Long extra-tropical planetary wave propagation in the presence of slowly varying mean flow and bottom topography. I: the local problem

Peter D. Killworth and Jeffrey R. Blundell

Southampton Oceanography Centre, Empress Dock, Southampton SO14 3ZH, England

ABSTRACT

One of the most successful theories to date to explain why observed planetary waves propagate westwards faster than linear flat-bottom theory predicts has been to include the effect of background baroclinic mean flow, which modifies the potential vorticity waveguide in which the waves propagate. (Barotropic flows are almost everywhere too small to explain the observed differences.) That theory accounted for most, but not all, of the observed wave speeds. A later attempt to examine the effect of the sloping bottom on these waves (without the mean flow effect) did not find any overall speed-up. This paper combines these two effects, assuming long (geostrophic) waves and slowly varying mean flow and topography, and computes group velocities at each point in the global ocean. These velocities turn out to be largely independent of the orientation of the wavevector. A second speed-up of the waves is found (over that for mean flow only). Almost no eastward-oriented group velocities are found, so that features which appear to propagate in the same sense as a subtropical gyre would have to be coupled with the atmosphere or be density-compensated in some manner.

1. Introduction

Planetary waves are the main mechanism whereby information about one part of the ocean is transferred to another part. Interest in their properties was rekindled when the availability of satellite altimetry permitted observation of planetary waves (Tokmakian and Challenor, 1993), and the discovery by Chelton and Schlax (1996) that these waves appeared to be propagating at speeds up to twice those predicted by the linear vertical normal mode theory. Although there has been some debate as to the extent of this discrepancy (e.g., Zang and Wunsch, 1999), it is now generally accepted that waves observed in altimetry *do* propagate significantly faster than linear theory would predict (and are also visible in other remotely sensed fields).

Theories have been put forward using a variety of mechanisms to explain the speed-up. One of the most successful to date (though not completely accounting for observations) is due to Killworth et al. (1997), who proposed that the background baroclinic east-west shear flow would alter the existing potential vorticity gradient sufficiently to modify the phase speed of small disturbances. (This had been earlier suggested by Kang and Magaard, 1979.) The Doppler shift produced by the barotropic flow was found to be almost everywhere far too small to account for the speed changes, as well as not necessarily being in the right direction. Killworth et al.'s baroclinic computations confirmed the general speed-up, save within about 10° of the equator. However, the speed-up predicted still remained slightly smaller than observed in higher northern latitudes, and noticeably smaller in higher southern latitudes. Dewar (1998), de Szoeke and Chelton (1999) and Liu (1999a, 1999b) have given simplified explanations why this should be the case, and Dewar and Morris (2000) have demonstrated the effect in a quasi-geostrophic model. Fu and Chelton (2001) extended the results to non-long waves, showing excellent agreement between observed and computed dispersion relations. Killworth and Blundell (1999) examined the role of topographic gradients in altering the phase and group velocity of long planetary waves in a continuously stratified model. While topographic effects were effective locally, Killworth and Blundell found that the effects tended to cancel out across an ocean basin, leaving the net effect of topography to be small. [This was not found by Tyler and Käse's (2001) in two-layer formulation of the same problem. There is a general, but unproven, belief that two-layer models appear to overestimate effects of topography; certainly in this case their results, showing significant net speed-ups, are not in accord with the equivalent continuously stratified calculation.]

Killworth and Blundell (2001), assuming that the cancellation of topographic effects across a basin might continue if baroclinic shear was included again, used a normal mode decomposition based on a locally flat bottom to produce approximate first and second mode phase and group velocities. [Liu (1999a, b) had been the first to draw attention to the second mode using a 2.5 layer model, both from the perspective of the surface temperature signal and from its direction of propagation.] They found similar results to those of Liu, including an apparent gyre-like structure in the second-mode group velocities. These results will be addressed later in this paper. Tailleux and McWilliams (2001) present an argument based on a different bottom boundary condition (one with zero pressure perturbation) and no mean flow to obtain a similar set of phase velocities to those found by Killworth et al. (1997) with a flat bottom and zonal mean flow. The waves are speeded up by a factor which is the same as that found by Samelson (1992) in a two-layer calculation, who found an element of surface trapping in waves over bumpy topography. While our work here addresses the traditional boundary condition of no normal flow, we shall also discuss this alternative condition later in this paper. Finally, Yang (2000) examined the propagation of a wave *packet*, rather than a coherent wave with a vertical structure as considered by the other cited papers. In the presence of a mean flow, the horizontal group velocity of the packet could differ in direction with depth (the effect of horizontal boundaries was neglected). Yang concluded that wave packets could account for most of the observed features in the satellite data, although his conclusion held for waves with a large vertical scale, formally invalid under the ray theory used.

Most papers cited used a WKBJ formulation (as we shall here), so that the background through which long waves propagate is considered slowly varying compared with the length scale of the perturbation (itself long compared with a deformation radius so that the long-wave approximation holds). However, the more general and realistic problem – how does a long planetary wave propagate through the background mean flow present in the ocean, across a slowly-varying version of the ocean topography? – has not been tackled, largely due to the computational and analytical difficulties (the problem depends on the orientation of the wavevector, and hitherto formulae for group velocities in such circumstances have been lacking). This paper attempts to remedy this omission, by producing local solutions for the first two vertical normal modes in these more general circumstances. Its companion paper uses ray theory and an eastern boundary wave initialisation to confirm that the solutions in this paper are legitimate.

Section 2 sets up the problem, and the algebra for its solution is given in Section 3. Section 4 discusses group velocities and a clustering algorithm to describe them. Section 5 examines some typical solutions, and Section 6 gives global results for group velocities.

2. Formulation

Following our earlier work, the planetary geostrophic approximation is employed, since this permits variation in stratification, unlike the quasi-geostrophic approximation. The equations of motion are cast onto Welander’s (1959) variable M. We write

$$p/\rho_0 = M_z \tag{2.1}$$

where p represents pressure, ρ the density with ρ_0 a reference value. The geostrophic velocity field (u, v, w) relative to axes (λ longitude, ϕ latitude, z upwards) is given by

$$u = -\frac{M_{z\phi}}{af} \tag{2.2}$$

$$v = \frac{M_{z\lambda}}{af\cos\phi} \tag{2.3}$$

$$w = \frac{M_{\lambda}}{a^2 f\sin\phi} \tag{2.4}$$

where a is the radius of the Earth, $f = 2\Omega\sin\phi$ is the Coriolis parameter, and Ω is the Earth’s rotation rate. The density is given by

$$\rho = -\frac{M_{zz}}{g}\rho_0 \tag{2.5}$$

where g is the gravitational acceleration. The (time-dependent) equation for density conservation is a single equation for M, with t representing time:

$$\rho_t + \frac{u\rho_\lambda}{a\cos\phi} + \frac{v\rho_\phi}{a} + w\rho_z = 0 \text{ or}$$

$$M_{zzt} - \frac{M_{z\phi}M_{zz\lambda}}{fa^2\cos\phi} + \frac{M_{z\lambda}M_{zz\phi}}{fa^2\cos\phi} + \frac{M_\lambda M_{zzz}}{fa^2\sin\phi} = 0. \tag{2.6}$$

We assume that the flow consists of a background mean field, denoted by an overbar, and a small perturbation, denoted by a prime. The background flow, together with the ocean topography, is permitted to vary only on the large (basin) scale L_{basin}, so that WKBJ analysis is permissible. This means that effects of abrupt topography, such as narrow ridges, and narrow boundary currents, must be excluded from analysis. The perturbation flow varies on a scale L_{pert} which is taken to be both short compared with the basin scale, but long enough for the geostrophic balance to hold. As Killworth and Blundell (2001) note, there are regions of the world ocean where the background flow changes sufficiently rapidly to formally invalidate the WKBJ assumption, though such theory continues to provide useable results even in such cases.

The background mean field is henceforth explicitly taken to be baroclinic (i.e., to have zero depth-integrated value), since any barotropic mean component merely induces a Doppler-shift to the solution. The barotropic component is seldom large enough to affect the solutions save for the Antarctic Circumpolar Current and intense western boundary currents; in the latter, a WKBJ formulation would anyway be questionable.

With these assumptions, the perturbation field satisfies

$$M'_{zzt} - \frac{1}{fa^2\cos\phi}(\bar{M}_{z\phi}M'_{zz\lambda} + M'_{z\phi}\bar{M}_{zz\lambda}) + \frac{1}{fa^2\cos\phi}(\bar{M}_{z\lambda}M'_{zz\phi} + M'_{z\lambda}\bar{M}_{zz\phi}) +$$
$$\frac{1}{fa^2\sin\phi}(\bar{M}_\lambda M'_{zzz} + M'_\lambda\bar{M}_{zzz}) = 0. \tag{2.7}$$

Because of the scale separation assumption, the first term in the last bracket ($\bar{w}\rho'_z$) is formally smaller than the second ($w'\bar{\rho}_z$) by a factor L_{pert}/L_{basin}; Killworth and Blundell (2001) note that the ratio is actually smaller still. With this assumption, (2.7) becomes, with a little rearrangement,

$$M'_{zzt} + \frac{\bar{u}}{a\cos\phi}M'_{zz\lambda} + \frac{\bar{v}}{a}M'_{zz\phi} - \frac{\bar{u}_z}{a\cos\phi}M'_{z\lambda} - \frac{\bar{v}_z}{a}M'_{z\phi} + \frac{N^2(x,y,z)}{fa^2\sin\phi}M'_\lambda = 0. \tag{2.8}$$

Equation (2.8) has boundary conditions at surface and floor. For free waves (Liu, 1999b considers a forced problem),

$$w' = 0,\ z = 0 \ \Rightarrow\ M' = 0,\ z = 0. \tag{2.9}$$

The ocean has varying depth $H(\lambda, \phi)$, at which there is no normal velocity;

$$w' + \frac{u' H_\lambda}{a \cos\phi} + \frac{v' H_\phi}{a} = 0,\ z = -H \ \Rightarrow$$

$$M'_\lambda + \tan\phi\,(M'_{z\lambda} H_\phi - M'_{z\phi} H_\lambda) = 0,\ z = -H. \tag{2.10}$$

The mean flow is also assumed to satisfy the boundary condition of no normal flow at the bottom.[1] (The bottom boundary condition is in contention: Tailleux and McWilliams, 2001 have recently suggested an alternative condition of zero bottom pressure perturbation. That used here is the standard, purely kinematic, condition.) Precisely what definition of bottom slope – and, indeed, of bottom depth – is applicable is not clear, and we shall use smoothings of high-resolution data on several different scales in what follows, as well as exploring in Section 5b how close our solutions come to satisfying the Tailleux and McWilliams bottom condition.

We now seek a wavelike solution

$$M' = F(\lambda, \phi, z) \exp i\,(k\lambda + l\phi - \omega t) \tag{2.11}$$

where (k, l) is a wavenumber in the (λ, ϕ) directions and ω is a frequency, assumed positive without loss of generality. The vertical structure F varies slowly laterally (i.e., on the basin scale) while the phase varies on the perturbation scale, as is usual in WKBJ theory. Substitution into (2.8) gives after a little algebra the self-adjoint (but nonlinear in frequency) eigenvalue system

$$\left(\frac{F_z}{R}\right)_z + \frac{S}{R^2} F \equiv L(F) = 0 \tag{2.12}$$

where

$$S(z; \lambda, \phi, k) = \frac{k N^2(\lambda, \phi, z)}{a^2 f \sin\phi} \tag{2.13}$$

$$Q(z; \lambda, \phi, k, l) = \frac{k\bar{u}}{a \cos\phi} + \frac{l\bar{v}}{a} \tag{2.14}$$

$$R(z; \lambda, \phi, k, l) = Q - \omega \tag{2.15}$$

---

[1] This is not strictly consistent with our assumption of purely baroclinic mean flow except when the ocean has a flat bottom. Put another way, it would require a specific mean Ekman pumping. Acquiring a consistent mean flow from observed density data is, of course, a long-standing inversion problem which will not be discussed here.

This has boundary conditions

$$F = 0,\ F_z = 1,\ z = 0 \text{ (assuming a rigid lid)} \tag{2.16}$$

$$F(-H) = -\alpha F_z(-H),\ \text{ where } \alpha = \tan\phi\,(H_\phi - \frac{l}{k}H_\lambda)\,. \tag{2.17}$$

The second surface condition is purely for scaling reasons, providing both the facility to treat (2.12) as an initial-value problem when convenient, and to enable differentiation of the problem to be well-posed.

If the frequency is known, then the phase velocities

$$\begin{gathered}\mathbf{c}^{\mathrm{p}} = (c^{\mathrm{px}}, c^{\mathrm{py}}) = \\ \left(\frac{\omega}{k}a\cos\phi,\ \frac{\omega}{l}a\right) \text{ (definition 1), or} \\ \left(\frac{\omega k}{k^2 + l^2}a\cos\phi,\ \frac{\omega l}{k^2 + l^2}a\right) \text{ (definition 2)}\end{gathered} \tag{2.18}$$

are known. Both definitions are used in the literature. The former is not a velocity, describes phase propagation well, but implies an infinite N-S phase speed for zero l. The latter is a vector, does not describe phase propagation well except in the direction of the wavevector, and implies a zero N-S phase speed for zero l.

We shall frequently ignore the geometrical factors in the following discussion for clarity, but all results will be quoted as actual speeds. If the derivatives ω_k, ω_l can be computed, then the group velocities

$$\mathbf{c}^{\mathrm{g}} = (c^{\mathrm{gx}}, c^{\mathrm{gy}}) = \left(\omega_k a\cos\phi,\ \omega_l a\right) \tag{2.19}$$

are also known. These are rather better behaved than either definition of phase velocity and will be mainly used in what follows.

The system (2.12), (2.16), and (2.17) provide an implicit dispersion relationship connecting ω with (k, l); the remainder of this paper will be concerned with computing it.

Before doing so, however, note that Appendix A demonstrates two properties of the dispersion relationship for a flat bottomed ocean and purely baroclinic mean flow: first, there are no real solutions for which the quantity R vanishes anywhere in the water column (i.e. for which a critical layer exists); and second, there are no solutions for positive k for which R does not vanish. The combination of these two results means that there can be no (real) eastward wave propagation for a flat-bottomed ocean, save for any possible Doppler effects of mean flow. This invalidates the second-mode solutions of Killworth and Blundell (2001), which sometimes had eastward phase velocity components. That paper had unintentionally avoided (apparently) the critical layer difficulties by casting the problem onto a finite number of vertical normal modes. Liu's (1999a, 1999b) papers cast the flow onto 2 or 3 layers (with a net barotropic flow) and, like Killworth and Blundell (2001), found a phase velocity which would have a critical layer somewhere in the water column. Thus none of these papers present solutions which could occur in the continuously stratified regime. (Complex solutions with eastward flow could occur; but these are not located using the present numerical approaches.)

There are a variety of cases which need to be treated, and we consider the matrix of six possibilities shown in Table 1. There are three background mean flow possibilities: no mean flow (N), zonal mean flow only (Z), to permit comparison with Killworth et al. (1997), and general mean flow (G). There are also two topographic possibilities: locally flat topography (F), for which (2.17) reduces to the familiar $F(-H) = 0$, and sloping topography (S). Some of these combinations have been addressed elsewhere: the NF case is the traditional flat-bottom normal mode calculation; the ZF case was examined by Killworth et al. (1997) and Fu and Chelton (2001); and the NS case is dealt with by Killworth and Blundell (1999). Of the remaining combinations, the fully general (GS) case will be the main point of focus, since it should be the case most relevant to the real ocean.

3. Differentiating the dispersion relation

Except in the simplest cases the dispersion relation (2.12) cannot be solved analytically. Because the long wave assumption has been made, the system is homogeneous, in that a solution set $(\omega; k, l)$ implies that $(A\omega; Ak, Al)$ is also a solution for any constant A. The lack of analytical solutions is hardly a problem numerically, save that in order to obtain group velocity we need to be able to compute derivatives of ω w.r.t. k, l. One could contemplate making small changes to k or l, recomputing the solution and finite-differencing to estimate gradients. However, for the ray theory in part II of this work, second derivatives are also required, and this brute force approach would be unsuitable.

We extend the approach of Killworth and Blundell (1999), which assumes that a solution has been found to (2.12), and uses that solution to compute frequency derivatives (in other words, the dispersion relation is not known exactly, but its derivatives can be computed exactly).[2] Here we show how to compute first derivatives of ω w.r.t. k, l from a knowledge of ω itself and its eigenvector F; the approach includes derivatives w.r.t. λ, ϕ as well – these latter are needed in Part II of this paper for ray tracing.

Let X be a component of $\mathbf{X} = \{\lambda, \phi, k, l\}$. Differentiate (2.12) w.r.t. X:

$$L(F_X) + L_X(F) + \omega_X L_\omega(F) = 0. \tag{3.1}$$

At the surface,

$$F_X = 0, z = 0 \tag{3.2}$$

$$F_{Xz} = 0, \; z = 0. \tag{3.3}$$

At the floor,

$$F_X(-H) + \alpha F_{Xz}(-H) - H_X F_z(-H) + \alpha_X F_z(-H) - H_X \alpha F_{zz}(-H) = 0.$$

Thus

$$F_X(-H) + \alpha F_{Xz}(-H) = H_X F_z(-H) - \alpha_X F_z(-H) + H_X \alpha F_{zz}(-H). \tag{3.4}$$

Here we note that

$$L_X(F) = \frac{\partial}{\partial z}\left\{\frac{-R_X}{R^2}F_z\right\} + \left(\frac{S}{R^2}\right)_X F$$

and since

$$R_\omega = -1, \; S_\omega = 0,$$

$$L_\omega(F) = \frac{\partial}{\partial z}\left\{\frac{F_z}{R^2}\right\} + \frac{2S}{R^3}F.$$

[2] The method to be described has some similarity to the Hellman-Feynman theorem of molecular quantum mechanics (Feynman, 1939).

Cross-multiply (2.12).F_X minus (3.2).F, and integrate top-to-bottom:

$$\left[F_X\frac{F_z}{R} - F\frac{F_{Xz}}{R}\right]_{-H}^{0} - \int_{-H}^{0} FL_X(F)\,dz - \omega_X \int_{-H}^{0} FL_\omega(F)\,dz = 0. \tag{3.5}$$

The surface values are zero, leaving

$$\frac{-1}{R(-H)}\left[F_X(-H)\,F_z(-H) - F(-H)\,F_{Xz}(-H)\right] = \int_{-H}^{0} FL_X(F)\,dz + \omega_X \int_{-H}^{0} FL_\omega(F)\,dz, \text{ or}$$

$$\frac{-1}{R(-H)}\left[F_X(-H)\,F_z(-H) + \alpha F_z(-H)\,F_{Xz}(-H)\right] = \int_{-H}^{0} FL_X(F)\,dz + \omega_X \int_{-H}^{0} FL_\omega(F)\,dz, \text{ or}$$

$$\frac{-F_z(-H)}{R(-H)}\left[F_X(-H) + \alpha F_{Xz}(-H)\right] = \int_{-H}^{0} FL_X(F)\,dz + \omega_X \int_{-H}^{0} FL_\omega(F)\,dz,$$

or, from (3.4),

$$\frac{-F_z(-H)}{R(-H)}\left[H_X F_z(-H) - \alpha_X F_z(-H) + H_X \alpha F_{zz}(-H)\right] =$$
$$\int_{-H}^{0} FL_X(F)\,dz + \omega_X \int_{-H}^{0} FL_\omega(F)\,dz. \tag{3.6}$$

We can now solve for ω_X provided merely that we can evaluate the integrals in (3.6). The first of these is of

$$FL_X(F) = F\left[\frac{\partial}{\partial z}\left\{\frac{-R_X}{R^2}F_z\right\} + \left(\frac{S}{R^2}\right)_X F\right]$$

so that after an integration by parts we have

$$\int_{-H}^{0} FL_X(F)\,dz = \left[F\left\{\frac{-R_X}{R^2}F_z\right\}\right]_{--H}^{0} + \int_{-H}^{0}\left\{\frac{R_X}{R^2}F_z^2 + \left(\frac{S}{R^2}\right)_X F^2\right\}dz, \text{ or}$$

$$\int_{-H}^{0} FL_X(F)\,dz = -\frac{\alpha R_X(-H)\,F_z^2(-H)}{R^2(-H)} + \int_{-H}^{0}\left\{\frac{R_X}{R^2}F_z^2 + \left(\frac{S}{R^2}\right)_X F^2\right\}dz \tag{3.7}$$

which is immediately computed. The second integral is

$$\int_{-H}^{0} FL_\omega(F)\,dz = \int_{-H}^{0} F\left[\frac{\partial}{\partial z}\left\{\frac{F_z}{R^2}\right\} + \frac{2S}{R^3}F\right]dz$$
$$= \left[F\frac{F_z}{R^2}\right]_{-H}^{0} + \int_{-H}^{0}\left\{\frac{2SF^2}{R^3} - \frac{F_z^2}{R^2}\right\}dz$$
$$= \frac{\alpha F_z^2(-H)}{R^2(-H)} + \int_{-H}^{0}\left\{\frac{2SF^2}{R^3} - \frac{F_z^2}{R^2}\right\}dz. \tag{3.8}$$

Combining (3.6–3.8) gives

$$\omega_X = \frac{\frac{-F_z(-H)}{R(-H)}\left[H_X F_z(-H) - \alpha_X F_z(-H) + H_X \alpha F_{zz}(-H)\right] + \frac{\alpha R_X(-H) F_z^2(-H)}{R^2(-H)} - \int_{-H}^{0}\left\{\frac{R_X}{R^2}F_z^2 + \left(\frac{S}{R^2}\right)_X F^2\right\} dz}{\frac{\alpha F_z^2(-H)}{R^2(-H)} + \int_{-H}^{0}\left\{\frac{2SF^2}{R^3} - \frac{F_z^2}{R^2}\right\} dz} \tag{3.9}$$

Eqn. (3.9) gives ω_X as required, and only needs F, integrals of F, and a knowledge of ω itself. The only problem occurs when the denominator in (3.9) vanishes, near caustics and changes from real to complex solutions. These are both legitimate breakdowns of the solution, though they occur rarely. Since in this paper we only solve the local problem, and $X = k, l$ only, we are not concerned with horizontal variation of the background fields except the topographic slope (present within α and α_X), so that the problem at each local position is only a vertical eigenvalue problem (see App. B for details).

Equation (3.9) is an important result, though complicated. In this paper, restricting X to k, l only means that the terms in H_X disappear in (3.9). In more simplified conditions it can be reduced to well-known results. For example, for a flat bottom and no flow, it is straightforward to show that $\omega_k = \omega / k$, $\omega_l = 0$. If east-west flow is added, $\omega_l = 0$ can still be demonstrated. However, in more general circumstances we have been unable to find simple relationships which can be deduced, other than the generic ones at the start of the following section.

4. Computing group and phase velocities

The frequency ω is found as an eigenvalue of (2.12) and its boundary conditions, as a function of the wavenumber (k, l) and, implicitly, of position (since $\bar{u}$, $\bar{v}$, N^2 all depend on all three spatial co-ordinates, and H depends on horizontal position). Because of the homogeneity of the system, frequency depends non-trivially only on the *orientation* ψ of the wavevector:

$$\omega = K \operatorname{fn}(\psi), \quad \text{where } \mathbf{k} = (k, l) = K(\cos\psi, \sin\psi). \tag{4.1}$$

Thus (Liu, 1999b)

$$\omega = k\omega_k + l\omega_l \tag{4.2}$$

which serves as a useful check on the computations. Eqn. (4.2) also implies that if $l = 0$, the east-west phase velocity (second definition) $\omega k a \cos \bullet / K^2$ equals the east-west group velocity $\omega_k a \cos \bullet$. This is a special case of a more general relation between phase and group velocity in the direction along the wavevector **k**. From (2.18), second definition,

$$\frac{k}{a\cos\phi}c^{\mathrm{px}} + \frac{l}{a}c^{\mathrm{py}} = \omega \tag{4.3}$$

while from (4.2) and (2.19),

$$\frac{k}{a\cos\phi}c^{\mathrm{gx}} + \frac{l}{a}c^{\mathrm{gy}} = \omega. \tag{4.4}$$

Thus the components of phase and group velocity oriented along **k** are identical for long waves. However, save in the special cases discussed at the end of the previous section, in general phase and group velocity are not identical (though they often are similar numerically); the complexity of (3.9) means that little progress can be made toward understanding these dispersive effects.

We concentrate on group rather than phase velocity in what follows, since the former is much less dependent on wavevector orientation (for example, the northwards phase velocity vanishes by definition when $l = 0$, but $\omega_l = 0$ only when $\bar{v}$ vanishes identically). Even so, the results remain formally dependent on ψ as well as position, so that some simplification has to be made for presentational purposes (especially since there are also six combinations of mean flow and topography possible, plus several vertical modes to consider).

The weak dependence on wavevector orientation found by Killworth and Blundell (2001) provides a useful simplification. At any horizontal point in the ocean, 19 equispaced orientations of the wavevector are considered, each with a westward component. The restriction to westward components is for two reasons: (a) Killworth and Blundell (2001, App. A) show that there is no solution with a locally flat bottom and a north-south-oriented wavevector; and (b) the current Appendix A shows that there is no solution with an eastward component of the wavevector, at least for a flat bottom. The argument could not be extended fully to include topographic slopes, but no occurrences of eastward phase velocity were found in any configuration. The argument in App. A does not apply to group velocity, but in fact we found only 1/3% of occurrences of any eastward *group* velocity (for the second mode), even when the full mean flow and topography were included. Henceforth we restrict attention to $k < 0$, i.e. westwards orientation of the wavevector. The 19 orientations are equispaced since the orientations which may occur for observed waves will depend crucially upon the forcing mechanism; lacking knowledge of this, we prefer to retain full generality. The orientations run between 90 + 2.9 and 270 − 2.9° to the eastward direction; the closeness to north-south orientations is necessary to include the low-latitude planetary wave behaviour in which k remains small but l increases westward (cf. Schopf et al., 1981).

For each of the 19 westward orientations, the eigenvalue problem (2.12) is solved for (real) frequency, and phase and group velocities are computed. (The methodology to do this is detailed in Appendix B.) As many modes as possible, up to four, are found, though we shall only report on the first two here (successively higher modes are confined to areas ever-nearer to the equator). A cluster algorithm is used on the group velocity vectors so produced. First, the set of group velocities is re-ordered slightly so that an entry with the maximum number of real group velocities located is first in the list. For the first orientation, a set of mean group velocities, which are simply the group velocities for that orientation, is stored. For each new orientation after the first, every group velocity is allocated to the most similar mean group velocity in the existing list of means on a one-to-one basis. This method has the advantage of robustness, in that it automatically filters out solutions which occur for unusual parameter combinations only. Such solutions would have difficulty propagating. (In the absence of mean flow, the solutions are naturally ordered by the number of zero-crossings of F_z, the eigenvector for the horizontal flow perturbation, but this ordering can disappear in the presence of mean flow. We are not aware of any systematic way to provide an ordering; which is another reason for our use of the clustering approach.) The mean and the standard deviation[3] are computed at each stage for each member of the list.

Because the number of real solutions may depend on orientation, the number of group velocities presented to the algorithm will vary (hence the initialisation by the maximum number of solutions). For some orientations, a solution may not exist. We require, arbitrarily, that a certain fraction, here one-quarter, of the 19 possible solutions must exist for a solution to be deemed valid.

3 Here computed as the obvious two-dimensional analogue of the one-dimensional standard deviation. More accurately, covariances should be stored so that an error ellipse can be produced, but for our purposes the simpler description suffices.

If not, that solution is discarded (thus removing solutions which only exist for certain wavevector directions, which would have trouble propagating).

After cluster analysis over the orientations, we have a set of modes, i.e., mean group velocities (and phase velocities, though we shall not show these) and standard deviations. The latter indicate the reliability of quoting a single mean group velocity; a small standard deviation compared with the mean value implies that the quoted value is representative, and a large standard deviation implies that the quoted value is not representative. The modes are arbitrarily numbered in order of east-west group velocity, so that the fastest westward mode is numbered 1. The N modes are properly entitled ‘normal modes’, while the Z and G modes are more correctly termed ‘shear modes’. For simplicity, here, both sets are referred to as normal modes.

5. Example profiles

a. Typical solutions

We discuss here three fairly typical solutions, for locations which are flat, slightly sloped, and steeply sloped. These locations are (a) flat: the N. Pacific (45° N, 150° W), with fractional depth gradients $H_\lambda / H = -0.06$, $H_\phi / H = 0.33$; (b) slightly sloped: the N. Atlantic (30° N, 30° W), with fractional depth gradients $H_\lambda / H = -0.55$, $H_\phi / H = -1.5$; (c) steeply sloped: the S. Indian (20° S, 70° E), with fractional depth gradients $H_\lambda / H = 8.6$, $H_\phi / H = 2.7$ (the latter chosen to have a large zonal gradient, which could amplify the effect of wavevector orientation). At each location, Fig. 1 shows the vertical structure of the first eigenmode and its vertical derivative for the GS case (full velocity and topographic gradients included), for wavevectors oriented at a subset of the 19 orientations used in the calculation, namely 120, 150, 180, 210 and 240° from east, together with the mean horizontal velocity profile.

The slopes impact on the bottom boundary condition through the value of α in (2.17). When α / H (which is nondimensional) is small, the boundary condition (2.17) becomes close to a flat bottom, while when α / H is large, (2.17) becomes close to one of zero pressure perturbation ($F_z = 0$), suggested by Tailleux and McWilliams (2001), which will be discussed below.

The flat case shows an eigenfunction F which is fairly similar across all orientations of wavevector, but which clearly resembles neither a flat-bottom solution nor a solution with zero F_z. The insensitivity of the eigenfunction is matched by the east-west group velocity, which takes the values respectively of (–0.0120, –0.0194, –0.0193, –0.0194, –0.0121) m s^{-1}, effectively independent of orientation due to the weak east-west topographic gradient.

The slightly sloping case again has an insensitive eigenfunction, but this time one which clearly does not resemble a flat bottom solution; with a relatively weak gradient at the floor, the solution is closer to one for zero derivative. The respective group velocities are (–0.0374, –0.0368, –0.0366, –0.0368, –0.0410) which again vary little, although interestingly showing a largest (westward, and 10% larger than the smallest value) value at 240° orientation, which, as Fig. 1 shows, is the mode least resembling one with zero bottom derivative.

The steeply sloping case shows a stronger effect of wavevector orientation relative to the slope direction. Three of the eigenfunctions (for 120, 150 and 180°) are all similar and satisfy a boundary condition approximating to zero F_z. Their group velocities are again similar (–0.112, –0.108, –0.105) m s^{-1}; their much larger value is because the location is closer to the equator. The other two orientations, in contrast, are qualitatively different (having a sign change in the vertical derivative), showing that at these orientations no monotonic first mode could be found. (This occurrence is familiar from the work of Rhines, 1970 for the case of no mean flow.) The group velocity at these two orientations is respectively much higher and much lower (–0.156, –0.032) m s^{-1} than for the other three orientations, showing again that little can be determined about the magnitude of the group velocity from the shape of the eigenfunction.

b. The role of the bottom boundary condition

In the absence of mean flow, Tailleux and McWilliams (2001) showed directly, and Killworth and Blundell (1999) showed indirectly, that as the bottom condition (2.17) moves from a condition of vanishing F towards a condition of vanishing F_z, the faster would be the westward phase velocity. Even under these restrictive conditions it is not known what the effect on the group velocity would be, and if mean flow is included even the induced change in phase velocity is not known. Tailleux and McWilliams then argued that the bottom condition is not well defined, since it is not obvious what length scales in bottom topography are relevant (though Part II will show that the changes in solutions produced by changing the smoothing length scale for topography are small). They thus suggested that the observed speed-up of westward propagation could be accounted for by replacing the sloping bottom condition with that of vanishing F_z. It is of interest to know how close, in some sense, our solutions come to satisfying such a boundary condition (and if they do, whether there is a speed-up associated with it). In the case of no mean flow, the answer is clear from Killworth and Blundell (1999). Since they found very little net speed-up across an ocean basin, then most of the world ocean cannot have slopes such that the sloping bottom boundary condition (2.17) resembles that of vanishing F_z. (It is possible *a priori* that since their solutions were for rays emanating from the eastern boundary, that modes might have changed during the ray propagation towards the second – and hence slower – mode. However, we shall see in Part II that this is not the case, at least for the full GS problem.) The cases shown in Fig. 1 also demonstrate that in the presence of mean flow, smaller bottom gradients are not necessarily associated with faster group velocity.

This can be made slightly more rigorous by examining the size of α. The operant value of α at a location depends on the wavevector orientation and hence on the mechanism producing the planetary waves. An average measure of α can be computed across the 19 orientations of wavevector, which of course cancels the l / k term by symmetry, giving a nondimensional estimate of α as $\bar{\alpha} = \tan\theta \, H_\theta / H$, perforce losing all longitudinal depth gradient information. Fig. 2 shows a histogram of the rates of occurrence of $\bar{\alpha}$ in the world ocean in the ‘planetary wave region’[4], over 1° squares and α bins of 0.1, with the number of values above 20 in magnitude indicated at the edge. (If the entire ocean is included, the large values of $\tan\theta$ at high latitudes strongly increase the number of large average α values, but these latitudes lie beyond most turning latitudes for annual and even biennial frequencies.) Killworth and Blundell’s (1999) calculations for a northward topographic gradient suggest that $\tan\theta \, H_\theta / H$ would have to be of order 2 or more for significant wave speed changes. If this remains true in the presence of mean flow, then Fig. 2 shows that the

[4] Defined as the world ocean between ±5 and ±50°, since only very long period waves are observed at higher latitudes.

majority of values of $\bar{\alpha}$ are not large. Visual inspection of contours of $\bar{\alpha}$, however, show that there are preferred locations for large values: the N. Atlantic between 30 and 50° N, small areas of the W. Pacific, and portions of the S. Indian at 30–50° S. So while there are coherent oceanic areas in which there are consistently high values of $\bar{\alpha}$, these do not occupy much of the ocean equatorward of 50°.

In general the ratio l / k must be found from ray theory, so that estimates of α cannot proceed. A very rough guess for the ratio, suggested by Tailleux, can be made, estimating the wavenumbers from the classical flat-bottom calculation, though because this is for a different problem, its validity is suspect. For rays starting at an eastern boundary (see Part II),

$$k \approx \frac{\omega}{c_{flat}}, \; l = -\frac{\omega \, dc_{flat} / dy}{c_{flat}^2 (y)} a \cos \phi \, (\lambda - \lambda_E)$$

where λ_E is the assumed starting longitude of the ray. The dashed curve in Fig. 2 shows the values of α derived from this approximation and from the local flat-bottomed solutions computed here.[5] The eastern half of all ocean basins now has α / H values under 2 in modulus, while the western side has values over 2 due to the linear increase in l westward. However, the sign of α in these latter regions changes one or more times with distance westward, and so α passes through zero, which would make problems for the propagation of a mode tied to high α values of one particular sign.

As a final test, the global calculations of the next section were repeated using depth gradients artificially increased by a factor of 10 over their ‘true’, i.e. 1° smoothed, values. A correlation of the artificial group velocities, point by point, against their original values shows that the original values are approximately 0.87 of the artificial ones (this holds for two separate calculations: including all velocities up to 1 m s^{-1} and including only velocities up to 0.01 m s^{-1}). If our solutions were essentially those with no bottom pressure perturbation (and if such were the fastest group velocities available), then increasing bottom slopes would have no effect on the solutions.

5 Rather than tracing some measure of the eastern boundary, we have defined a pseudoboundary as the easternmost point of the three longitude bands defined in the figure caption.

We conclude, therefore, that it is difficult to draw any straightforward conclusions about the behaviour of individual solutions. The problem is complicated, and global computations must therefore be made.

6. Global results

Solutions were found for the six cases in Table 1, using values of topography (and its gradient) smoothed with a 1° length scale. The baroclinic fields and buoyancy frequency are computed directly from the World Ocean Data atlas (Antonov et al., 1998; Boyer et al., 1998) and so have been horizontally smoothed during the production of those data. Since the effects of lateral smoothing are felt most strongly in the depth gradients, and it is far from clear at what scale long waves 'see' topography, a set of values with 3° smoothed data was also computed. The results are almost indistinguishable by eye from the 1° results to be presented here. Thus our results are robust to detail in the fields on these scales.

Several general statements can be made concerning the results.

First, over the vast majority of the ocean, results for the first vertical mode using the zonal mean flow (Z in our notation) overwhelmingly resemble their general mean flow counterparts (G). (The arguments of Killworth and Blundell (2001) can be used to give a partial reason why this may be the case, though extending the arguments to bottom slope requires weak east-west slopes, which are not observed.) This finding applies less strongly for the second mode.

Second, the group velocities are largely dominated by the east-west component. For example, for the first vertical mode in the fully general GS case, in 94% of the planetary wave region, $|c^{gx}| / |\mathbf{c}^{g}|$ exceeds 0.9; for the second mode, the figure is slightly smaller, at 87%.

Third, over most of the ocean, the (mean) group velocities found by this approach indeed possess little dependence on wavevector orientation (i.e., have small standard deviations). For the first vertical mode, in 84% of the world ocean in the planetary wave region, the coefficient of variation (i.e., the mean group speed divided by the standard deviation) was more than 1.96, so that each group velocity in the list can be thought of as statistically different from zero. In 52% of the ocean, the coefficient of variation exceeded 5, so that the mean group velocity provides a very accurate estimate independent of orientation of the wavevector. Fig. 3 shows the spatial structure of this ratio. The locations where the ratio is smaller than either 5 or 1.96 tend to be near steeper topography (mid-ocean ridges, etc.), in higher poleward latitudes, and near ocean boundaries. For the second mode, a similar but weaker result was found: in 76% of points, the coefficient of variation exceeded 1.96, and in 43% of locations exceeded 5. (There are fewer locations where a real second mode can be found). Thus *for about three-quarters of the world ocean, we can think of there being a single well-defined group velocity for each vertical mode*. This permits us, henceforth, to quote only mean values with only minor caveats in what follows.

a. The first vertical mode

The first-mode westward group velocity has already been shown for the NF and ZF cases in Killworth et al. (1997), subject to minor changes due to small differences in the datasets used.[6] Since all such diagrams look similar (they are dominated by the latitudinal variation due to f), we show only the fully general GS case in Fig. 4. The results are slightly noisier than those found by Killworth et al. (1997), and show a consistent poleward movement of contours relative to their (ZF) case. The north-south group velocity is, as noted, mainly much smaller than the westward component. Fig. 5 shows this clearly (experiments with contouring direction were less clear than this representation owing to noisiness). Almost the only locations in which the group velocity is not strongly directed westwards are those over topographic features such as mid-ocean ridges (which can almost be picked out by eye from the figure).

This fully general case is a combination of mean flow and bottom topographic effects. Killworth et al. (1997) found a speed-up of the westward (phase) speed over the traditional zero-flow, flat-bottom normal mode speed, which was sufficient to explain most of Chelton and Schlax's (1996) observations of an increase in phase speed. However, Killworth and Blundell (1999) found little systematic effect on phase or group speed from topographic effects alone. It is therefore interesting that the combination of *both* effects in the GS results demonstrates a further

6 Since in those calculations $\bar{v}$ vanished identically, $\omega_l = 0$ and so east-west group and phase velocity are identical.

speed-up of the westward group velocity over that already found by Killworth et al. for the ZF case. Fig. 6 summarises this, by showing a logarithmic plot of the longitudinally averaged westward group velocity as a function of latitude. (In all summary plots, it should be remembered that the number of points entering the average decreases in high latitudes where fewer solutions could be found.) The results fall into three groups, of increasing speed. The slowest group is the no-flow solutions, with and without topographic slopes (NF, NS). The next fastest speeds are ZF (not shown) and GF, i.e., inclusion of mean flow but retaining a locally flat bottom. The fastest speeds are found for ZS (not shown) and GS. The differences are small near the equator, but become much more noticeable poleward. Contour plots of the group velocity ratio GS/NF (not shown) indicate that there is much variation of the ratio with longitude, including small areas of ratios under unity even at high latitudes.

Figure 7 uses the same data to show ratios of the mean speeds to those in the NF case (i.e., to the traditional normal mode calculation). An almost identical diagram is produced if one plots the ratio of the means instead of the mean of the ratios, save at high latitudes where the signal is noisy. The speed-up found by Killworth et al. (1997) and later authors is represented by the dashed line (GF, almost identical to ZF which is not shown). As Killworth and Blundell (1999) found, there is little net speed-up by topographic effects alone (dash-dotted line, NS). However, the second speed-up when both mean flow and topographic gradients are included is clearly evident; it approximately increases in magnitude poleward, reaching values of order 2 in both hemispheres.

It can be argued that a more careful interpretation of longitudinally-averaged speeds should use the harmonic mean rather than the arithmetic, since what is observed is a travel time, proportional to an integral of the inverse speed (Tailleux, personal communication). The results of this (not shown) are essentially similar, though made more noisy by small areas of low velocity which contribute heavily to the harmonic mean (in addition, the NS/NF ratio is mainly under unity).

b. The second vertical mode

Figure 8 shows contours of the second-mode east-west group velocity, again for the GS case. It resembles the first-mode in structure, though of course with much reduced speeds, and a further encroachment of areas where no solution could be found. Fig. 9 shows the direction of the group velocity. This shows more variability from the strongly westward orientation of the first mode, though still dominated by the westward direction, and indicates the location of the 1/3% of vectors with an eastward component.

Figure 10 shows the zonally-averaged group speeds for the second mode. The slowest speeds remain, as before, the NF and NS cases (i.e., those with no mean flow). However, all the remaining cases (whether zonal or full mean flow, and whether there is a locally flat or sloping bottom) show similar speeds, showing a strong degree of speed-up over the linear normal mode, except near the equator. Fig. 11 shows the ratios to the NF case. Here there is a systematic separation of the different cases. The smallest ratio, never differing far from unity, remains the NS case. The cases with mean flow all have much larger ratios, and in increasing order of ratio (apart from a small amount of noise at high latitudes) are: GF/NF, ZF/NF, GS/NF, and ZS/NF, so that the maximum speed-up experienced by the second mode occurs for zonal flow and sloping topography. The additional speed-up for mean flow and topography over mean flow alone is, however, much smaller for this second mode than for the first. For the full case (GS), ratios of over 2 are typical poleward of 35°S. (The ratio of the means is again similar to the mean of the ratios, and is not shown.)

7. Discussion

This paper has sought to produce as general a theory as possible for the propagation of small-amplitude long planetary waves, subject only to the WKBJ assumption of slowly-varying background fields such as baroclinic shear and topography. These are necessary assumptions without which it would make little sense to compute local solutions. However, they are not necessarily well satisfied in the real ocean. Killworth et al. (1997) give some discussion of variability in the baroclinic signal, and Killworth and Blundell (1999) do the same for topographic slopes. The test mentioned earlier (with 3° smoothing rather than 1°) shows that results are essentially independent of any *larger*-scale variability; and, indeed, 1° is quite a small smoothing scale for global WKBJ theory. Nonetheless, more work needs to be done on whether (and how) background flow and topographic variation on scales small compared with the waves can interact with the waves themselves on a global basis.

The second speed-up of westward propagation induced by the combination of baroclinic mean flow and topographic slopes has been mentioned several times without an explanation being offered. It is hardly likely that *any* topographic slope will act to speed up a planetary wave (certainly this is not the case without mean flow, cf. Killworth and Blundell, 1999). Why, then, should there be such a consistent speed-up? In the earlier case of the speed-up induced by baroclinic mean flow over a locally flat bottom, Dewar (1998) and de Szoeke and Chelton (1999) were able to offer suggestions based on potential vorticity structures. These structures had their basis in the dynamics of the large-scale ocean circulation. Modifying this theory to include what can be thought of as an arbitrary bottom slope seems difficult.

As an experiment, we re-examined the three typical ocean locations used in Fig. 1. At each of these, the absolute bottom slope was computed. Then a collection of modified bottom slopes was created, by taking the given slope and orienting it along an angle varying between 0° and 360° to the eastward direction. For each angle, the eigenvalue problem (2.12) was solved (over a variety of wavevector orientations again) and the mean group velocities computed as before. Solutions for the three locations are shown in Fig. 12. The flat area (45° N, 150° W) shows a noticeable degree of variability of the first mode speed with angle: some orientations of topography speed the wave up over its locally flat-bottomed GF value, and some slow it down [plus, at 234°, what may be close to a zero of the denominator of ω_k in (3.9)]. The actual orientation (shown by the vertical dashed line) is such as to slow the wave down over its locally flat-bottomed GF value. The second mode speed is almost independent of slope orientation. The slightly sloped area (30° N, 30°W) again shows slope orientations both increasing and decreasing the first-mode wave speed; the actual orientation is such as to speed up the first mode, but has little effect on the second. The steeply sloped location (20° S, 70°E) is less typical among those we have examined, in that *most* orientations of the topographic slope increase the wave speed over its GF value, though this only holds for the first mode.

There is thus little obvious reason why the speed-up appears so consistent, unless, of course, the topographic slope has acted to produce a mean flow which is such as to induce a speed-up of planetary waves. This is an intriguing possibility worthy of study, but the set-up of the mean flow must remain beyond the scope of this paper.

The lack of eastward group velocities in our results means that explanations for eastward propagating sea surface temperature signals (e.g., Sutton and Allen, 1997, but with caveats discussed in Killworth and Blundell, 2001) cannot rely on ocean-only dynamics unless the signal is density-compensated, at least at the large-scale and for real solutions (except perhaps at high latitudes where the barotropic Doppler shift can operate). Smaller-scale features such as mesoscale eddies, etc. could of course be involved, but are beyond the scope of this paper. Thus the 'information loop' within the ocean must be closed by Kelvin or coastal waves taking the planetary wave signal at a western boundary and re-propagating it around the basin circumference to re-initialise the wave. No pseudo-gyral dynamics are permitted by our results.

Acknowledgments. Our thanks to the SOC satellite team for their enthusiasm and keenness to compute strange quantities (almost) on demand, and to an anonymous referee and to Rémi Tailleux for their helpful comments and queries. This work is part of the James Rennell Division Core Project CSP1.

APPENDIX A

Existence of solutions

This appendix proves three results concerning the existence of real solutions.

a. Non-existence of critical layers

Suppose that ω is such that R vanishes, for some z_0, $0 > z_0 > -H$, i.e., that there is a real critical layer within the fluid column. Near $z = z_0$, $R \sim A(z - z_0) + \ldots$. If R had no higher terms, and S were independent of depth, the two solutions of (2.12) would simply be

$$(z - z_0)\left\{\mathrm{J}_2\left(2\left[S(z - z_0)\right]^{1/2}\right),\ \mathrm{Y}_2\left(2\left[S(z - z_0)\right]^{1/2}\right)\right\},$$

where J_2, Y_2 are Bessel functions of second order. The first of these solutions is well behaved on both sides of the zero of R. The expansion of the second of the solutions begins with a constant, and contains a term proportional to $(z - z_0)^2 \ln|z - z_0|$. The solution would not be physical since the vertical velocity is proportional to F, the pressure is proportional to F_z and so the density is proportional to F_{zz}, containing a term in $\ln|z - z_0|$, clearly unphysical.

Now it is possible that the higher terms in the expansion of R and S about the zero of R could be such as to remove the logarithmic term in the expansion of F, and render the second solution physical. In general, however, this cannot occur. To see this most easily, evaluate (2.12) and $\partial / \partial z$ (2.12) at z_0, assuming only well-behaved solutions:

$$\{RF_{zz}\} - R_z F_z + SF = 0,\ z = z_0$$

$$\{RF_{zzz}\} - R_{zz} F_z + S_z F + SF_z = 0,\ z = z_0$$

where the terms in curly brackets vanish at z_0 for well-behaved solutions. Thus F and F_z must satisfy two homogeneous equations at $z = z_0$:

$$SF - R_z F_z = 0$$

$$S_z F + (S - R_{zz}) F_z = 0.$$

In general, these have no solution other than the null solution, unless the determinant of the coefficients vanishes, i.e., if

$$\left(\frac{R_z}{S}\right)_z = 1,\ z = z_0. \qquad (A1.1)$$

If this occurs at all, it can only occur at some specific value of z_0. Unlike the numerical approach in Appendix B, we here take the usual view that the eigenvalue problem is one for ω given k, l. Until the boundary conditions are applied to close the eigenvalue problem, z_0 depends on ω (as well as on k and l).

To close the problem, then, two cases can occur. The first, more usual, case is that only one well-behaved solution exists at the critical layer, when $F \sim (z - z_0)^2$. This solution can only vanish at the surface (satisfying the boundary condition there) for some specific z_0 (in turn implying a specific ω). But there are no further degrees of freedom left to satisfy the bottom boundary condition (save in exceptionally special circumstances). In the second, special, case (A1.1) holds above, and both well-behaved solutions exist, but for a pre-specified eigenvalue ω [z_0 and ω are jointly determined by the requirements that R must vanish and (A1.1) must hold]. The boundary conditions at top and bottom, both homogeneous, must then additionally both be satisfied. With no degrees of freedom, again in general the solution can only be null.

Thus any physically acceptable solution with an internal critical layer must be have complex frequency (and so not really possess a critical layer). Killworth et al. (1997) reported cases of complex solutions (typically at high latitudes), but the solution method employed here cannot locate complex solutions. We concentrate also on group velocities for robustness, and the definition of group velocity when the mode concerned has a complex frequency still remains ill-defined (cf. the review by Pierrehumbert and Swanson, 1995, which shows that the full non-longwave dispersion relation would be needed).

b. The sign of R

Since the mean flow is baroclinic,

$$\int_{-H}^{0} Q dz \equiv \int_{-H}^{0} \left(\frac{k\bar{u}}{a \cos \phi} + \frac{l\bar{v}}{a} \right) dz = 0,$$

it follows that Q takes both positive and negative values in the water column. Further, as $\omega > 0$ by supposition, $R = Q - \omega$ must be somewhere negative in each fluid column. We showed above that R cannot vanish for real solutions. It therefore follows that

$$R < 0 \text{ everywhere.}$$

c. Non-existence of eastward propagating waves for a flat-bottomed ocean

Multiply (2.12) by F and integrate from bottom to top:

$$\left[\frac{FF_z}{R} \right]_{-H}^{0} - \int_{-H}^{0} \frac{F_z^2}{R} dz + \int_{-H}^{0} \frac{SF^2}{R^2} dz = 0. \qquad (A1.2)$$

The first term vanishes at the surface, and for a flat-bottomed ocean also vanishes at the floor. The second term is positive (since $R < 0$ has just been proven). Thus the third term must be negative. The only adjustment to the sign can be made by selecting k, which must therefore be *negative* to satisfy (A1.2). Thus there can be no flat-bottomed waves propagating with an eastward phase velocity (and indeed, if there is no north-south mean flow, the same holds for group velocity).

This has immediate ramifications for simplified solutions such as those by Liu (1999a, 1999b) and Killworth and Blundell (2001). These articles both found higher vertical modes with eastward propagation (and frequently possessing what would be real critical layers as well) using, respectively a 2.5-layer ocean and a modal decomposition. *These solutions cannot occur in the continuous case.*

If the ocean possesses topography, then the first term in (A1.2) gives an additional term, $\alpha F_z^2 / R$, evaluated at the floor. Thus if $\alpha < 0$ the same restriction against eastward orientation of phase velocity holds. Thus an eastward orientation of the phase velocity can only occur when $\alpha > 0$, which depends on the orientation of the wave vector (k, l), the topographic gradient, and the sign of the latitude.

APPENDIX B

The data treatment and numerical method

We use the 1998 World Ocean Atlas (Antonov et al., 1998; Boyer et al., 1998) for temperature and salinity data, and the ETOPO5 dataset (National Geophysical Data Center, 1988) for topography. The ETOPO5 data were smoothed by applying a Lanczos sigma factor filter (Lanczos, 1957, 1966) to the 2-D Fourier transform of a 1/4° average of the original data (averaged for storage reasons), and the inverse transform is averaged onto the same 1° squares as the WOA data. The smoothed FFT method is used so that we can obtain a consistent set of H and its derivatives. This paper only requires H and its first derivatives, but the ray tracing and caustic checking of part II will require up to third derivatives, and we wish to use the same method in both papers. With a 1° filter width, the resulting topography is very close to the simple averaged topography used in Killworth et al. (1997).

We use the UNESCO 1981 equation of state (Gill, 1982) to compute $N^2(z)$ at the reference levels, and the thermal wind equations to derive the vertical velocity shear, from which we infer baroclinic velocity components u and v by requiring their vertical integrals to vanish. These calculations are done at all 1° gridpoints with depth at least 1000 m. We thus produce tabulated fields suitable both for splining to higher resolution in the vertical, and for bilinear interpolation in the horizontal.

We now wish to compute ω_X from (3.9), for $X = (k, l)$. This entails first solving (2.12) subject to (2.16) to give F, F_z and F_{zz}. We choose a direction ψ for the wavevector, and adjust the value of $|\mathbf{k}|$ using a NAG library (Numerical Algorithms Group Ltd, 1999) zero-finding routine until the bottom boundary condition – either (2.17), or the simpler condition $F(-H) = 0$ for the flat-bottom case – is satisfied. Eqn. (2.12) is integrated downwards using a simple predictor-corrector scheme with up to 126 gridpoints for both F and F_z, viewed as elements of a coupled ODE system. Both F and F_z are tabulated on the refined vertical grid onto which N^2, u and v have been splined within the water column, and at the bottom. Several numerical schemes involving different tabulations of F and F_z, and different integration schemes were tried before settling on that described, which gave the most consistent approximation to the bottom boundary condition as H varied with position. (Note that this required a variable-length step from the last tabulated point to the bottom.)

Having thus obtained F and its vertical derivatives, we can compute all the terms in (3.9) using the trapezoidal approximation for the integrals, and thus obtain the group velocity. Note that for $X = (k, l)$ the derivatives of R and S appearing in (3.9) do not involve the horizontal derivatives of N^2, u or v, only those of H (via α and α_X).

TABLE 1

Combinations of mean flow and topography

	Topography	
Mean flow	Locally Flat (F)	Topographic Slopes (S)
None (N)	NF	NS
Zonal mean flow only (Z)	ZF	ZS
General mean flow (G)	GF	GS

REFERENCES

Antonov, J., Levitus, S., T. P. Boyer, M. Conkright, T. O'Brien, C. Stephens, 1998: *World Ocean Atlas 1998 Vols. 1-3: Temperature of the Atlantic/Pacific/Indian Ocean*. NOAA Atlas NESDIS 27. U.S. Gov. Printing Office, Wash., D.C., 166 pp.

Boyer, T. P., Levitus, S., J. Antonov, M. Conkright, T. O'Brien, and C. Stephens, 1998: *World Ocean Atlas 1998 Vols. 4-6: Salinity of the Atlantic/Pacific/Indian Ocean*. NOAA Atlas NESDIS 30. U.S. Gov. Printing Office, Wash., D.C., 166 pp.

Chelton, D. B., and M. G. Schlax, 1996: Global observations of Oceanic Rossby Waves. *Science*, **272**, 234-238.

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.

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

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

Feynman, R. P., 1939: Forces in molecules. *Phys. Rev.*, **56,** 340.

Fu, L. L., and D. B. Chelton, 2001: Large-scale ocean circulation. *Satellite altimetry and earth sciences*, Fu, L. L., and A. Cazenave, Eds, Academic Press, 133-169.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press, 662 pp.

Kang, Y. Q., and L. Magaard, 1979: Stable and unstable Rossby waves in the North Pacific Current as inferred from the mean stratification. *Dyn. Atmos. Ocean.*, **3**, 1-14.

Killworth, P. D., and J. R. Blundell, 1999: The effect of bottom topography on the speed of long extra-tropical planetary waves. *J. Phys. Oceanogr.*, **29**, 2689-2710.

Killworth, P. D., and J. R. Blundell, 2001: Large-scale propagating disturbances: approximation by vertical normal modes. *J. Phys. Oceanogr.*, **31,** 2852-2870.

Killworth, P. D., 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.

Lanczos, C., 1957: *Applied Analysis*. Pitman, 539 pp.

Lanczos, C., 1966: *Discourse on Fourier series*. Oliver and Boyd, 255 pp.

Liu, Z., 1999a: Planetary wave modes in the thermocline: non-Doppler-shift mode, advective mode and Green mode. *Quart. J. R. Meteorol. Soc.*, **125**, 1315-1339.

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

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

Numerical Algorithms Group Ltd, UK, 1999: Nag Fortran Library Manual, Mark 19, Numerical Algorithms Group Ltd, Oxford, UK.

Pierrehumbert, R. T. and K. L. Swanson, 1995: Baroclinic instability. *Ann. Rev. Fluid Mech.*, **27**, 419-467.

Rhines, P. B., 1970: Edge-, bottom-, and Rossby waves in a rotating stratified fluid. *Geophys. Fl. Dyn.*, **1**, 273-302.

Samelson, R. M., 1992: Surface-intensified Rossby waves over rough topography. *J. Mar. Res.*, **50**, 367-384.

Schopf, P., D. L. T. Anderson, and R. Smith, 1981: Beta-dispersion of low frequency Rossby waves. *Dyn. Atmos. Ocean.*, **5**, 187-214.

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

Tailleux, R., and J. C. McWilliams, 2001: The effect of bottom pressure decoupling on the speed of extratropical, baroclinic Rossby waves. *J. Phys. Oceanogr.*, **31**, 1461-1476.

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

Tyler, R. H. and R. Käse, 2001: A string function for describing the propagation of baroclinic anomalies in the ocean. *J. Phys. Oceanogr.*, **31**, 765-776.

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

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.

Captions

1. Vertical structure of eigenvector F (corresponding to the vertical velocity) and its vertical derivative F_z (corresponding to pressure) for locations which are (a) flat (45° N, 150° W), (b) slightly sloped (30° N, 30° W), and (c) steeply sloped (20° S, 70° E). At each location, five orientations of the wavevector are considered: 120, 150, 180, 210 and 240° from east. F has been arbitrarily normalized to a maximum value of unity; F_z, for convenience, has been renormalized to make its maximum modulus also be unity (so that some curves do not now have the original normalisation $F_z(0) = 1$). The right-hand panel shows the mean horizontal velocity profile at that location, in m s^{-1}. The eastward component of group velocity differs between the orientations of **k** and is given in the text.
2. Histogram of the number of occurrences (counted by 1° × 1° areal bins and 0.1 bins for α) of a measure of the nondimensionalised average value of α, i.e., $\tan\theta H_\theta / H$, (firm line) and the nondimensionalised value of $\alpha = \tan\theta \left(H_\theta - lH_\lambda / k\right) / H$, using an approximate expression in the text for l / k (dashed line) in the world ocean between 5 and 50° latitude. The area for the latter calculation is limited to 40–100°, 160–260°, and 300–360° E. The larger black circles show accumulated counts beyond the ±20 extrema for the average α, the smaller ones for the approximate expression.
3. The coefficient of variation (mean group velocity divided by its standard deviation over the wavevector orientations where it is defined) for the first internal mode. Contours are 1.96 (ratios above this are statistically different from zero) and 5 (an arbitrary value for which the mean group velocity can be considered independent of wavevector orientation). Light grey areas represent regions where a real solution could not be found.
4. Mean east-west first-mode group velocity, for the GS (general mean flow, topographic slopes) case, in m s^{-1}. Contour intervals are nonuniform: 0.30, 0.20, 0.15, 0.10, 0.08, 0.06, 0.04, 0.02, 0.01 m s^{-1} westward for comparison with Killworth et al. (1997). Values are masked within 5° of the equator, where equatorial, rather than long planetary wave, theory should hold. Light grey areas represent regions where a real solution could not be found.