Reply

ROBERT HALLBERG
NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

2 May 2000 and 6 September 2000

1. Introduction

Hallberg (1997) used a combination of quasigeostrophic ray tracing and primitive equation numerical simulations to study the propagation of topographic and planetary Rossby waves around an ocean basin with slowly varying but arbitrary topography. When the topographic and planetary vorticity gradients are not parallel to each other, the classical baroclinic and barotropic modes are replaced by modes whose vertical structure varies strongly with the orientation of the wavevector. The low-frequency motions are more aptly described in terms of surface- or bottom-intensified modes. Ray tracing suggests that the surface- and bottom-intensified motions should be coupled in localized regions because of the changing physical nature of a single branch of the dispersion relation. At low frequencies and long wavelengths, Hallberg (1997) found that the coupling was confined to regions where the topographic slopes become nearly meridional. But the zonal changes in the orientation of the slopes were essential to this coupling. The assumptions leading to the WKB approximation are strongly violated in the coupling locations in the examples considered by Hallberg (1997), and WKB theory does not provide a useful estimate for the strength of the coupling, either between surface- and bottom-intensified motions or between the two branches of the dispersion relation. At low frequency, an adiabatic constraint does give an estimate for the strength of the coupling. Hallberg (1997) quantitatively verified this estimate for the strength of the coupling, as well as the location of the coupling, with a series of primitive equation numerical simulations.

Vanneste (2001) points out that a well-established extension of WKB theory can be used to estimate the strength of coupling between wave modes. Vanneste then reexamines the strength of coupling between wave modes in the special case of meridionally varying, meridionally sloping topography. Vanneste argues, based on expressions that apply to this special case and the assertion that this extension of WKB theory is generally applicable to the ocean, that the estimates for the coupling strength between surface- and bottom-intensified motion in Hallberg (1997) are incorrect.

In this reply, a more appropriate, but still analytically tractable, physical model is proposed to describe the coupling regions for low-frequency topographic and planetary waves. This alternate simple model is examined using essentially the same techniques as in Vanneste (2001). It is explicitly demonstrated that the WKB approximation is not justified in these regions. But the extension of WKB theory cited by Vanneste indicates the likelihood of strong (but not total) coupling between modes; with an appropriate model, these techniques do not contradict the original Hallberg (1997) estimate of the coupling strength between modes (and between layers). This coupling strength has already been verified in a series of primitive equation numerical simulations, whose only assumption is that the aspect ratio is small (as is true of the ocean). The physical argument for estimating the strength of coupling in Hallberg (1997) is clarified in response to (erroneous) criticisms in Vanneste (2001). In this reply, it is shown that while the mathematical techniques applied by Vanneste are quite general and valuable, the physical conclusions that he draws regarding the ocean circulation stem from a very specific physical model, and the physical description from Hallberg (1997) of localized coupling between surface- and bottom-intensified low-frequency flow and between wave modes in the ocean is valid.

2. Rossby waves over zonally varying topography

Vanneste (2001) considers only the case where topographic slopes are purely meridional and vary only in the meridional direction. This is the most mathematically tractable orientation of the topography, and it has been extensively discussed by Veronis (1980), for example. It is also a relatively uninteresting limit. In this one special limit, at any location each of the two branches of the dispersion describes a single physical mode [as can be seen in Figs. 1a and 4c of Hallberg

Corresponding author address: Dr. Robert Hallberg, Geophysical Fluid Dynamics Laboratory, Princeton University, Forrestal Campus, U.S. Route 1, P.O. Box 308, Princeton, NJ 08542.
E-mail: rwh@gfdl.gov

(1997)]. The mode conversion that Vanneste describes does not apply to low-frequency long waves since these propagate zonally in this special case. By contrast, for nonmeridional topographic slopes, a single branch of the dispersion relation describes both surface- and bottom-intensified motions at different orientations of the wavevector. In short, the special case described by Vanneste (2001) is very different from the general case discussed in Hallberg (1997).

The vertical structure of mixed topographic–planetary waves in a two-layered fluid depends strongly on the relative magnitudes of the potential vorticity gradients perpendicular to the wavevector (Hallberg 1997). Since the typical length scale of topographic slopes is much less than the radius of the earth, topographic potential vorticity gradients at the margins of an ocean basin are much larger than planetary gradients. The ratio of the lower-layer topographic and planetary vorticity gradients over continental margins is $(R_{\rm earth}\alpha\tan(\theta))/H_2$, typically of order 30 in midlatitudes (for slope $\alpha = 0.01$, lower-layer thickness $H_2 = 2$ km, and latitude $\theta = 45°$). So, even slowly changing orientations of topography with respect to the wavevector can be expected to lead to abrupt changes in the vertical structure associated with a wave mode. Furthermore, in the ocean the complicated coastlines essentially guarantee that every topographic ray will encounter the regions where significant coupling between wave modes and between surface- and bottom-intensified motions will occur.

Low-frequency, long planetary waves are of primary importance for understanding the ocean circulation, as evidenced by the prevalence of theories based on the planetary geostrophic equations, which essentially consider only waves in this limit. Hallberg (1997) described mode conversion and coupling between surface- and bottom-intensified flow in the low-frequency long-wave limit. The coupling occurs near the point where the zonal planetary vorticity gradient goes from being positive to negative (or, in other words, in regions of meridional slopes). The changes along a ray in the relative orientation of the planetary and topographic vorticity gradients are important to the coupling of low-frequency long waves. This can be clearly illustrated in a model with a zonally varying zonal slope.

The two quasigeostrophic vorticity equations can be combined to give a quartic equation in the zonal wavenumber k. This equation is too complicated to give any physical insight. But, in the coupling regions, low-frequency waves can be qualitatively described as having long zonal wavelengths compared to the deformation radius. This yields a quadratic equation for k. With this approximation, the two equations are

$$\frac{\partial}{\partial t}\left[\frac{\partial^2\psi_1}{\partial y^2} - \lambda_1^{-2}(\psi_1 - \psi_2)\right] + \beta_{1,y}\frac{\partial\psi_1}{\partial x} = 0 \quad (2.1)$$

$$\frac{\partial}{\partial t}\left[\frac{\partial^2\psi_2}{\partial y^2} - \lambda_2^{-2}(\psi_2 - \psi_1)\right] + \beta_{2,y}\frac{\partial\psi_2}{\partial x} - \beta_{2,x}\frac{\partial\psi_2}{\partial y} = 0 \quad (2.2)$$

near the coupling regions. The notation here is the same as in Hallberg (1997), except that the zonal and meridional components of β gradients are explicitly separated. By assumption the layer deformation radii λ_n and

$$\beta_{n,y} = \frac{df}{dy} - \frac{f}{H_n}\frac{dH_n}{dy}$$

are constant, but

$$\beta_{2,x} = -\frac{f}{H_2}\frac{dH_2}{dx}$$

is a linear function of x. These assumptions are reasonable for a local approximation to the northern or southern end of a circular, bowl-shaped basin, but they should apply equally well to the pertinent area for low-frequency coupling in an arbitrary basin.

Assuming a wavelike structure $\psi_n(x, y, t) = \Psi_n(x) \exp\{i[\int k(x)\,dx + ly - \omega t]\}$, where l and ω are constants, but k and Ψ_n can vary slowly in x (to the extent that the WKB approximation holds), (2.1) and (2.2) can be manipulated to give

$$\begin{aligned}\beta_{1,y}(k - k_1)\Psi_1 &= \omega\lambda_1^{-2}\Psi_2,\\ \beta_{2,y}(k - k_2)\Psi_2 &= \omega\lambda_2^{-2}\Psi_1.\end{aligned} \quad (2.3)$$

Here

$$\begin{aligned}k_1 &\equiv -\omega\frac{(l^2 + \lambda_1^{-2})}{\beta_{1,y}},\\ k_2 &\equiv -\omega\frac{(l^2 + \lambda_2^{-2})}{\beta_{2,y}} + \frac{\beta_{2,x}}{\beta_{2,y}}l\end{aligned} \quad (2.4)$$

are the long zonal wavenumbers that the waves would have without coupling between the layers. Equations (2.3) can be combined to give a quadratic equation for k, the solutions of which are

$$k = \frac{1}{2}\left[k_1 + k_2 \pm \sqrt{(k_1 - k_2)^2 + 4\omega^2/(\lambda_1^2\lambda_2^2\beta_{1,y}\beta_{2,y})}\right]. \quad (2.5)$$

One of the criteria for the WKB approximation to hold is that $(k^{-2}dk/dx \ll 1)$ (Lighthill 1978). The only variable appearing on the right-hand side of (2.5) that is not constant (by assumption) is k_2. So,

$$\frac{dk}{dx} = \frac{dk_2}{dx}\frac{1}{2}\left[1 \pm \frac{(k_2 - k_1)}{\sqrt{(k_1 - k_2)^2 + 4\omega^2/(\lambda_1^2\lambda_2^2\beta_{1,y}\beta_{2,y})}}\right], \quad (2.6)$$

where

$$\frac{dk_2}{dx} = \frac{l}{\beta_{2,y}}\frac{d\beta_{2,x}}{dx} = \frac{l}{R_C}, \quad (2.7)$$

and R_C is the radius of curvature of the topography, assuming that topographic potential vorticity gradients are much larger than planetary gradients.

Over topography that slopes up toward the equator, the second term inside the square root in (2.5) is negative, and the entire argument can become negative. Since $\sqrt{\beta_{2,x}^2 + \beta_{2,y}^2} \gg \beta_{1,y}$, this occurs roughly a distance

$$s \approx \left|\frac{\omega}{\omega_1^{\max}}\frac{(1 + l^2\lambda_1^2)}{2l\lambda_1}R_C\right| \tag{2.8}$$

from the point where the topographic slope is due southward. Here

$$\omega_1^{\max} = \beta_{1,y}\lambda_1/2 \tag{2.9}$$

is the maximum Rossby wave frequency that can be obtained in the upper layer without coupling with the lower layer. The closest distance occurs when $l = 1/\lambda_1$, for which wavenumber

$$s = \left|\frac{\omega}{\omega_1^{\max}}R_C\right| \tag{2.10}$$

from the point where the topographic slope is due southward. This is the point where the mode conversion would occur, and at this point dk/dx is infinite, while $k = (k_1 + k_2)/2$ is finite. WKB theory is obviously degenerate at this point, although this transition is accommodated by the theory described by Flynn and Littlejohn (1994) and cited by Vanneste (2001).

Further, the physical separation between wavelike solutions in the x direction with the same value of l is

$$\Delta s = 2R_C\frac{\omega}{\omega_1^{\max}}\frac{1}{l\lambda_2}\sqrt{\frac{\beta_{1,y}}{\beta_{2,y}}}. \tag{2.11}$$

The integral of the imaginary part of the wavenumber over the separation between the wavelike solutions produces (not surprisingly) the same estimate for the transmission of wave activity as do the formulas of Flynn and Littlejohn (1994).

The transitions are much less abrupt over topography that slopes up poleward. The transition occurs where $k_2 \approx k_1$. At this point

$$\begin{aligned} k &= \left(k_1 \pm \frac{\omega}{\lambda_1\lambda_2}\sqrt{\frac{1}{\beta_{1,y}\beta_{2,y}}}\right) \\ &= \left[-\frac{1}{2\lambda_1}\frac{\omega}{\omega_1^{\max}}(1 + l^2\lambda_1^2) \pm \frac{1}{2\lambda_2}\frac{\omega}{\omega_1^{\max}}\sqrt{\frac{\beta_{1,y}}{\beta_{2,y}}}\right] \\ &= -\frac{1}{2\lambda_1}\frac{\omega}{\omega_1^{\max}}\left(1 + l^2\lambda_1^2 \mp \frac{\lambda_1}{\lambda_2}\sqrt{\frac{\beta_{1,y}}{\beta_{2,y}}}\right) \\ &\approx -\frac{1}{2\lambda_1}\frac{\omega}{\omega_1^{\max}}(1 + l^2\lambda_1^2) = k_1 \quad \text{and} \end{aligned} \tag{2.12}$$

$$\frac{dk}{dx} = \frac{1}{2}\frac{dk_2}{dx} = \frac{1}{2}\frac{l}{\beta_{2,y}}\frac{d\beta_{2,x}}{dx} = \frac{l}{2R_C}. \tag{2.13}$$

The criterion for the validity of WKB theory is then

$$\begin{aligned} k^{-2}\frac{\partial k}{\partial x} &= \left(\frac{1}{2\lambda_1}\frac{\omega}{\omega_1^{\max}}(1 + l^2\lambda_1^2)\right)^{-2}\frac{l}{2R_C} \\ &= \left(\frac{\omega_1^{\max}}{\omega}\right)^2\frac{2l\lambda_1}{(1 + l^2\lambda_1^2)^2}\frac{\lambda_1}{R_C} \ll 1. \end{aligned} \tag{2.14}$$

For sufficiently low frequencies, WKB theory does not apply, although the inapplicability of the theory is less decisive than in the case when topography slopes up toward the equator.

The strength of transmission between modes can be estimated for this physical model as it was in Vanneste (2001), and these estimates support the physical description in Hallberg (1997). Equation (2.3) can be cast in matrix form, as in Vanneste (2001), giving

$$\begin{aligned} \mathbf{D\Psi} &= \begin{bmatrix} \beta_{1,y}(k - k_1) & -\omega\lambda_1^{-2} \\ -\omega\lambda_2^{-2} & \beta_{2,y}(k - k_2) \end{bmatrix}\begin{bmatrix}\Psi_1 \\ \Psi_2\end{bmatrix} \\ &= \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22}\end{bmatrix}\begin{bmatrix}\Psi_1 \\ \Psi_2\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix}. \end{aligned} \tag{2.15}$$

The transmission coefficient from the expressions of Flynn and Littlejohn [1994, their Eq. (4.12)] can be expressed (again as in Vanneste) as $T = \exp(-2\pi\nu)$, where (in the special case that D_{12} and D_{21} are independent of wavenumber or position)

$$\nu = \left.\frac{D_{12}D_{21} - D_{11}D_{22}}{\sqrt{\left(\frac{\partial D_{11}}{\partial x}\frac{\partial D_{22}}{\partial k} - \frac{\partial D_{22}}{\partial x}\frac{\partial D_{11}}{\partial k}\right)^2}}\right|_{x=x_0,k=k_0}. \tag{2.16}$$

Equation (2.16) is evaluated at the wavenumber and position where the coupling would be expected to occur, namely at the point x_0 where $k_1 = k_2 \equiv k_0$. For either the coupling in the north or the south,

$$\nu = \frac{\omega^2\lambda_1^{-2}\lambda_2^{-2}}{|\beta_{1,y}\beta_{2,y}(\partial k_2/\partial x)|} = \frac{R_C}{4\lambda_2^2|l|}\left(\frac{\omega}{\omega_1^{\max}}\right)^2\left|\frac{\beta_{1,y}}{\beta_{2,y}}\right|. \tag{2.17}$$

Typical values for the various quantities in (2.17) might be $R_C = 1000$ km (for a basin roughly the size of the Argentine Basin; a much smaller radius of curvature is warranted in many locations, and this would give stronger coupling between modes), $\lambda_2 = 50$ km, $l = 1/\lambda_2$, and $\beta_{1,y}/\beta_{2,y} = 1/30$. With these values,

$$\nu = \frac{1}{6}(\omega/\omega_1^{\max})^2, \tag{2.18}$$

and the transmission coefficient becomes

$$T \approx \exp[-(\omega/\omega_1^{\max})^2], \tag{2.19}$$

It has already been assumed that frequency under consideration is smaller than $\omega_1^{\max}$, so the transmission coefficient may be quite close to 1. At frequencies not dramatically less than $\omega_1^{\max}$, both transmission between modes and mode conversion would be expected to be significant. For any values of the various constant coefficients, the coupling between modes is nearly complete for sufficiently low-frequency waves. All this is exactly as described in Hallberg (1997).

3. Clarification of the mass-conservation transmission estimate

Vanneste's (2001) claim that significant transmission should be considered exceptional simply does not apply to low-frequency waves in an ocean basin with arbitrary topography. For sufficiently low frequencies or for sufficiently abruptly varying topographic orientations, WKB theory is inapplicable to the coupling over zonally varying topography. The very special case of meridionally varying, meridionally sloping topography, upon which Vanneste is basing this claim, does not give insight into the more general case. Vanneste further suggests that the estimate of Hallberg (1997) for the strength of coupling is incorrect since it disagrees with his assertion that coupling should be weak. Here validity of the mass conservation argument of Hallberg (1997) is clarified.

Linear quasigeostrophic waves do transport mass. [This is not the same as causing a significant displacement of individual fluid parcels (see Hallberg and Rhines 1996).] If they did not, they could not excite any geostrophic flow. The ageostrophic convergence of mass is intimately connected with the excitation of geostrophic motion, and, in the limit that wavelengths are long, a perfectly sensible mass balance can be obtained at scales that are much shorter than that wavelength.

The velocities associated with low-frequency Rossby waves tend to be aligned with isopleths of potential vorticity (PV). Flows across PV isopleths create the restoring forces that lead to Rossby waves, and these across-PV-isopleth flows must be relatively weak in low-frequency waves. While it is true that there is no net mass flux when averaging over a wave period, along-contour gradients in the convergence of mass fluxes ''*even over just a fraction of a wave period*'' (Hallberg 1997, p. 986) will lead to cross-contour geostrophic flows. For oscillatory forcing with a given frequency, it is inconsistent to have convergence of along-PV-contour mass fluxes (and the resulting baroclinic Montgomery potential gradients) with spatial scales along the contour that are much shorter than the wavelength along the contour of topographic or planetary waves with that frequency. Unless there is coupling between modes, these small-scale Montgomery potential gradients will occur when

$$\left|\frac{1}{k\Psi_1}\frac{\partial}{\partial x}(\Psi_1)\right| \geq O(1). \tag{3.1}$$

It can be shown from conservation of wave activity, and starting from expressions in the previous section and in Hallberg (1997) (after much tedious algebra) that, in the case where $\lambda_1 = \lambda_2$ and at the coupling point (where $k_1 = k_2$ and $\Psi_1 = \pm\Psi_2$),

$$\begin{aligned}\left|\frac{1}{k\Psi_1}\frac{\partial}{\partial x}(\Psi_1)\right| &\approx \left|\frac{1}{2k}\left[-\frac{1}{c_g}\frac{dc_g}{dx} \mp \frac{d}{dx}\left(\frac{\Psi_2}{\Psi_1}\right)\right]\right| \\ &= \left|(l^2\lambda_1^2 + 1 \mp 1)\left(\frac{\beta_{2,y} + \beta_{1,y}}{4\beta_{1,y}}\right)\frac{1}{k^2}\frac{dk}{dx}\right|.\end{aligned} \tag{3.2}$$

[The two branches of the dispersion relation take on the two signs in (3.2)]. Since typically $|\beta_{2,y}| \gg |\beta_{1,y}|$, (3.2) can be approximated [also using (2.14)] as

$$\begin{aligned}\left|\frac{1}{k\Psi_1}\frac{\partial}{\partial x}(\Psi_1)\right| &\approx \left|\left(\frac{\omega_1^{\max}}{\omega}\right)^2\frac{l\lambda_1(l^2\lambda_1^2 + 1 \mp 1)}{2(1 + l^2\lambda_1^2)^2}\frac{\lambda_1}{R_C}\left(\frac{\beta_{2,y}}{\beta_{1,y}}\right)\right| \\ &\approx \left|\left(\frac{\omega_1^{\max}}{\omega}\right)^2\left(\frac{1}{4} \mp \frac{1}{8}\right)\frac{\lambda_1}{R_C}\left(\frac{\beta_{2,y}}{\beta_{1,y}}\right)\right|.\end{aligned} \tag{3.3}$$

The final expression applies when $l\lambda_1 = 1$, which is a typical meridional length scale. Equation (3.3) shows that this mass conservation argument becomes applicable essentially when (2.17) indicates that significant transmission should occur. From (3.3), relatively small-scale Montgomery potential gradients would occur without coupling between the modes for sufficiently low-frequency waves at the point where mode conversion would occur. For these waves, coupling between the modes prevents the convergence of mass in either isopycnal layer, and the strength of the coupling can be estimated from this constraint, as described in Hallberg (1997).

This argument applies to low-frequency motions and it predicts the strongest transmission for the lowest frequencies or for relatively short waves (in the cross-slope direction). Hallberg (1997) stated exactly this, and these are exactly the criterion for strong transmission from (2.17). In essence, Hallberg (1997) qualified this argument to apply in instances where WKB theory becomes inapplicable. This argument is not contradicted by the work in Vanneste (2001). Similar behavior is found from either the mass-flux conservation argument or from mode conversion theory. Hallberg (1997) has demonstrated the value of his theory for predicting the strength of transmission and mode conversion with primitive equation numerical simulations in idealized ocean basins in instances where WKB theory would be expected to fail. Further, the mass-flux conservation theory may be particularly useful as it permits a rough estimate of coupling strength in the real ocean based only on observations of velocity structures.

4. Conclusions

Nonmeridional variations in topography have quite a different effect on Rossby wave propagation from the very special case of meridionally varying meridional slopes. The criticism by Vanneste (2001) of the work in Hallberg (1997) is inaccurate, simply because that critique is based on exactly the wrong physical model. Order 1 transmission occurs between low-frequency wave modes, and is virtually inevitable as these waves propagate around the complex bathymetry of the real ocean. Sufficiently low-frequency waves strongly violate the assumptions behind WKB theory in the coupling regions, but the extensions to WKB theory cited by Vanneste (2001) support the strong coupling between modes described by Hallberg (1997), when applied to an appropriate physical model. Mode conversion (and associated coupling of surface- and bottom-intensified motion) is relatively weak for low-frequency motions, but is stronger at frequencies approaching the maximum baroclinic planetary wave frequency. That transmission between modes and the coupling between low-frequency surface- and bottom-intensified motion occur where the topographic slope becomes nearly meridional, but the variations of the orientation of the slope along a ray are essential in that process.

The theory discussed here and in Hallberg (1997) may ultimately prove useful for explaining a variety of oceanic observations, such as, perhaps, the observation from satellite altimetry that baroclinic Rossby waves are much stronger to the west of the Hawaiian Ridge and Emperor Seamount chain than they are to the east (Chelton and Schlax 1996). Similar behavior has been seen in numerical models (e.g., Gerdes and Wübber 1991), but explanations have typically been cast in terms of bottom torques or the joint effect of baroclinicity and relief (JEBAR: Mertz and Wright 1992). Explicit consideration of the properties of wave modes may provide much clearer explanations. But, progress in such an explanation is unlikely without consideration of nonmeridional topography and non-WKB effects.

REFERENCES

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

Flynn, W. G., and R. G. Littlejohn, 1994: Normal forms for linear mode conversion and Landau–Zener transitions in one dimension. *Ann. Phys.,* **234,** 334–403.

Gerdes, R., and C. Wübber, 1991: Seasonal variability of the North Atlantic Ocean—A model intercomparison. *J. Phys. Oceanogr.,* **21,** 1300–1322.

Hallberg, R., 1997: Localized coupling between surface- and bottom-intensified flow over topography. *J. Phys. Oceanogr.,* **27,** 977–998.

——, and P. B. Rhines, 1996: Buoyancy-driven circulation in an ocean basin with isopycnals intersecting the sloping boundary. *J. Phys. Oceanogr.,* **26,** 913–940.

Lighthill, M. J., 1978: *Waves in Fluids.* Cambridge University Press, 504 pp.

Mertz, G., and D. G. Wright, 1992: Interpretations of the JEBAR term. *J. Phys. Oceanogr.,* **22,** 301–305.

Vanneste, J., 2001: Mode conversion for Rossby waves over topography: Comments on ''Localized coupling between surface- and bottom-intensified flow over topography.'' *J. Phys. Oceanogr.,* **31,** 1922–1925.

Veronis, G., 1980: Dynamics of large-scale ocean circulation. *Evolution of Physical Oceanography.,* B. A. Warren and C. Wunsch, Eds., The MIT Press, 140–183.