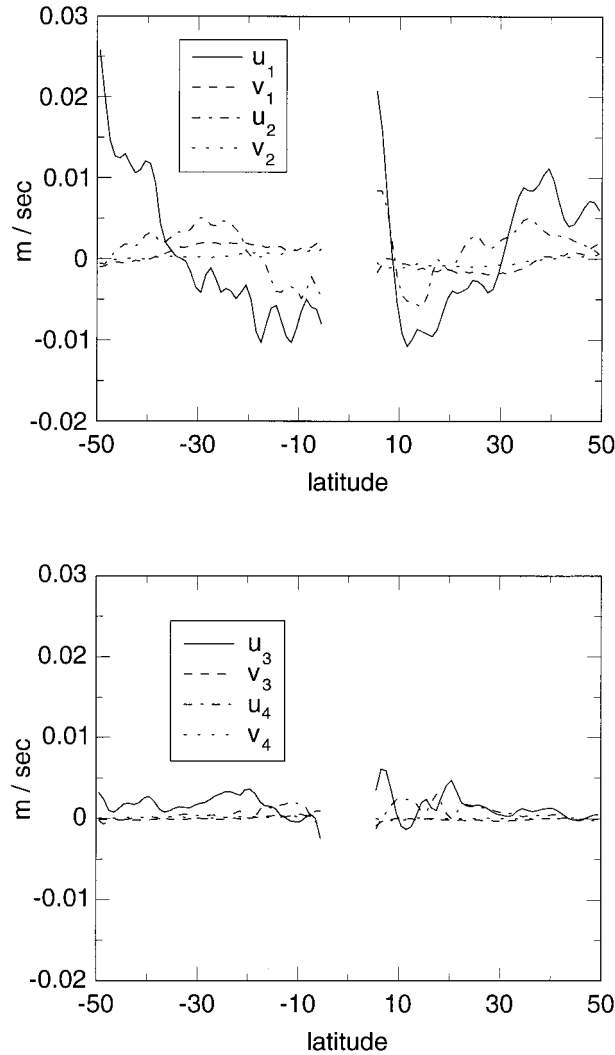
FIG. 2. The zonal average of modal coefficients (a) u_1, v_1, u_2, v_2 and (b) u_3, v_3, u_4, v_4 as functions of latitude. Values within $5°$ of the equator are not shown.

is positive; it will turn out that this is in competition with the westward propagation of mode 2, permitting a second mode to move more freely north–south.

The coefficients v_1, v_2 take the signs given from east–west thermal wind applied to warm poleward western boundary currents and colder equatorward interior flow. The higher modes are uniformly small, so that the truncation of the total number of modes computed (K in the above notation) can be at a small value—we shall typically use 2—without loss of accuracy. Put another way, the values of u_{kn}^{net}, v_{kn}^{net} are mainly independent of K, providing $K \geq 2$.

b. The values of the pseudovelocities $\boldsymbol{u}_{kn}^{net}$

We turn now to consideration of the pseudovelocities themselves. From (2.13), these are sums of the $\mathbf{u}_i$, weighted by the γ coefficients. These latter are com-

TABLE 1. Values of $\mathbf{U}^{net}$ (cm s^{-1}) at $35.5°$N, $150.5°$E (North Pacific). The bold items are values that should be theoretically small in most places. For shear mode 1, the east–west phase speed computed from the vertically stratified problem is -3.7 cm s^{-1}, and that from the 4 $\times$ 4 matrix (computed with $K = 4$) given is also -3.7 cm s^{-1}. The eigenvalue from the 2 $\times$ 2 submatrix is -3.3 cm s^{-1}. The mode 1 and 2 loadings for the eigenvector are 0.99 and 0.13, respectively. For shear mode 2, the stratified phase speed is -1.3 cm s^{-1}, and that from the 4 $\times$ 4 matrix is -0.8 cm s^{-1}. The eigenvalue from the 2 $\times$ 2 submatrix is -0.6 cm s^{-1}. The mode 1 and 2 loadings for the eigenvector are 0.82 and 0.56, respectively.

u_{ij}^{net}: i/j	1	2	3	4
1	−5.2	7.2	6.1	1.2
2	**−1.2**	1.3	5.8	3.3
3	**−0.8**	**0.8**	3.0	5.7
4	**−0.3**	**−0.1**	**1.8**	4.2
v_{ij}^{net}: i/j	1	2	3	4
1	−0.2	−0.6	−0.1	−0.2
2	**−0.0**	−0.3	−0.5	0.1
3	**−0.0**	**−0.1**	−0.3	−0.4
4	**−0.0**	**0.0**	**−0.1**	−0.4

puted in situ from the vertical structure, but in fact do not differ strongly from place to place. If an approximate WKBJ form is used for the vertical normal modes (cf. KCdS, p. 1963, for the first mode; the others are similar) then it is straightforward to show that (i) γ_{abc} is only sizeable if in some order $a = b + c$, and (ii) γ_{abc} decreases as min (a, b, c) increases. Point (i) is because the stretched sinusoidal nature of the eigenvectors means that the triple integrals to obtain α_{ijk}, and hence γ_{ijk}, have reasonable nonzero values only when sums and differences of sinusoids interact. Point (ii) is because the increasingly oscillatory nature in the vertical as the mode number increases means that the integrals naturally decrease. These remarks hold reasonably well for low orders.

Combining this information with the previous findings that high modal loadings $\mathbf{u}_i$ are small puts strong restrictions on the coefficients $\mathbf{u}_{kn}^{net}$. The diagonal east–west terms will be large due to the β term, though this decreases rapidly with mode number. In particular, the first column below the diagonal will generally be small, since $\mathbf{u}_{k1}^{eff}$ is $\Sigma_m \mathbf{u}_m \gamma_{km1}$. The $m = 1$ term in this sum vanishes since γ_{k11} is zero. The remaining terms only involve $\mathbf{u}_m$, $m > 1$ and so are smaller in magnitude. The second column above the diagonal ($\mathbf{u}_{12}^{eff}$) can be a reasonable size since it involves $\gamma_{112}\mathbf{u}_1$. The remainder of the second column, below its diagonal, is mainly small for similar reasons, save for $\mathbf{u}_{32}^{eff}$, which involves $\gamma_{312}\mathbf{u}_1$, neither term of which is necessarily small.

Similar arguments extend to higher rows and columns: values *above* the diagonal can have reasonable sizes, while those below tend to be small. Two typical instances are shown in Tables 1 and 2, which show the 4×4 $\mathbf{u}_{kn}^{net}$ coefficients (i.e., including the planetary wave diagonal term) for locations in the North Pacific and

North Atlantic. The sizes are in good agreement with the arguments presented here.

In the particular case of a 2×2 truncation, u_{kn}^{eff} becomes simply

$$\begin{bmatrix} -\gamma_{112}u_2 & \gamma_{112}u_1 \\ -\gamma_{212}u_2 & \gamma_{212}u_1 \end{bmatrix},$$

where the γ coefficients shown are both positive. Over latitudes between ± 5 and $\pm 50°$, on average $\gamma_{112} = 2.54$, and $\gamma_{212} = 0.83$. Note that the u_{22}^{eff} term is proportional to u_1. This is positive at mid- and high latitudes, as noted previously. As the size of the planetary wave correction decreases with distance poleward, at some latitude it may be possible for u_{22}^{eff} to become positive; that is, the background flow may be sufficient to alter the propagation velocity for the second mode from westward, its natural direction, to eastward.

These predictions are confirmed by Fig. 3, which shows longitudinal averages of the (vector) 2×2 components of the $\mathbf{U}^{\text{net}}$ matrix.[2] [There is no ideal way to show these data due to the large variation in size. An alternative (Fig. 4) shows longitudinally averaged *moduli* of the (vector) 2×2 components of the $\mathbf{U}^{\text{net}}$ matrix. Moduli are used since the strong planetary wave component makes the diagonal u terms dominant save at high latitudes; the strong value of u_{11}^{net} near the equator requires a logarithmic scale to show detail at higher latitudes.] Either presentation shows again that the above-diagonal terms are not negligible, while the below-diagonal—here only the (2,1) term—tends to be small. As in the modal casting of the background velocities, u dominates v almost everywhere. As predicted, poleward of about 35°, u_{22}^{eff} is positive (though only a few millimeters per second in the Northern Hemisphere), raising the possibility that anomalies may be able to propagate with an eastward, rather than westward, component.

4. Shear modes

The behavior of the advective modal system (2.17), or of the original continuously stratified problem (2.2), is generally complicated. For the moment, we write M_i as a horizontal Fourier decomposition with wavenumber $\mathbf{k} = (k, l)$ and frequency ω:

$$M_i = \iint a_i(k, l)e^{i(kx+ly-\omega t)} \, dk \, dl. \quad (4.1)$$

If the velocity field were horizontally uniform, or if using a WKBJ formulation, assuming as before that the background fields vary on the length scale $L_{\text{basin}} \gg L_{\text{pert}}$, then each Fourier mode would satisfy an eigenvalue relationship

---

[2] These terms were computed using a four-mode expansion, but almost identical answers occur with two, three, or more than four modes included.

TABLE 2. Values of $\mathbf{U}^{\text{net}}$ (cm s^{-1}) at 40.5°N, 35.5°W (North Atlantic). For shear mode 1, the east–west phase speed computed from the vertically stratified problem is -1.2 cm s^{-1}, and that from the 4×4 matrix (computed with $K = 4$) given is -1.1 cm s^{-1}. The eigenvalue from the 2×2 submatrix is -0.9 cm s^{-1}. The mode 1 and 2 loadings for the eigenvector are 0.98 and 0.20, respectively. There is no real stratified phase speed for shear mode 2. That from the 4×4 matrix is -0.8 cm s^{-1}. The eigenvalue from the 2×2 submatrix is $+0.3$ cm s^{-1}. The mode 1 and 2 loadings for the eigenvector are 0.93 and 0.37, respectively.

u_{ij}^{net}: i/j	1	2	3	4
1	−1.7	2.2	2.1	0.2
2	**−0.4**	−0.0	1.7	0.4
3	**−0.3**	**0.4**	1.0	1.9
4	**−0.1**	**−0.1**	**0.6**	1.3
v_{ij}^{net}: i/j	1	2	3	4
1	−0.1	−0.4	−0.1	−0.2
2	**0.1**	−0.0	−0.4	0.1
3	**0.1**	**−0.1**	−0.3	−0.4
4	**0.1**	**0.0**	**−0.1**	−0.3

$$(\mathbf{k} \cdot \mathbf{U}^{\text{net}})\mathbf{M} = \omega\mathbf{M}, \quad (4.2)$$

where $\mathbf{M}$ represents the vector M_i'. Modes that satisfy (4.2) we term "shear modes," oriented along (k, l), using Killworth and Anderson's (1977) terminology. They are the natural vertical structures (or, here, modal structures) that can propagate locally. Equation (4.2) is an implicit dispersion relation for the waves. To be solved, the number of modes would be truncated. If the decomposition (2.16) is truncated at L modes (with $L \leq K$), the eigenvalue relation becomes the matrix relation

$$\sum_{j=1}^{L} (ku_{ij}^{\text{net}} + lv_{ij}^{\text{net}})M_j' = \omega M_i', \quad i = 1, 2, \cdots, L. \quad (4.3)$$

Since the matrix operator in (4.3) is not symmetric, the shear modes are not orthogonal. In the continuously stratified problem, we set $M' = F(x, y, z) \exp i(kx + ly - \omega t)$, where F is a slowly varying function of the horizontal, and obtain

$$(ku + lv - \omega)F_{zz} - (ku_z + lv_z)F_z + \frac{\beta N^2}{f^2}kF = 0 \quad (4.4)$$

together with boundary conditions

$$F(0) = F(-H) = 0 \quad (4.5)$$

since we restrict attention to a flat bottom. The system [(4.4), (4.5)] is not self-adjoint, though progress can be made, as will be shown elsewhere. The solutions of KCdS are the first shear mode for the continuous problem, oriented east–west.

Both systems can yield real or complex eigenvalues, the latter corresponding to long-wave instability (cf. KCdS; Liu 1999a for discussions). Over the World Ocean between $\pm(5–50)°$, a negligible 0.08% of 1° squares possess no real shear mode solutions to the 4×4 matrix problem, and 27% possess only two real
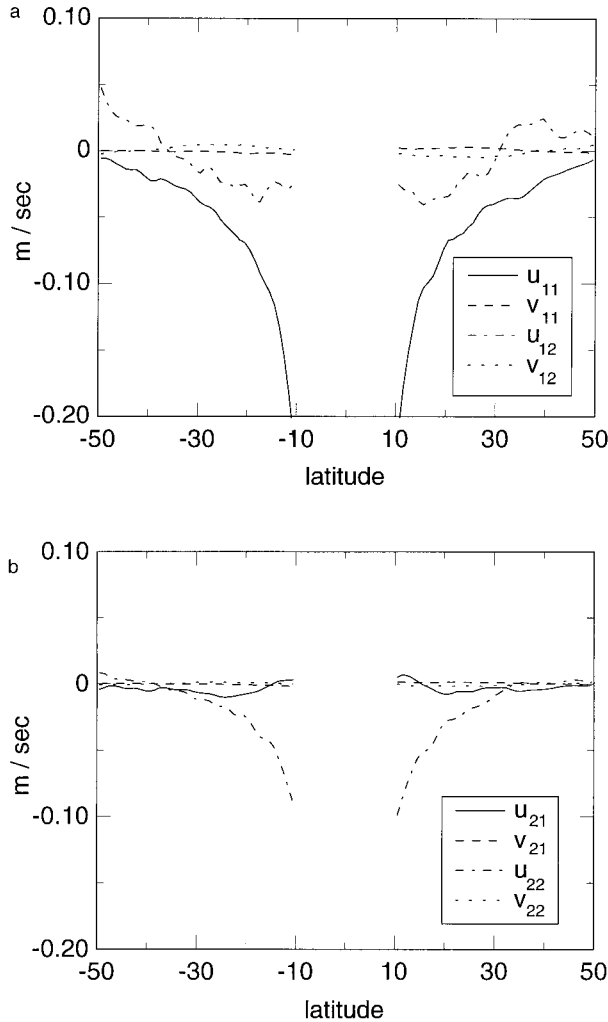
FIG. 3. Zonal averages of pseudovelocities u_{kn}^{net}, v_{kn}^{net} for (a) k, n = 1,1; 1,2 and (b) k, n = 2,1; 2,2. Values within $10°$ of the equator or above 0.20 m s^{-1} in magnitude are not shown.

FIG. 4. The zonal average of $|u_{kn}^{net}|$, $|v_{kn}^{net}|$ for (a) k, n = 1,1; 1,2 and (b) k, n = 2,1; 2,2. A logarithmic scale is used so as to show high-latitude detail that would otherwise be swamped by near-equatorial values. Values within $5°$ of the equator are not shown.

shear mode solutions (the remainder of the squares, 73%, having the full four real solutions). (The 3×3 solutions yield similar figures, e.g., 82% squares with all three real solutions. The 5×5 solutions are again similar, though now 99% of the solutions have at least three real shear modes, so that some details of the reality of the solutions depend on the truncation used.) Although around 28% of $1°$ squares possess some complex solutions, with $L \geq 4$ there remain at least two real solutions (corresponding to the gravest shear modes) in these locations. So at least the first two shear modes can propagate throughout the ocean, but higher shear modes will move into regions where there potentially is growth or decay at small wavenumbers. Such areas are not easily treated with WKBJ analysis, since there will typically be caustics surrounding the areas of complex solutions, and linkage techniques must be employed to cross the real–complex boundary.
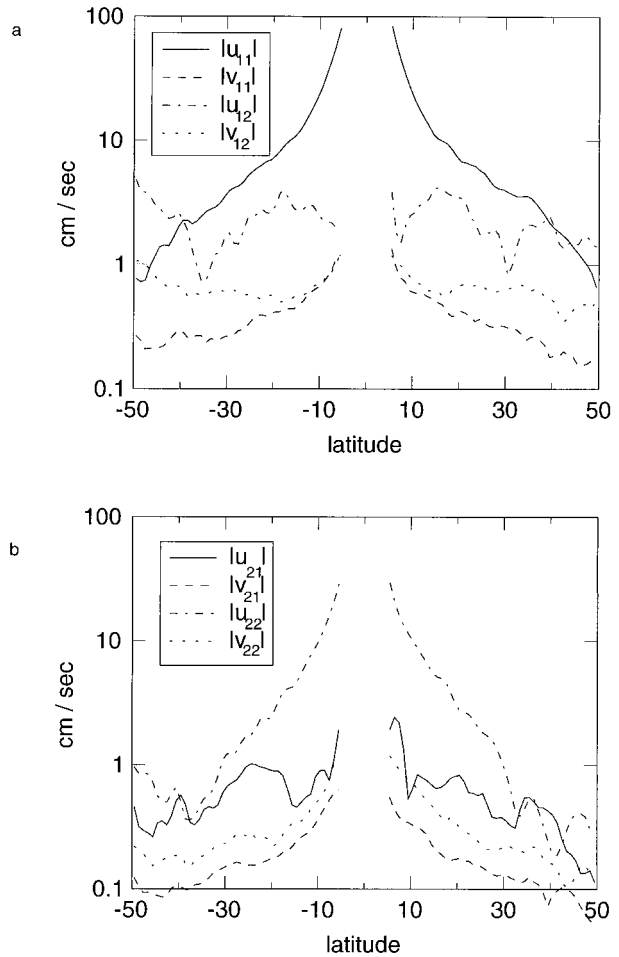
The small number of complex modes does, however, remain interesting. A referee raised the question of whether the number might be underestimated because of a singular limit, due to the order in which the limits are taken in the usual baroclinic instability problem (the background relative vorticity is neglected, and the perturbation wavenumber is set to zero for the geostrophic problem). While we have not investigated this fully, we do not believe that any singular limit can occur. First, the limit that neglects background relative vorticity can occur independently of the perturbation length scale being large compared with a (in our case, this neglect is doubly justified because the background flow varies on the basin scale only, so that the relative vorticity, of order U/L_{basin}^2, is negligible compared with β, which is retained). Second, Fu and Chelton (2001) have computed the non-long-wave solutions for a background east–west flow and found that the problem is fully continuous in the geostrophic limit.

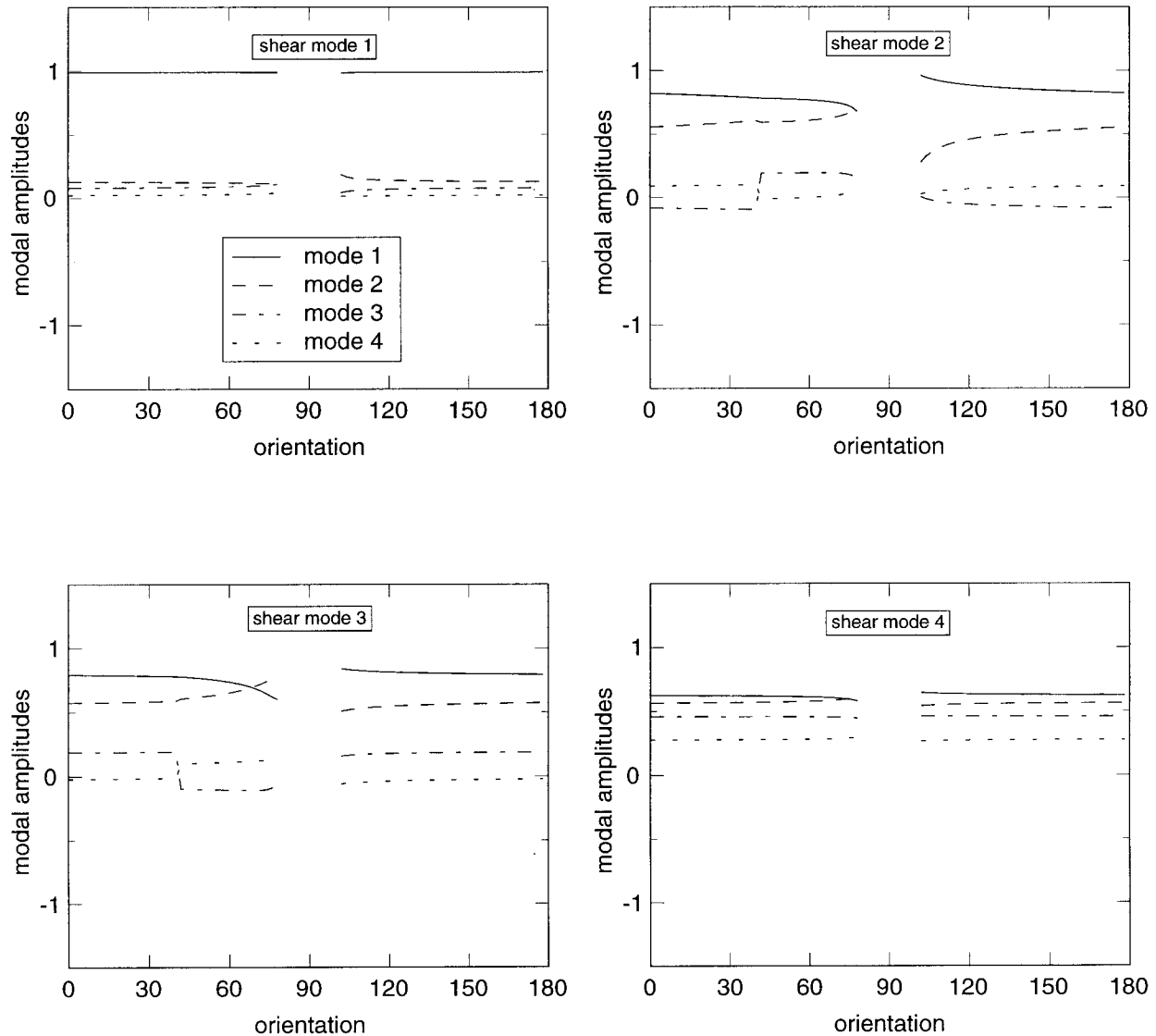The areas of complex solutions for third and fourth

FIG. 5. The variation of modal amplitudes present in the first four shear modes at 35.5°N, 150.5°E, as the orientation of the wavenumber is altered. The values were computed at multiples of 2°; between 80° and 100° at least one shear mode is not present, and so none are shown.

shear modes in the 4 × 4 matrix case are mainly located at high latitude (specifically in the Antarctic Circumpolar Current) and with occasional latitudinal bands (e.g., around 20°N in the Pacific). In the 5 × 5 case, the areas of two unstable shear modes are similar; however, the third shear mode is now stable there. We shall concentrate almost exclusively on the first two shear modes henceforth, so that complex solutions are not an issue.

The continuous problem [(4.4), (4.5)] in general has a finite number of solutions, the number decreasing as the magnitude of the mean flow increases. The decomposition approach, however, always possesses L solutions. This is a familiar problem. For example, if one poses the one-dimensional continuous Eady problem, which has two solutions, *either* as a finite-difference

problem with K points *or* by casting it onto K internal modes, then K apparent solutions will be found, all but two of which are numerical in nature.[3] Thus, we expect that not all the modal behavior—or by analogy, the behavior of an K-layered system—will be a good approximation to the behavior of the original stratified system. Appendix A gives a specific, though not realistic, example (when the east–west wavenumber $k = 0$) in which the modal decomposition gives incorrect solutions.

At any event, the propagation of a large feature involves more complexity than the simple (4.3). This eigenvalue relation clearly depends on the *orientation* of

---

[3] The continuous Eady problem has delta-function solutions, discussed by Pedlosky (1964), but these cannot be represented by discrete numerical solutions.
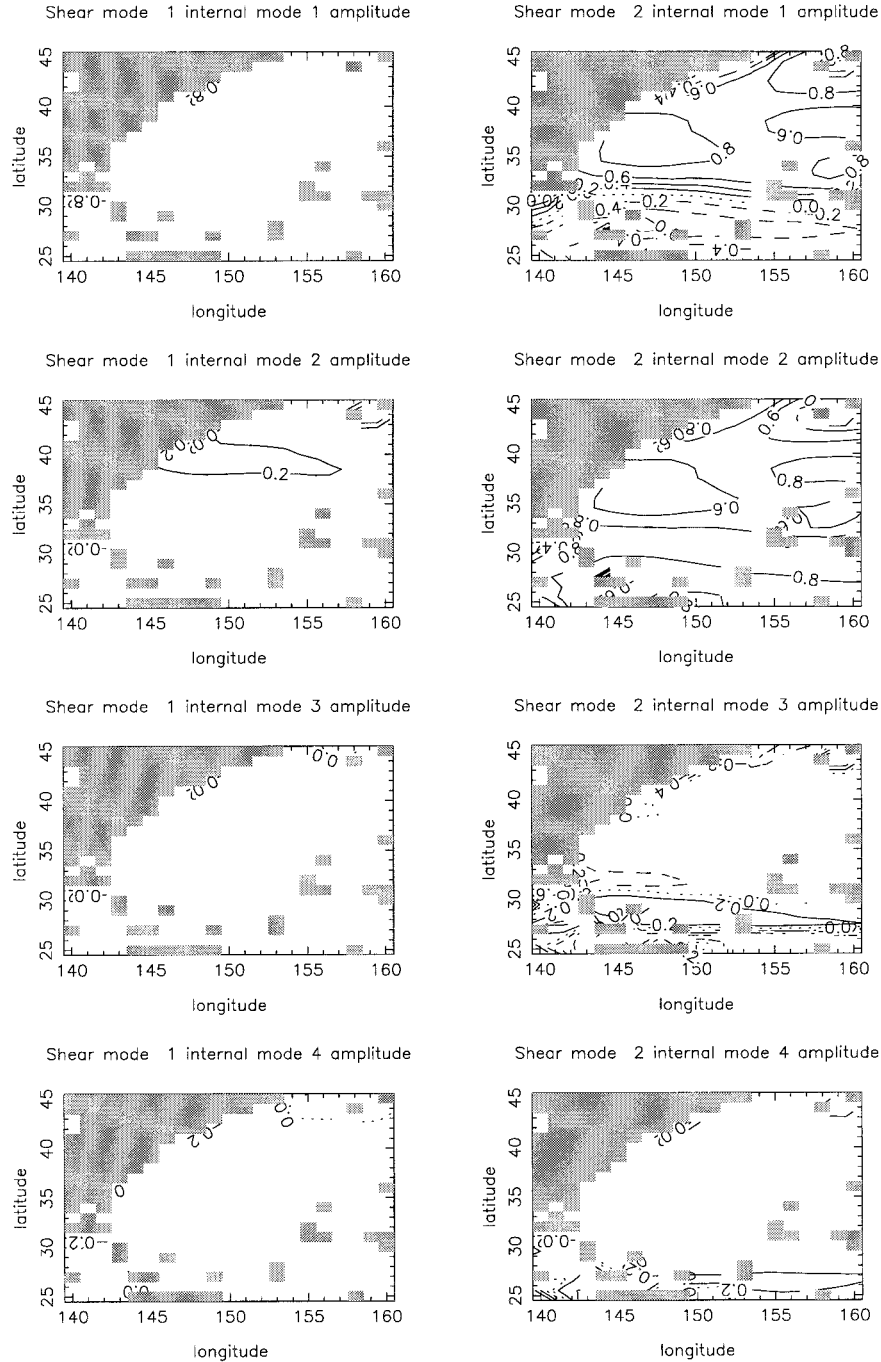
Shear mode 1 internal mode 1 amplitude

Shear mode 2 internal mode 1 amplitude

Shear mode 1 internal mode 2 amplitude

Shear mode 2 internal mode 2 amplitude

Shear mode 1 internal mode 3 amplitude

Shear mode 2 internal mode 3 amplitude

Shear mode 1 internal mode 4 amplitude

Shear mode 2 internal mode 4 amplitude

FIG. 6. The modal amplitudes of the first two shear modes in the Northwest Pacific, with east–west orientation, as determined from a four-mode decomposition. The contour interval is 0.2; negative contours dashed, zero contour dotted. Gray areas denote land (in the northwest corner) or regions where a real (stable) shear mode cannot be found.

the wave vector. It does not depend, other than a simple rescaling of ω, on its magnitude, due to the original long-wave assumption, a feature shared by the continuous problem. Even if the entirety of a feature were describable by a single vertical structure, this structure would not necessarily be the eigenstructure for all wavenumber ori-

entations and so would tend to spread as it propagated. Thus, whether a feature can maintain its shape depends on how sensitive the eigenvector solutions to (4.3) are to orientation of the wavenumber vector.

Figure 5 shows how the loading of the first four individual modes in shear modes 1–4 [i.e., the eigenvector
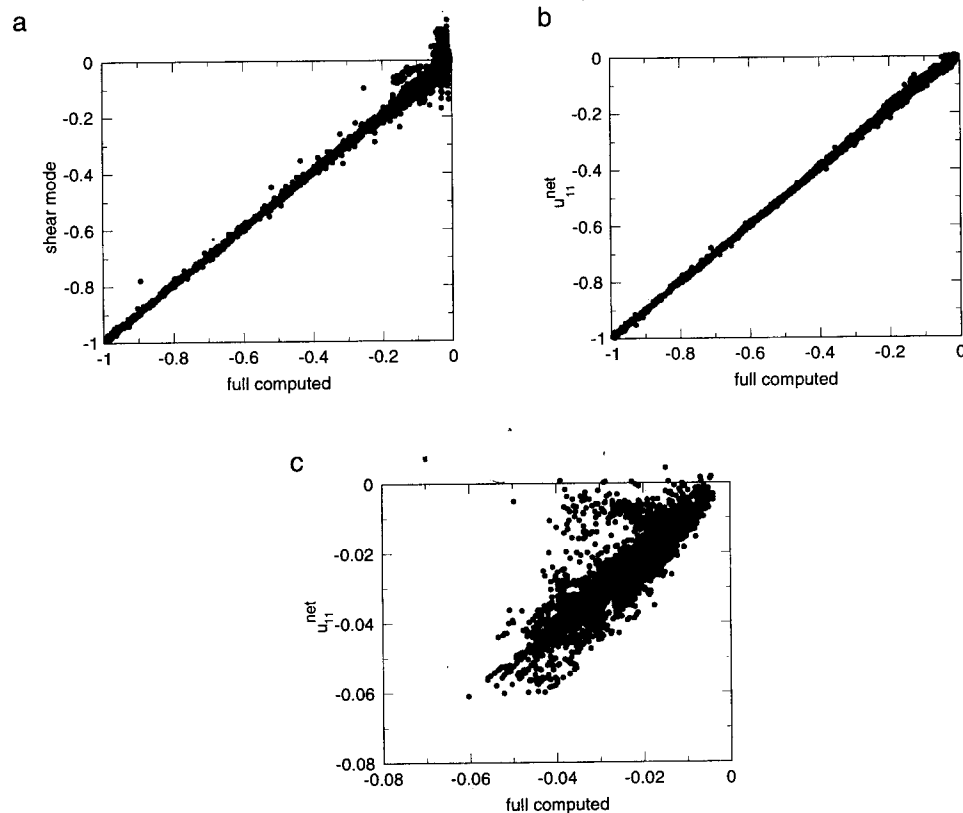
FIG. 7. Comparisons of east–west oriented phase speeds (m s^{-1}) of shear mode 1 between the full continuous local vertical calculation, the shear mode based on decompositions using two internal modes (that with three or more is indistinguishable), and the simple identification of u_{11}^{net} with shear mode 1. The scatterplots are over all (global) points with real solutions in the latitude range $\pm(5°–50°)$. (a) The full calculation against the shear mode. (b) The full calculation against u_{11}^{net}. (c) Replots of (b), limiting attention to smaller phase velocities by restricting the latitude range to $\pm(30°–50°)$.

solutions to (4.3)] depends on the orientation of the wavevector at one location; the figure is typical of most midlatitude locations. Solutions are presented in terms of an *eigenspeed* $\omega/|\mathbf{k}|$, where $\mathbf{k} = |\mathbf{k}|(\cos\phi, \sin\phi)$, rather than the less easily interpreted frequency ω. We solve (4.3) using four modes as ϕ varies between 0° and 180° (beyond 180° the sign of the problem is merely changed). The shear modes are ordered on the eigenspeed from the most negative, so that shear mode 1 at orientation 0° would be the fastest westward-propagating mode as discussed for the continuous problem by KCdS.

The results show that the loadings of individual modes are remarkably invariant to orientation in this long-wave limit, other than occasional mode switches (e.g., the shear mode 2 and 3 exchange at around 40° orientation in the example). Shear mode 1 is dominated almost exclusively by internal mode 1 at all orientations, showing that the approximation used by KCdS (p. 1962) of replacing the shear mode by the internal mode was valid. Internal mode 1 is still mainly dominant in the second and third shear modes, although internal mode

2 becomes of similar importance. All four internal modes used have similar loadings in shear mode 4.

Near 90° orientation, the number of modes decreases rapidly, and so values are not shown between 80° and 100° orientation.

Appendix B gives approximate arguments based on the smallness of v compared with u that indicate that this invariance to orientation should be expected for both pseudovector and continuous representations.

The second factor that might act to prevent coherent propagation is if the shear mode changed its shape rapidly in the horizontal. (Within a WKBJ framework this cannot happen: the shape is part of the slowly varying structure. However, a large feature, to be recognizable as such from data, needs to be horizontally coherent across L_{pert}.) Figure 6 examines how the internal mode loadings (for zero, i.e., eastward, orientation) vary with position in the Northwest Pacific for the first two shear modes. Clearly shear mode 1 is universally dominated by internal mode 1 as before, so that there is weak variation with position of the modal loadings. At lower latitudes there is evidence of some rapid modal loading

changes in shear mode 2 (and higher shear modes, not shown), but north of about 32° there is little change in modal loadings, even in these shear modes. Thus, at such latitudes it is reasonable to expect that, to a good degree of approximation, a single shear mode can be coherent over the scale of a feature and can propagate following the solution of (4.3) while slowly changing its vertical shape in a WKBJ-like fashion.

5. Strongly truncated representations

Starting with the continuous system (2.2), the problem has reduced successively to an infinite or finite matrix advection system for a collection of modes, and

then to an eigenvalue problem for propagation at a specific angle. At all times through this procedure numerical solutions could be sought, and will be shown in a companion paper. In terms of understanding the propagation of baroclinic features, however, even the eigenvalue problem is not of itself enlightening. We therefore seek further simplifications, by *truncating* the matrix system heavily.

Such truncations give worthwhile results because of the arguments in section 3. There it was argued that diagonal terms in $\mathbf{U}^{\text{net}}$—and hence in $\mathbf{k} \cdot \mathbf{U}^{\text{net}}$—are large, and above-diagonal terms may be large, but below-diagonal terms are mainly small. The solution to the matrix eigenvalue relation (4.3) for some reasonably large number of modes is thus

$$\begin{vmatrix} ku_{11}^{\text{net}} + lv_{11}^{\text{net}} - \omega & ku_{12}^{\text{net}} + lv_{12}^{\text{net}} & ku_{13}^{\text{net}} + lv_{13}^{\text{net}} & \cdots \\ \boldsymbol{ku_{21}^{\text{net}} + lv_{21}^{\text{net}}} & ku_{22}^{\text{net}} + lv_{22}^{\text{net}} - \omega & ku_{23}^{\text{net}} + lv_{23}^{\text{net}} & \cdots \\ \boldsymbol{ku_{31}^{\text{net}} + lv_{31}^{\text{net}}} & \boldsymbol{ku_{32}^{\text{net}} + lv_{32}^{\text{net}}} & ku_{33}^{\text{net}} + lv_{33}^{\text{net}} - \omega & \cdots \\ \bullet\bullet\bullet & \bullet\bullet\bullet & \bullet\bullet\bullet & \cdots \end{vmatrix} = 0, \tag{5.1}$$

where boldface indicates values which tend to be small. This determinant can be expanded approximately down the first column as

$$(ku_{11}^{\text{net}} + lv_{11}^{\text{net}} - \omega)$$

$$\times \begin{vmatrix} ku_{22}^{\text{net}} + lv_{22}^{\text{net}} - \omega & ku_{23}^{\text{net}} + lv_{23}^{\text{net}} & \cdots \\ \boldsymbol{ku_{32}^{\text{net}} + lv_{32}^{\text{net}}} & ku_{33}^{\text{net}} + lv_{33}^{\text{net}} - \omega & \cdots \\ \bullet\bullet\bullet & \bullet\bullet\bullet & \cdots \end{vmatrix}$$

$$+ \text{ (small)} = 0. \tag{5.2}$$

Continuing to expand, we have approximately

$$(ku_{11}^{\text{net}} + lv_{11}^{\text{net}} - \omega)(ku_{22}^{\text{net}} + lv_{22}^{\text{net}} - \omega)$$

$$\times (ku_{33}^{\text{net}} + lv_{33}^{\text{net}} - \omega) \cdots = 0. \tag{5.3}$$

This gives approximate solutions

$$\omega_i = ku_{ii}^{\text{net}} + lv_{ii}^{\text{net}}, \qquad i = 1, 2, 3, \cdots, \tag{5.4}$$

which is simple horizontal advection by the *i*th diagonal entry of $\mathbf{U}^{\text{net}}$. (This would not be expected to hold for all shear modes, since complex solutions would enter at some stage.)

a. The first shear mode

Equation (5.4) predicts that the gravest shear mode is approximately advected by $(u_{11}^{\text{net}}, v_{11}^{\text{net}})$. This is consistent with the findings of KCdS, [their Eq. (22)], who showed to a good approximation (for east–west propagation, but now seen to be more general) that there was an effective Doppler shift from the resting planetary wave speed by an amount proportional to the second background mode, precisely equivalent to $\mathbf{u}_{11}^{\text{eff}}$.

The eigenvector corresponding to this can be evaluated approximately. Since modal loadings fall off rapidly after mode 2, we here consider only the first two modes. We use the shorthand

$$r_{ij} = ku_{ij}^{\text{net}} + lv_{ij}^{\text{net}}$$

and note that r_{21} can be considered small compared with the other entries by the previous arguments. The solution of $|\mathbf{r} - \omega\mathbf{I}| = 0$ is then approximately (5.4), but showing the small corrections:

$$\omega_1 = r_{11} + \frac{r_{12}r_{21}}{r_{11} - r_{22}}; \qquad \omega_2 = r_{22} - \frac{r_{12}r_{21}}{r_{11} - r_{22}}.$$

The first eigenvector is then to the same degree of approximation,

$$\mathbf{M}_1 = \begin{pmatrix} 1 \\ \left[\dfrac{r_{21}}{r_{11} - r_{22}} \right] \end{pmatrix},$$

and this, given the smallness of r_{21}, is approximately the vector $(1, 0)^{\text{T}}$ and so is dominated by the first internal mode. The entries in Tables 1 and 2 for east–west orientation give some detail. The continuously stratified first shear mode velocity in Table 1 is -3.7 cm s^{-1}, which is also the eigenvalue from the 4×4 matrix shown; that for the 2×2 submatrix is -3.3 cm s^{-1}, which is very similar. However, at this location the u_{11}^{net} is less accurate, at -5.2 cm s^{-1}. (As we shall see, this degree of inaccuracy is not typical.) The first two modal loadings bear out the predicted domination by the first internal mode, again confirming the approxi-

mate approach of KCdS. The Table 2 results have similar properties, even to the overestimation of the shear mode speed by u_{11}^{net}. This mode is similar to the graver mode in two-layer models such as Liu's (1999b) "non-Doppler-shift" mode, and has connections with characteristic velocities in such models (e.g., Luyten and Stommel 1986).

These examples notwithstanding, on a global basis $\mathbf{u}_{11}^{net}$ is a superior estimate of the phase speed than the matrix eigenvalue solution, for reasons that are unclear. Figure 7 shows scatterplots comparing (between latitudes 5°–50° from the equator) the exact continuous first shear mode phase speed with either (Fig. 7a) the 2 × 2 eigenvalue phase speed (that using more modes is indistinguishable) or (Fig. 7b) u_{11}^{net}. Because the high speeds near the equator are included, all correlations are near unity. Figure 7c, restricted to the latitude range ±(30°–50°), shows that at low speeds the correlations are less good: between continuous and shear mode the correlation is 0.5, and between continuous and u_{11}^{net} it is 0.9. It should be noted that u_{11} tends to overestimate the phase speed. (KCdS show a similar result for model data.)

Figure 8 shows a similar diagram for propagation at 45° in the eastward direction. The phase velocities are now much reduced (as they are in classical planetary wave theory, so that the phases move east–west at similar speeds to before). The scatter at low speeds shown by the shear mode phase speed remains, with many eastward speeds not found in the full calculation. The superiority of $u_{11}^{net} \cos 45 + v_{11}^{net} \sin 45$ to the shear mode is similar to the case of eastward orientation, with a similar degree of overestimation.

b. The second shear mode

From (5.4), an approximation for the second shear mode phase speed would be $(u_{22}^{net}, v_{22}^{net})$, again acting as an effective Doppler shift to the undisturbed phase speed. The eigenvector would then take the approximate form

$$\mathbf{M}_2 = \begin{pmatrix} -r_{12} \\ r_{11} - r_{22} \end{pmatrix}$$

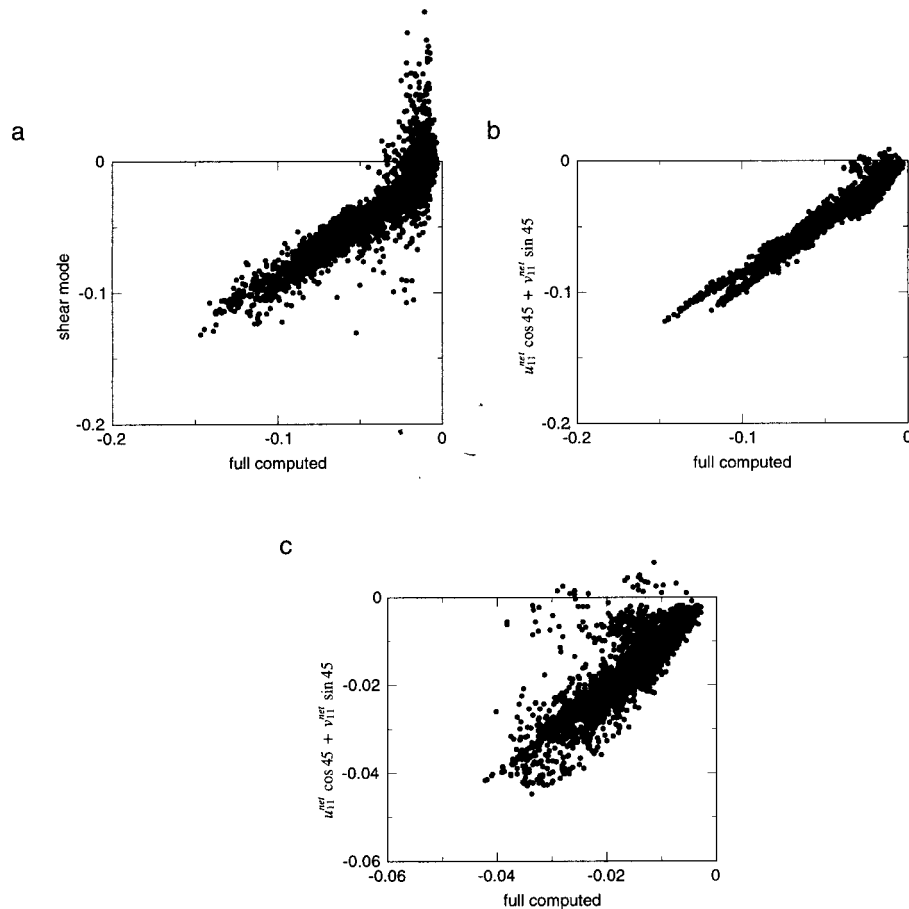so that loadings from modes 1 and 2 are of similar sizes.

FIG. 8. As Fig. 7, but for phase speeds of the first shear mode oriented at 45° to the eastward direction, and using the identification of $u_{11}^{net} \cos 45 + v_{11}^{net} \sin 45$ with shear mode 1.
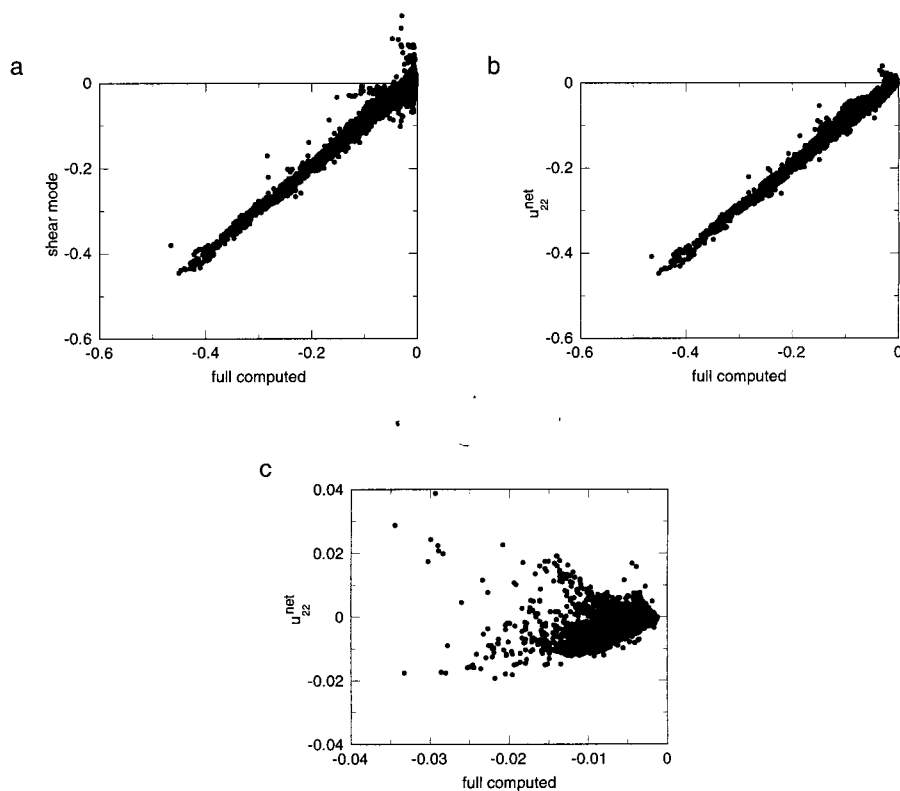
FIG. 9. As Fig. 7, but for the second shear mode, and using the identification of u_{22}^{net} with shear mode 2.

This is confirmed by the entries in Tables 1 and 2: the second shear mode does *not* take the form of the second internal mode, but possesses large loadings from the first internal mode. The mode 1 loading means that it may be possible for shear-mode-1-like behavior to "leak" faster disturbances westward from a slightly unbalanced second shear mode.

Tables 1 and 2 also indicate that the above approximation is less good than for the first shear mode (indeed, no continuously stratified shear mode was located at the location in Table 2). Figure 9 confirms this. There is somewhat more scatter between the 2×2 eigenvalue and the full continuous problem than for the first shear mode. Over the full latitude range, u_{22}^{net} serves as a valid estimate. At low speeds (Fig. 9c) over substantial areas of the ocean a positive velocity is predicted, although most points are well correlated with the continuous solution. This mode identifies roughly with Liu's (1999a,b) "advective," or "A," mode.

6. Propagation behavior

The previous section has shown that the diagonal pseudovelocities give reasonable, but not perfect, estimates of the propagation speed for a shear mode, for wavenumbers oriented in any direction save near 90° (a rare occurrence in practice). Unfortunately, since most propagation seems to occur at latitudes above 30°, it is the slower speeds estimated by the pseudovelocities that contain more scatter, since the domination of the planetary wave speed over the background velocities becomes less at higher latitudes. Appendix B gives arguments for why the shear mode propagation should resemble that of a single velocity vector, based on the smallness of v compared with u in most of the ocean.

Figures 10 and 11 show the speeds and directions of propagation of the first two modes for the North Atlantic and North Pacific, respectively. The fields have been interpolated using the methods and software of Fieguth et al. (1998). Because the amplitude of the velocities differs strongly between high and low latitudes, it is hard to show detail at higher latitude. To aid this, velocities are not shown south of 15°N, and the square root of the velocity (correctly oriented, taking cosine shrinkage of longitude and diagram aspect ratio into account) is shown rather than the velocity itself.

Except at latitudes above 50° (where wave propagation would have to involve periods above a year in classical planetary wave theory), the first shear mode is predicted to travel almost entirely westward. This is due to a combination of reasons, as we have seen. First, the propagation of the gravest shear mode is dominated by westward planetary wave dynamics. Second, the east–west background velocity is larger than the north–
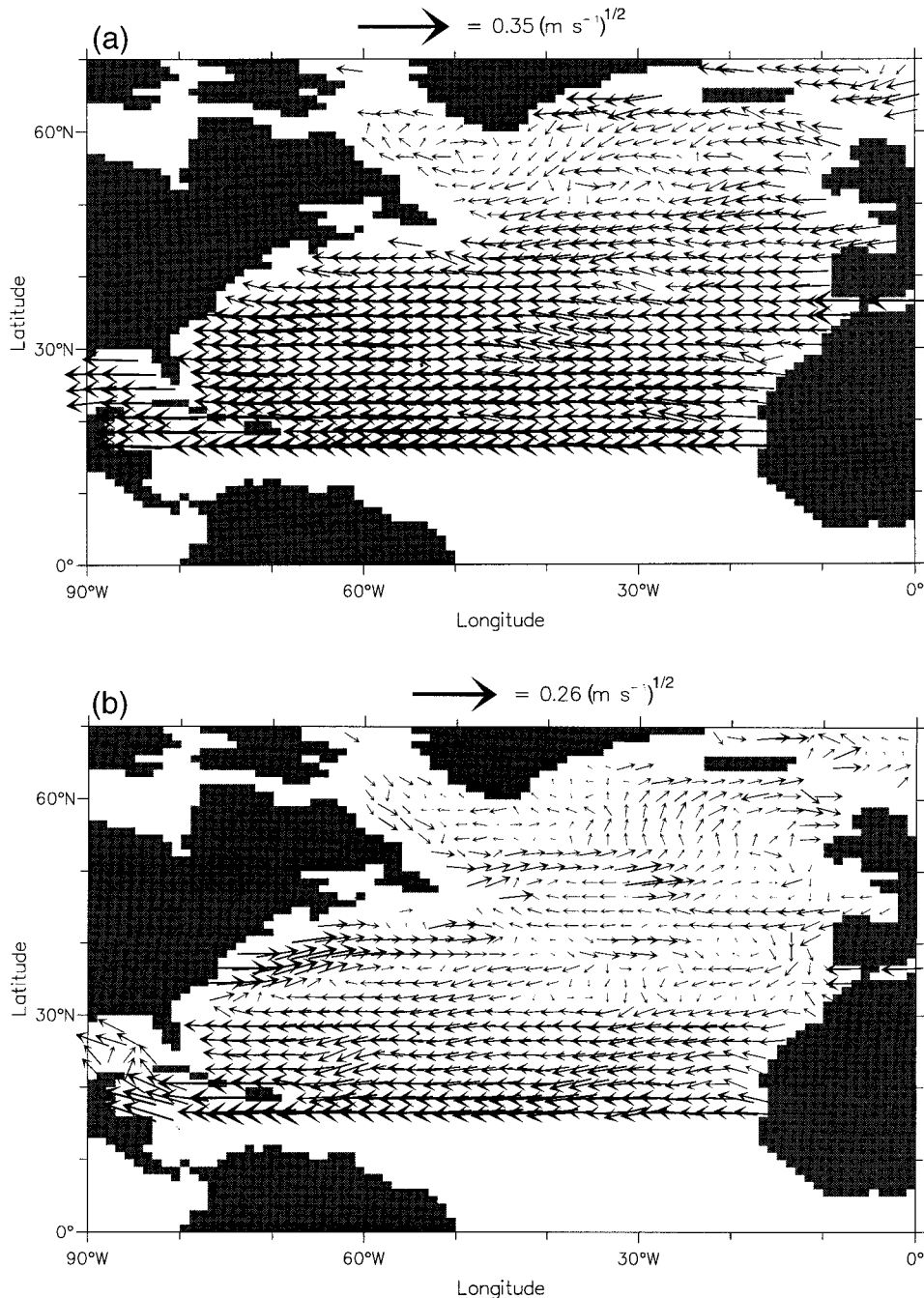
FIG. 10. Diagonal entries of the pseudovelocity matrix (i.e., estimates of the speed and direction in which free shear modes propagate) for the North Atlantic. The pseudovelocities have been interpolated using the Fieguth et al. (1998) software. Although computed on a 1° grid, velocities are only shown every 2° for clarity. Velocities are not shown south of 15°N, and the square root of the velocity amplitude is shown, so that detail is visible in the small velocities at higher latitudes. The orientation takes the diagram aspect ratio and longitudinal convergence at higher latitudes into account. (a) The first shear mode pseudovelocities ($\mathbf{u}_{11}^{\mathrm{net}}$). (b) The second shear mode pseudovelocities ($\mathbf{u}_{22}^{\mathrm{net}}$).

south, so that u_{11}^{eff} itself dominates over v_{11}^{eff}. This combination yields the westward behavior evident in Figs. 10a and 11a.

In contrast, a much more gyrelike circulation is ev-

ident for the second shear mode (Figs. 10b and 11b). Up to about 30°N the pseudovelocity is again westward (since the second undisturbed planetary wave is still dominating at these latitudes). North of this, however,
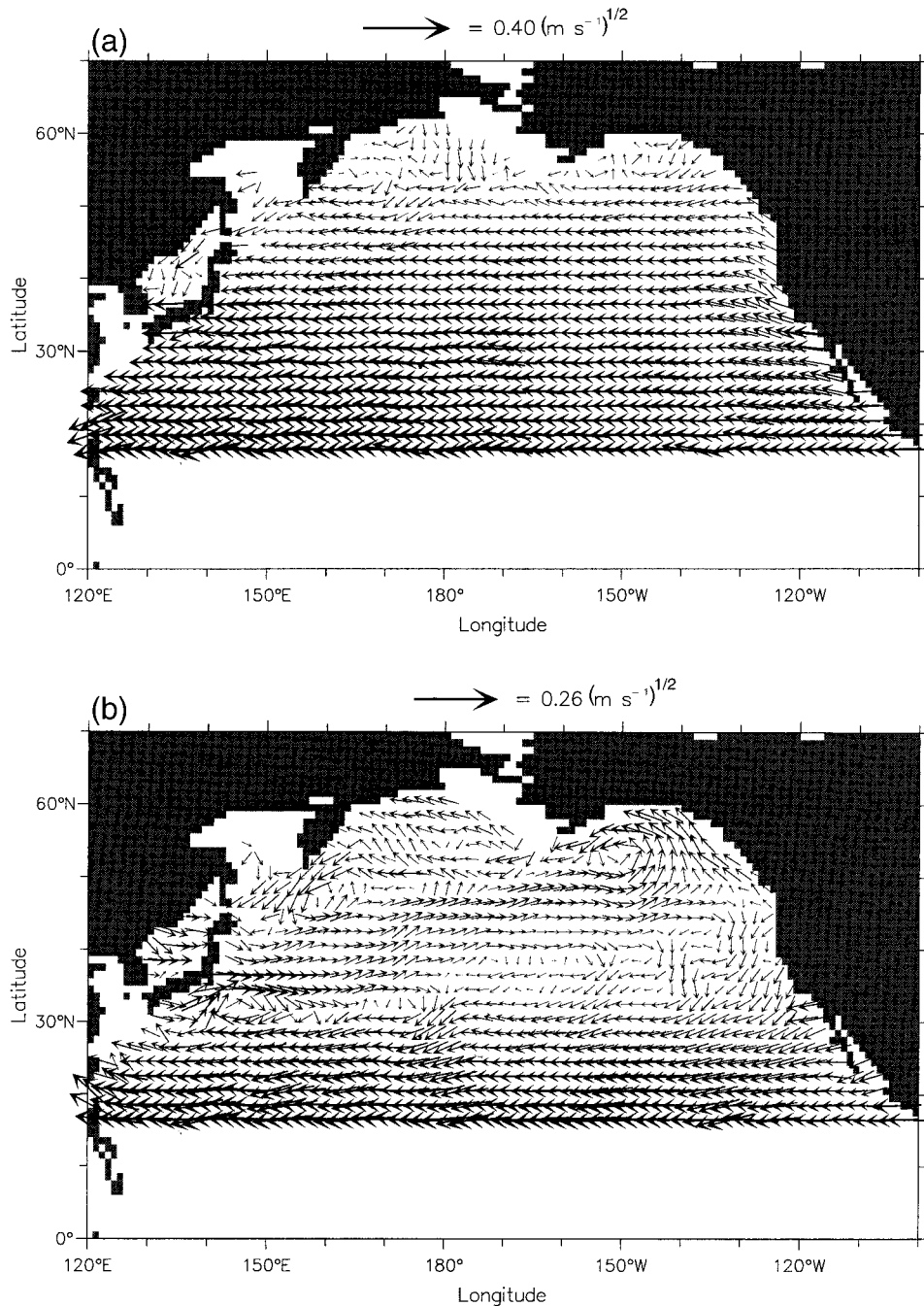
FIG. 11. As Fig. 10, but for the North Pacific.

the planetary wave is slower, and the pseudovelocity is dominated by the contribution from the first background mode. Although, as before, the v contribution is small compared with u, the width of the ocean basins means that north–south propagation can occur. Both diagrams show evidence for increased western boundary flow despite the smoothing used in the temperature and salinity data. The predicted flow is similar to, but not the same as, that predicted by the mean anomaly of geopotential thickness between 500 and 1000 m in these oceans and latitude bands (Levitus 1982). There is little direct correlation between the baroclinic velocities computed at these depths and $\mathbf{u}_{22}^{net}$ if a wide latitude is included in the computation. Figures 10b and 11b indicate why: the beta term produces a dominant westward trend at lower latitudes. If attention is restricted to latitudes above 30°,

the correlations improve markedly, being around 0.7 for depths between 500 and about 1200 m. However, amplitudes are poorly estimated: $\mathbf{u}_{22}^{\text{net}}$ is approximately $0.3\mathbf{u}(500 \text{ m})$ and $0.5\mathbf{u}(1000 \text{ m})$, but nowhere down the water column is there a depth at which the way the second shear mode propagates is well estimated by the local velocity. Nonetheless, the directionality is well represented. For example, the southward part of the splitting of the eastward motion at 48°N, 150°W in the Pacific is reminiscent of Liu and Zhang's (1999) Fig. 3 (mislabeled in the paper), although their southward motion appears to be caused by the effective barotropic flow averaged in the thermocline (the model being 2.5-layer quasigeostrophic).

7. Discussion

This paper has argued the need for a fairly accurate descriptor of how small-amplitude, large-scale oceanic internal disturbances propagate. Subject only to the flat-bottom boundary condition, the linear internal modes form a useful basis for expanding the oceanic shear modes of propagation. Remarkably, the shear modal structure is largely independent of orientation of the flow. The resulting advective velocities, termed pseudovelocities here, are a mixture of (i) background flow decomposed onto normal modes and (ii) planetary wave propagation speeds. The diagonal entries of the matrix of pseudovelocities prove to be good, though not perfect, descriptors of the speed and direction of propagation of the shear modes. They appear to be more accurate estimators than the eigenvalue solutions of the pseudovelocity matrix itself, for reasons that are unclear. (Arguments are given that suggest why the propagation of shear modes should resemble that by a single velocity vector.)

The first diagonal entry contains no contribution from the first baroclinic mode of the background flow; the second diagonal entry is mainly from the first background mode. The first shear mode is dominated by westward propagation, and is the mode first discussed by KCdS, showing the midlatitude speed-up over the undisturbed linear first mode planetary wave. The pseudovelocity for the second shear mode, in contrast, while still dominated by westward propagation at lower latitudes, shows a gyrelike structure at latitudes above 30° where the (second) planetary wave speed is smaller and can be countered by the first baroclinic mode background flow contribution. This resembles in *shape* and *direction* the geostrophic baroclinic flow between about 500- and 1000-m depth. The speed, however, is half or less of the geostrophic flow at those depths, showing that features may well be able to propagate some distance around a subtropical or subpolar gyre, but they will not in general propagate at the speed of the circulation. This is in agreement with, and extends quantitatively, the simpler 2.5-layer quasigeostrophic model findings of Liu (1999a).

The way in which features propagate around or across basins will be dealt with in a companion paper, which will compare propagation by a diagonal $\mathbf{u}^{\text{net}}$ entry, propagation by a truncated modal set of equations, and propagation interpreted through WKBJ analyses of the full continuous problem.

APPENDIX A

Incorrect Solutions with Zero East–West Wavenumber

A useful example of modal decompositions yielding erroneous behavior is demonstrated by setting the east–west wavenumber k to zero. The continuous problem (4.4) then becomes

$$(lv - \omega)F_{zz} - lv_z F_z = 0,$$

the solution of which is

$$F = A\left\{l\int_{-H}^{z} v \, dz - \omega(z + H)\right\},$$

where A is an arbitrary constant. Requiring the solution to satisfy the upper boundary condition gives, assuming v to be baroclinic, $\omega = 0$. Thus, no horizontal WKBJ solution can achieve a zero east–west wavenumber at finite frequency. As $k \to 0$, solutions disappear or coalesce, although there a solution can be found of similar form to the large u case in KCdS, in which $\omega = lv_{\min} - k^{2/3}\hat{\omega}$, where $v_{\min}$ is the minimum v in the water column, and $\hat{\omega}$ can be computed. The solution resembles mode 1 in the vertical, though possessing a weak internal boundary layer at the minimum of v, of width $k^{1/3}$. Clearly both east–west phase and group velocities become singular in the limit $k \to 0$.

The behavior of solutions to the decomposed problem when k is zero depends on how the truncation is constructed. If the summation into modes uses K internal modes, and the solution is cast onto $L = K$ modes, then there is always a zero eigenvalue present when k vanishes (or, equivalently, when β vanishes). In the case of a single mode ($K = 1$) this is trivial: then $\mathbf{u}_{11} = \gamma_{111}\mathbf{u}_1 = 0$. For larger K, this is because the matrix A is of the form $A_{kj} = \sum_{i=1}^{K} lv_i\gamma_{kij}$, and immediately we have $\sum_j v_j A_{kj} = \sum_{i=1}^{K}\sum_{j=1}^{K} lv_i v_j\gamma_{kij} = 0$. Thus, the columns of A are linearly dependent and so a zero eigenvalue exists as in the continuous problem. However, there will still be $K - 1$ other eigenvalues that are nonzero and have no physical counterpart in the continuous solution.

If the number of modes predicted (L) is less than K, then there is no zero eigenvalue at all when $k = 0$. The trivial case of this is when $K = 2$ (retaining modes 1 and 2) and $L = 1$ (considering the shear mode 1 only). Since the background mode 2 has strongest effect on

the shear mode 1 (cf. KCdS), using this combination would be a plausible option. However, in such a case $\mathbf{u}_{11}^{\text{eff}} \neq 0$ since it now contains the second mode. Thus, with such a truncation there would apparently be propagation when $k = 0$, with a nonzero frequency, which is incorrect. Thus, in general, not all the decomposed modal solutions—and therefore the finite layer truncations—generate reliable solutions.

APPENDIX B

Why the Shear Modes Resemble the Response to a Single Velocity Field

This appendix suggests (i) why the eigenvector tends to be insensitive to orientation of the wavevector, and (ii) why the eigenvalue responds to change in orientation approximately as if the advection were by a single velocity field. The argument uses the fact that the background north–south velocity is mainly small compared with the east–west velocity. (Clearly v is not small compared with u everywhere, by the statistics given earlier, so this argument can only be suggestive in the few regions where $v \gg u$.) The argument is given for (i) the matrix of pseudovectors and (ii) for the continuous problem.

a. Pseudovectors

If the pseudovectors were all oriented precisely east–west, (4.2) would have the solution $(\omega_0, \mathbf{M}_0)$:

$$\omega \equiv \omega_0 = c|\mathbf{k}| \cos\phi \qquad \mathbf{U} \cdot \mathbf{M}_0 = c\mathbf{M}_0. \quad \text{(B.1)}$$

Here we write $\mathbf{k} = |\mathbf{k}|(\cos\phi, \sin\phi)$, and the $\cos\phi$ factor on ω_0 merely reflects the vector component on the east axis. We now permit a small north–south set of velocities $\sigma\mathbf{V}$, where σ is simply a reminder of small quantities, and write

$$\omega = c|\mathbf{k}| \cos\phi + \sigma c\psi \sin\phi$$
$$\mathbf{M} = \mathbf{M}_0 + \varepsilon\mathbf{M}_1. \quad \text{(B.2)}$$

(The form for the change in ω could be left completely open; this representation is for clarity.) Substitution into (4.2) and removal of the leading terms gives

$$\sin\phi\mathbf{V} \cdot \mathbf{M}_0 + \cos\phi\mathbf{U} \cdot \mathbf{M}_1$$
$$= c\psi \sin\phi\mathbf{M}_0 + c \cos\phi\mathbf{M}_1. \quad \text{(B.3)}$$

(This will not be the leading solution if ϕ is near 90°, consistent with the numerical solutions.) We introduce the left eigenvector of $\mathbf{U}$ that corresponds to $\mathbf{M}_0$, here termed $\mathbf{N}$, which satisfies

$$\mathbf{N}^{\text{T}} \cdot \mathbf{U} = c\mathbf{N}^{\text{T}}, \quad \text{(B.4)}$$

where superscript T denotes the transpose. Left-multiplying (B.3) by $\mathbf{N}^{\text{T}}$ means that the second terms on each side of (B.3) cancel, leaving

$$\mathbf{N}^{\text{T}} \cdot \mathbf{V} \cdot \mathbf{M}_0 = c\psi\mathbf{N}^{\text{T}} \cdot \mathbf{M}_0 \quad \text{or}$$

$$\psi = \frac{\mathbf{N}^{\text{T}} \cdot \mathbf{V} \cdot \mathbf{M}_0}{c\mathbf{N}^{\text{T}} \cdot \mathbf{M}_0}. \quad \text{(B.5)}$$

This shows that the approximation used is consistent.

Now suppose that the solution of the problem related to a single velocity c at an angle ψ to the east axis, where ψ is small because of the dominance of u over v. In such a case, we would have

$$\frac{\omega}{|\mathbf{k}|} = c \cos(\phi - \psi)$$

for an orientation ϕ. If ψ is indeed small, this reduces to $\omega/|\mathbf{k}| = c \cos\phi + c\psi \sin\phi$, which is of the form (B.2). Thus, the eigenvalue responds as it would to a single velocity field.

The eigenvector, meanwhile, remains $\mathbf{M}_0$ to leading order, and thus is approximately independent of orientation.

b. Continuous problem

The arguments are similar in this case. We pose

$$\omega = c|\mathbf{k}| \cos\phi + \sigma c\psi \sin\phi$$

$$(u - c)F_{0zz} - u_z F_{0z} + \frac{\beta N^2}{f^2}F_0 = 0. \quad \text{(B.6)}$$

We again expand using a small v field, so that $F = F_0 + \sigma F_1$, etc. The first-order correction to (4.4) becomes

$$\cos\phi\left\{(u - c)F_{1zz} - u_z F_{1z} + \frac{\beta N^2}{f^2}F_1\right\}$$
$$+ \sin\phi\{(v - c\psi)F_{0zz} - v_z F_{0z}\} = 0. \quad \text{(B.7)}$$

Note that (B.6) can be turned into a self-adjoint system by dividing by $(u - c)^2$. Thus, we may eliminate F_1 by taking $(1/(u - c)^2 \cos\phi)$ times (B.7) $- 1/(u - c)^2$ times (B.6) and integrating top to bottom. This leaves, after cancellation,

$$c\psi \int_{-H}^{0} \frac{F_{0zz}}{(u - c)^2} dz = \int_{-H}^{0} \frac{dz}{(u - c)^2}(vF_{0zz} - v_z F_{0z}),$$

so that again there is a solution for ψ and, hence, the assumed form for ω again holds. Similarly, the F_0 eigenvector still dominates the eigensolution.

REFERENCES

Chelton, D. B., and M. G. Schlax, 1996: Global observations of oceanic Rossby waves. *Science,* **272,** 234–238.

Cipollini, P., D. Cromwell, M. S. Jones, G. D. Quartly, and P. G. Challenor, 1997: Concurrent altimeter and infrared observations of Rossby wave propagation near 34 degrees N in the Northeast Atlantic. *Geophys. Res. Lett.,* **24,** 889–892.

——, ——, P. G. Challenor, and S. Raffaglio, 2001: Rossby waves detected in global ocean colour data. *Geophys. Res. Lett.,* **28,** 323–326.

Deser, C. M., A. Alexander, and M. S. Timlin, 1996: Upper ocean

thermal variations in the North Pacific during 1970–1991. *J. Climate,* **9,** 1840–1855.

de Szoeke, R. A., and D. B. Chelton, 1999: The modification of long planetary waves by homogeneous potential vorticity layers. *J. Phys. Oceanogr.,* **29,** 500–511.

——, and M. Y. Morris, 2000: On the propagation of baroclinic waves in the general circulation. *J. Phys. Oceanogr.,* **30,** 2637–2649.

Dewar, W. K., 1998: On ''too fast'' baroclinic planetary waves in the general circulation. *J. Phys. Oceanogr.,* **28,** 1739–1758.

Dickson, R. R., J. Meincke, S. A. Malmberg, and A. J. Lee, 1988: The ''Great Salinity Anomaly'' in the northern North Atlantic, 1968–1982. *Progress in Oceanography,* Vol. 20, Pergamon, 103–151.

Ezer, T., 1999: Decadal variability of the upper layers of the subtropic North Atlantic: An ocean model study. *J. Phys. Oceanogr.,* **29,** 3111–3124.

Fieguth, P. W., D. Menemenlis, T. Ho, A. Willsky, and C. Wunsch, 1998: Mapping Mediterranean altimeter data with a multiresolution optimal interpolation algorithm. *J. Atmos. Oceanic Technol.,* **15,** 535–546.

Flierl, G. R., 1978: Models of vertical structure and the calibration of two-layer models. *Dyn. Atmos. Oceans,* **2,** 341–382.

Fu, L.-L., and D. B. Chelton, 2001: Large-scale ocean circulation. *Satellite Altimetry and Earth Sciences,* L.-L. Fu and A. Cazenave, Eds., Academic Press, 133–169.

Hill, K., I. R. Robinson, and P. Cipollini, 2000: Propagation characteristics of extratropical planetary waves observed in the ATSR global sea surface temperature record. *J. Geophys. Res.,* **105,** 21 927–21 945.

Jacobson, A. R., and J. L. Spiesberger, 1998: Observations of El Niño–Southern Oscillation induced Rossby waves in the northeast Pacific using in situ data. *J. Geophys. Res.,* **103,** 24 585–24 596.

Kessler, W., 1990: Observations of long Rossby waves in the northern tropical Pacific. *J. Geophys. Res.,* **95,** 5183–5218.

Killworth, P. D., and D. L. T. Anderson, 1977: Meaningless modes? Mode Hot-Line News, No. 72, 8 pp.

——, and J. R. Blundell, 1999: The effect of bottom topography on the speed of long extratropical planetary waves. *J. Phys. Oceanogr.,* **29,** 2689–2710.

——, D. B. Chelton, and R. A. de Szoeke, 1997: The speed of observed and theoretical long extratropical planetary waves. *J. Phys. Oceanogr.,* **27,** 1946–1966.

Latif, M., and T. P. Barnett, 1994: Causes of decadal climate variability over the North Pacific and North America. *Science,* **266,** 634–637.

Levitus, S., 1982: *Climatological Atlas of the World Ocean.* NOAA Professional Paper 13, 173 pp.

——, 1989: Interpentadal variability of temperature and salinity in the deep North Atlantic, 1970–1974 versus 1955–1959. *J. Geophys. Res.,* **94,** 16 125–16 131.

——, and T. Boyer, 1994: *World Ocean Atlas 1994.* Vol. 4: *Temperature,* NOAA Atlas NESDIS 4, 117 pp.

——, R. Burgett, and T. Boyer, 1994: *Salinity, World Ocean Atlas 1994.* Vol. 3: NOAA Atlas NESDIS 3, 99 pp.

Liu, Z., 1999a: Forced planetary wave response in a thermocline gyre. *J. Phys. Oceanogr.,* **29,** 1036–1055.

——, 1999b: Planetary wave modes in the thermocline: Non-Doppler-shift mode, advective mode and Green mode. *Quart. J. Roy. Meteor. Soc.,* **125,** 1315–1339.

——, and R.-H. Zhang, 1999: Propagation and mechanism of decadal upper-ocean variability in the North Pacific. *Geophys. Res. Lett.,* **26,** 739–742.

Luyten, J., and H. Stommel, 1986: Gyres driven by combined wind and buoyancy flux. *J. Phys. Oceanogr.,* **16,** 1551–1560.

National Geophysical Data Center, 1988: Digital relief of the Surface of the Earth. NOAA, National Geophysical Data Center, Data Announcement 88-MGG-02.

Pedlosky, J., 1964: An initial value problem in the theory of baroclinic instability. *Tellus,* **16,** 12–17.

——, 1987: *Geophysical Fluid Dynamics.* 2d ed. Springer Verlag, 710 pp.

Sutton, R. T., and M. R. Allen, 1997: Decadal predictability of North Atlantic sea surface temperature and climate. *Nature,* **388,** 563–567.

Tokmakian, R. T., and P. G. Challenor, 1993: Observations in the Canary Basin and the Azores Frontal using Geosat data. *J. Geophys. Res.,* **98,** 4761–4773.

Welander, P., 1959: An advective model of the ocean thermocline. *Tellus,* **11,** 309–318.

——, and C.-T. Liu, 1976: On a mid-ocean thermocline regime. *J. Phys. Oceanogr.,* **6,** 592–595.

White, W. B., 1977: Annual forcing of baroclinic long waves in the tropical North Pacific. *J. Phys. Oceanogr.,* **7,** 50–61.

——, and R. G. Peterson, 1996: An Antarctic circumpolar wave in surface pressure, wind, temperature and sea–ice extent. *Nature,* **380,** 699–702.

Yang, H., 2000: Evolution of long planetary wave packets in a continuously stratified ocean. *J. Phys. Oceanogr.,* **30,** 2111–2123.

Zang, X., and C. Wunsch, 1999: The observed dispersion relationship for North Pacific Rossby wave motions. *J. Phys. Oceanogr.,* **29,** 2183–2190.

Zhang, R.-H., and S. Levitus, 1997: Structure and cycle of decadal variability of upper ocean temperature in the North Pacific. *J. Climate,* **10,** 710–727.

——, and Z. Liu, 1999: Decadal thermocline variability in the North Pacific Ocean: Two pathways around the subtropical gyre. *J. Climate,* **12,** 3273–3296.