What is the "Near-Inertial" Band and Why Is It Different from the Rest of the Internal Wave Spectrum?

CHRIS GARRETT

Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 12 December 1999, in final form 23 June 2000)

ABSTRACT

The "near-inertial" part of the internal wave continuum is dominant and also different from the rest of the spectrum. A simple possible reason for the difference is that waves generated at the surface are not reflected or scattered from the seafloor until they have propagated equatorward to a latitude where their frequency exceeds the local inertial frequency. This excess is easily estimated and is of the order of 10% of f at midlatitudes. The estimate is in reasonable agreement with data on the depth dependence of the peak frequency over smooth topography and on the frequency band within which there is little upward propagating energy. Internal wave propagation and interactions with bottom topography may thus be just as important as wave–wave interactions in controlling the energetic parts of the internal wave spectrum and, hence, in determining mixing rates in the ocean.

1. Introduction

Freely propagating internal waves are possible in the open ocean at all frequencies between the Coriolis frequency f and the buoyancy frequency N. Near f the fluid motion is close to horizontal, with the Coriolis force dominating, whereas near N the motion is close to vertical, with buoyancy providing the main restoring force. There is a gradual transition between these two limits, with no obvious transition frequency separating different parts of the continuum.

On the other hand, it does seem that the "near-inertial" band of waves with a frequency close to f does have different characteristics from the rest of the frequency spectrum. Near-inertial waves seem to have a more variable energy level than the rest of the spectrum (e.g., Fu 1981), a predominance of downward propagating energy (e.g., Leaman and Sanford 1975; D'Asaro and Perkins 1984; Pinkel 1984; Sanford 1991), and a redder vertical wavenumber spectrum than the rest of the frequency spectrum (Sanford 1991). Part of this different spectral slope was shown by Sherman and Pinkel (1991) to be a consequence of the vertical displacement, by the internal waves themselves, of a spectrum that is more nearly separable in a "semi-Lagrangian" frame in which the motion is referenced to isopycnals rather than to fixed depths. Even in this frame, however, the vertical wavenumber spectrum is steeper at near-inertial frequencies than at high frequencies (see their Fig. 8c).

These differences are important, especially because, in typical internal wave spectra, most of the energy and most of the shear (which influences mixing) are at frequencies close to f. A counterargument (E. Kunze 1999, personal communication) is that the near-inertial part of the spectrum is not as important as it seems, since the vertical energy *flux* involves multiplying the energy *density* by the vertical component of the group velocity, and this tends to zero near f. For example, if an energy spectrum with the same vertical wavenumber dependence at all frequencies is, as often assumed, proportional to $(\omega^2 - f^2)^{-1/2}\omega^{-1}$ as a function of frequency ω and is also vertically unidirectional, then the vertical energy flux is proportional to $(\omega^2 - f^2)^{1/2}\omega^{-2}$ for frequencies much less than the buoyancy frequency N, and so is spread over a broad frequency range. The near-inertial band is then of minor importance. Given the directional properties of the spectrum, however, with higher-frequency waves being fairly isotropic in the vertical, near-inertial waves may again emerge as the dominant contributors to the net vertical energy flux. The so-called GM spectrum (Garrett and Munk 1972; Munk 1981), which is vertically isotropic and has the same vertical wavenumber dependence at all frequencies, may not be a very good basis for internal wave studies.

It seems more appropriate to start focusing on the near-inertial band, treating this as different from the rest of the spectrum and thus introducing the questions asked in the title of this paper. The questions are, of course, connected, with many possible answers. Some, to be

Corresponding author address: Dr. Chris Garrett, Department of Physics and Astronomy, University of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada.

discussed in this paper, arise from consideration of the generation of inertial waves in the upper ocean, the types of nonlinear interaction that are possible, and the role of latitudinal turning points in wave propagation. The main suggestion to be made, though, is connected with the fact that near-inertial waves exist on a β plane and must therefore propagate equatorward. If they conserve their frequency, it becomes greater than the local f. The amount by which the wave frequency exceeds the local f when the waves reach the seafloor may then be a measure of what determines the near-inertial band within which the energy propagates downward and has not had its vertical wavenumber spectrum modified by reflection and scattering off a sloping and irregular seafloor.

There is, of course, extensive literature on near-inertial waves, including the comprehensive treatise of Fu (1981) and the other papers referred to above. Details of these and papers not yet cited will be introduced later as appropriate.

2. The near-inertial bandwidth

Let us suppose that the near-inertial band extends from f to $(1 + \epsilon)f$ and ask what value ϵ might take. An obvious answer might be that ϵ is about 1, on the grounds that at a frequency $2f$ the Coriolis and buoyancy forces are of equal importance in determining the wave frequency. This answer, however, appears to lack any dynamical significance or implication. We seek alternatives.

a. Generation

Much of the internal wave activity in the ocean appears to originate with the wind and other surface forcing. As discussed by Gill (1984), Zervakis and Levine (1995), and others, the speed of atmospheric fronts is typically much faster than the phase speed of internal waves in the ocean so that the oceanic response occurs in two stages. In the first, short, stage, currents are generated in the surface layer of the ocean. In the second, longer, stage, after the forcing has ceased, these currents undergo ‘‘Rossby adjustment.’’ In this process, waves of near-inertial frequency are generated, with the difference between their frequency ω and the local inertial frequency depending on the horizontal wavenumber (k, l) of the disturbance left behind by the storm, according to the formula

$$\omega^2 - f^2 = c_j^2(k^2 + l^2), \qquad (1)$$

where c_j is the eigenvalue associated with the jth vertical mode. If $\omega = (1 + \epsilon)f$ and ϵ is small, then for the jth mode

$$\epsilon = \epsilon_j \simeq \frac{1}{2}\left(\frac{c_j}{f}\right)^2(k^2 + l^2), \qquad (2)$$

where c_j/f is the associated Rossby radius. A typical midlatitude value might be 30 km for the first mode and less for higher modes, giving $\epsilon \simeq 0.05$ for mode 1 if $(k^2 + l^2)^{-1/2} \simeq 100$ km, decreasing with mode number.

Young and Jelloul (1997) have shown that (1) can effectively be changed by the interaction of near-inertial waves with small-scale quasigeostrophic motions, but we ignore this here, perhaps on the grounds that the eddies become weaker with increasing depth. Also, Kundu (1993) has shown that slower-moving storms will tend to shift the peak of the generated internal wave spectrum to higher frequencies. On the whole, however, it can be argued that the dominant inertial peak is a consequence of the forcing, in which case the question becomes one of accounting for the rest of the frequency spectrum.

b. The latitudinal turning point

Near-inertial waves in a medium of constant f are really a theoretical idealization; they cannot exist in the ocean because their wavelength exceeds the distance over which f varies significantly. In reality, waves propagating poleward in varying f will encounter a turning latitude where the meridional component of their group velocity tends to zero and they have a greatly enhanced energy, as discussed in detail by Fu (1981) and Munk (1980, 1981). Over rough topography and in the upper ocean, Fu (1981) found that the observed energy is enhanced more than predicted for free propagation, suggesting the importance of some sort of local process for transferring energy out of the near-inertial band before it has propagated far. In the deep ocean over smooth topography, however, he demonstrated that the near-inertial cusp in typical frequency spectra of horizontal kinetic energy may be purely a turning point effect, as if the waves had been generated at lower latitudes, at frequencies greater than the Coriolis frequency there, and had then freely propagated poleward.

The reverse situation is also possible, with the waves being generated near their turning latitude and propagating freely equatorward. This seems more likely if the waves being generated tend to have a frequency close to the local Coriolis frequency, though it cannot then account for internal waves with a frequency higher than the maximum value of f in the basin, which would be that at the highest latitude.

The width of the near-inertial band associated with enhancement at the latitudinal turning point can be estimated by considering the equatorward distance Y_{tp} from the turning point at which this distance is equal to the reciprocal of the meridional wavenumber. This location roughly marks the transition from an equatorward zone where ray theory is appropriate to a poleward zone where solutions need to be expressed in terms of Airy functions. We use the internal wave dispersion relation

$$\omega^2 - f^2 = \frac{(N^2 - f^2)(k^2 + l^2)}{k^2 + l^2 + m^2} \simeq \frac{N^2(k^2 + l^2)}{m^2} \qquad (3)$$

away from N and with $N^2 \gg f^2$. Here f is a function of latitude, with $\omega - f \simeq \beta Y_{tp}$ at a small distance Y_{tp} equatorward of the latitude where $\omega = f$. Then, with $k = 0$ and taking $l = Y_{tp}^{-1}$,

$$Y_{tp} = \left(\frac{N^2}{2\omega\beta m^2}\right)^{1/3} = \left(\frac{N_0^2 b^2}{2\omega\beta j^2\pi^2}\right)^{1/3} = \left(\frac{N_0^2 b^2 R}{4\Omega^2 j^2\pi^2 \sin 2\phi}\right)^{1/3} \quad (4)$$

for $f = 2\Omega \sin\phi$ and $\beta = 2\Omega R^{-1} \cos\phi$ at latitude ϕ with $\Omega = 7.292 \times 10^{-5}$ s^{-1} the angular rotation rate of the earth and $R = 6371$ km its radius. We have also used the connection $m = j\pi N/(N_0 b)$ between the vertical wavenumber m and the mode number j for an ocean with $N = N_0 e^{z/b}$, where $z = 0$ at the surface, and we assume that the ocean depth H is much greater than the scale depth b (Munk 1981). The associated value of ϵ is ϵ_{tp} from $\omega = f(0) = (1 + \epsilon_{tp})f(-Y_{tp})$, where f is expressed as a function of the poleward distance from the turning point. Hence, ignoring a term of order ϵ_{tp}^2,

$$\epsilon_{tp} \simeq \left(\frac{\beta}{f(0)}\right)Y_{tp} = \frac{1}{2} j^{-2/3}\pi^{-2/3}\left(\frac{N_0 b \cos\phi}{\Omega R \sin^2\phi}\right)^{2/3}. \quad (5)$$

Taking $N_0 = 5.24 \times 10^{-3}$ s^{-1} and $b = 1.3$ km,

$$Y_{tp} = 112 j^{-2/3}(\sin 2\phi)^{-1/3} \text{ km},$$
$$\epsilon_{tp} = 0.014 j^{-2/3}(\cos\phi/\sin^2\phi)^{2/3}, \quad (6)$$

which are $120j^{-2/3}$ km and $0.032j^{-2/3}$, respectively, at latitude 30°. (Within the approximations used here, ϕ can refer to the latitude at the turning point or at a distance Y_{tp} equatorward.)

These rather small values for Y_{tp} and ϵ_{tp} suggest that ray theory is likely to be appropriate to within a fairly small distance of the turning latitude. The main point to be made, though, is that there seems to be no reason why the dynamics of the waves should change character as this boundary is crossed. We thus seek other definitions of the near-inertial bandwidth.

c. *Nonlinear interactions*

A possible dynamical reason for treating near-inertial waves separately from the rest of the spectrum is that only waves with frequency greater than $2f$ are susceptible to parametric subharmonic instability (PSI) (McComas and Müller 1981), which transfers energy to waves with half the frequency of the original waves. The frequency band from f to $2f$ is thus different from the rest of the spectrum in that it can receive energy by PSI but cannot lose it. It is not clear, however, that this transfer could account for the observed predominance of downward propagating near-inertial energy, or for the redder vertical wavenumber spectrum. On the contrary, one might expect that PSI would generate high vertical wavenumber near-inertial energy, thus giving rise to a less red, rather than redder, vertical wavenumber spectrum.

McComas and Müller (1981) found that induced diffusion also tends to transfer energy to lower frequency and higher wavenumber, but the rate at which this happens for a wavenumber–frequency spectrum that has different vertical wavenumber dependencies at near-inertial and higher frequencies requires further attention. Moreover, all the calculations have been carried out for an f plane; it is possible that the β effect ''detunes'' some of the interactions for near-inertial waves. Overall, the behavior of near-inertial waves in weakly nonlinear wave–wave interaction theory seems uncertain.

Noting that weak interaction theory breaks down for much of the internal wave field, Henyey et al. (1986) adopted an alternative approach in which high vertical wavenumber waves propagate through a much larger-scale background shear representing the rest of the internal wave field. They found a general tendency for refraction to higher vertical wavenumber, and higher frequency, leading ultimately to dissipation. The theory has been extended by Polzin et al. (1995) to allow for different spectral shapes and by Sun and Kunze (1999a, b) to include the effects of vertical divergence. In all these studies the background shear is clearly provided to a large extent by the near-inertial waves, but the back effect of the interactions on them is unclear.

It seems that there are good grounds for revisiting the two different approaches to nonlinear internal wave interactions in the ocean, paying more attention to the energetically dominant near-inertial band, but this is not the immediate object of this short paper. The conclusion here is that the existing theories of nonlinear interactions present no compelling reason for considering a narrow frequency band near f to be special.

3. Meridional propagation of near-inertial waves

Several authors (e.g., Kroll 1975; Gill 1984; Zervakis and Levine 1995) have used ray theory to trace the paths of near-inertial wave packets, but their emphasis has been largely on the escape of near-inertial energy from the near-surface region of the ocean. The point to be made here is illustrated in Fig. 1: inertial energy generated at the sea surface at the turning latitude propagates equatorward and vertically downward until it reaches the seafloor, where it is reflected and scattered. The waves then have the same frequency as at the generation site, but this is now larger than the local Coriolis frequency, by an amount ϵf, say, with ϵ a small number that varies with latitude. If seafloor reflection and scattering are only in the equatorward direction (as sketched in Fig. 1), then near the seafloor the waves in the frequency band from the local f to $(1 + \epsilon)f$ are only propagating downward and, furthermore, have a vertical wavenumber spectrum that has not been modified by bottom reflection and scattering. The waves in this band

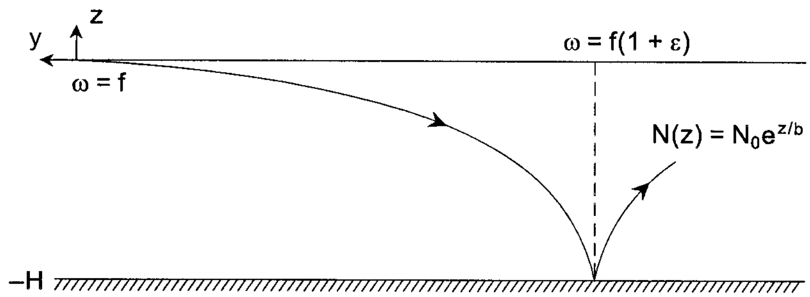

FIG. 1. Schematic of the ray for inertial waves generated at the surface and propagating equatorward and downward. The frequency ω is fixed, but the Coriolis frequency is a function of meridional distance y.

are thus distinct from the rest of the internal wave continuum.

If backscattering occurs, as is possible if the bottom slope is steep enough, this can give rise to upward propagating energy in the defined frequency band, but perhaps with generally shorter wavelengths for the upward propagating waves than for the downward propagating ones. We return to this issue later, but now proceed to evaluate ϵ according to the definition above.

It is simple to use the ray equations to determine ϵ. We start from the dispersion relation

$$\omega = \left(f^2 + \frac{N^2(k^2 + l^2)}{m^2}\right)^{1/2} = F(k, l, m;\ f(y), N(z)), \quad (7)$$

and have

$$\frac{dy}{dt} = \frac{\partial F}{\partial l} = \frac{N^2 l}{m^2\omega}, \qquad \frac{dz}{dt} = \frac{\partial F}{\partial m} = -\frac{N^2(k^2 + l^2)}{m^3\omega} \quad (8)$$

for the velocity of a wave group, taking y positive poleward and z positive upward, with the origin at the location where $\omega = f$.

The wavenumber k in the zonal direction does not change, whereas the meridional wavenumber l and vertical wavenumber m satisfy

$$\frac{dl}{dt} = -\frac{\partial F}{\partial y}, \qquad \frac{dm}{dt} = -\frac{\partial F}{\partial z}, \quad (9)$$

where the differentiation is with respect to the explicit dependence of F on y and z through $f(y)$ and $N(z)$.

Evaluating (9) gives

$$\frac{dl}{dt} = -\frac{\beta f}{\omega}. \quad (10)$$

This is approximately $-\beta$ if we are assuming $f \simeq \omega$, so that

$$l \simeq l_0 - \beta t \quad (11)$$

for $l = l_0$ at $t = 0$. It is interesting that this decrease of l with time, because of the β effect, emerges from the ray equations as well as being implicit in the momentum equations, for which D'Asaro (1989) pointed out that current vectors are rotated at different rates at different latitudes, generating a finite meridional wavenumber if $l_0 = 0$. The ray equations make it clear that this change of l with time is associated with the simultaneous propagation of a wave packet.

In the vertical

$$\frac{dm}{dt} = -\frac{N(k^2+l^2)}{m^2\omega}\frac{dN}{dz} = \frac{m}{N}\frac{dN}{dz}\frac{dz}{dt}, \quad (12)$$

giving the standard WKB result that $m \propto N$. In the model ocean chosen, with $N = N_0 e^{z/b}$ and $b \ll H$, we may write

$$m = \frac{j\pi N}{N_0 b}, \quad (13)$$

where j is the mode number for modes on an f plane. Here, though, we use j only as a label for a vertically propagating wave on a β plane, without implying waves of equal amplitude propagating up as well as down.

These simple ray equations are not new [see, e.g., Kroll (1975) or Zervakis and Levine (1995)] but seem not to have been fully exploited. It should be emphasized that only three of the four ray equations for y, z, l, and m are independent, given the dispersion relation (7) connecting y, z, l, and m with the constant k and constant ω. There are different pathways to the solution using different combinations of the ray equations and the dispersion relation. We shall concentrate first on the geometry of the ray, leaving the time dependence until later. We proceed by using (7) and (8) to write

$$\frac{dy}{dz} = \mp N(\omega^2 - f^2)^{-1}\left[\omega^2 - f^2 - \left(\frac{N_0 bk}{j\pi}\right)^2\right]^{1/2}, \quad (14)$$

where the sign is negative for $l > 0$ and positive for $l < 0$. Since $\omega^2 - f^2$ is a function of the meridional distance y, the right-hand side of (14) is a function only of y and z and may be integrated to give the ray.

On a β plane $f = \omega + \beta y$, taking the origin for y at the latitude where $f = \omega$, so we may write $\omega^2 - f^2 \simeq -2\omega\beta y$ and (14) becomes

$$\frac{dy'}{dz} = \pm\frac{N}{2\omega\beta y'}\left[2\omega\beta y' - \left(\frac{N_0 bk}{j\pi}\right)^2\right]^{1/2}, \quad (15)$$

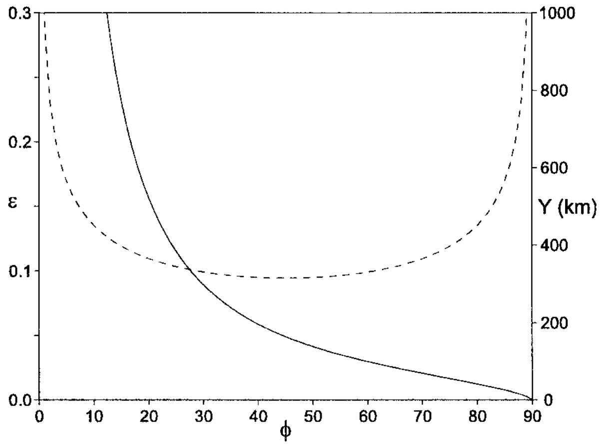

FIG. 2. The equatorward propagation distance Y (dashed curve, right-hand axis) and fractional frequency difference ϵ (solid line, left-hand axis), from (19) as functions of the latitude ϕ.

where we have defined $y' = -y$ for convenience, and the sign is now positive for $l > 0$ and negative for $l < 0$.

We first investigate the simplest case in which $k = l_0 = 0$, so that waves initially have a frequency equal to the local f and originate at the turning latitude where $y' = 0$. In this case (15) becomes

$$\frac{dy'}{dz} \simeq -\frac{N}{(2\omega\beta y')^{1/2}}. \quad (16)$$

For the model ocean already considered, with $N = N_0 e^{z/b}$, the ray reaches the seafloor at $z = -H$ at $y' = Y$ where, for $H \gg b$,

$$Y \simeq \frac{1}{2}\left(\frac{9N_0^2 b^2}{\omega\beta}\right)^{1/3} = \frac{1}{2}\left(\frac{9N_0^2 b^2 R}{2\Omega^2 \sin 2\phi}\right)^{1/3}, \quad (17)$$

which is similar in form to (4). The corresponding value of ϵ is

$$\epsilon \simeq \frac{1}{2}\left(\frac{3N_0 b \cos\phi}{2\Omega R \sin^2\phi}\right)^{2/3}, \quad (18)$$

which is likewise similar to (5). Taking $N_0 = 5.24 \times 10^{-3}\ \mathrm{s}^{-1}$ and $b = 1.3$ km as before, we may write

$$Y = 315(\sin 2\phi)^{-1/3}\ \mathrm{km},$$
$$\epsilon = 0.039(\cos\phi/\sin^2\phi)^{2/3}, \quad (19)$$

as shown in Fig. 2, and giving values of 330 km and 0.09, respectively, at latitude 30°. (As in the earlier discussion of turning points, ϕ can be taken here to be either the generation latitude or the latitude where the waves reach the seafloor.)

The formulas are similar to (6), but with larger values of both ϵ and Y, especially for high modes, thus illustrating the validity of ray theory over most of the ray path, at least as far as the meridional wavenumber l is concerned. Using ray theory is more questionable for low vertical mode numbers for which the vertical wavelength is not small compared with the buoyancy scale height b. We do note, however, that within the limitations of ray theory the ray path is independent of the equivalent vertical mode number for the waves originating at the turning latitude. These formulas use the approximation $H \gg b$; for the assumed stratification both Y and ϵ are reduced at a depth of 4 km by a factor $(1 - e^{-4/1.3})^{2/3} = 0.97$.

It should be remarked that, while ϵ from (18) gives the frequency bandwidth for downward-propagating waves observed near the bottom, upward-propagating waves will not be seen higher in the water column until the wave ray has progressed farther equatorward (assuming reflection from a flat bottom or only forward scattering). The reflected ray is described by (16), with the sign reversed since the vertical wavenumber is reversed on reflection at a flat bottom. Simple integration shows that the ray reaches the sea surface at a distance $2^{2/3}Y = 1.6Y$ from the generation site, with Y given by (17) or (19) and corrected by a factor 0.97 for finite depth. The next encounter with the bottom is at $3^{2/3}Y = 2.1Y$, and so on; a wave group originating at its turning latitude clearly has many bottom encounters before it reaches the latitude where its frequency is greater than twice the local f and can decay by parametric subharmonic instability (Hibiya et al. 1999; Nagasawa et al. 2000).

a. The pole and equator

In evaluating Y and ϵ we have made approximations that are not valid near the pole or equator. We may examine these regions separately, again for $k = l_0 = 0$, but using (14) rather than the approximation (15). In these cases we write $\omega = 2\Omega \sin\phi_0$, $f = 2\Omega \sin\phi$, and $dy = R d\phi$, leading to

$$\int_{\phi_b}^{\phi_0} (\sin^2\phi_0 - \sin^2\phi)^{1/2}\, d\phi = \left(\frac{N_0 b}{2\Omega R}\right)(1 - e^{-H/b}), \quad (20)$$

where ϕ_0 is the latitude at which inertial waves are assumed to be generated at the surface, and ϕ_b is the latitude at which the wave ray reaches the seafloor.

For a wave packet originating at the North Pole with $k = l = 0$, we have $\phi_0 = \pi/2$, and the left-hand side of (20) integrates to $1 - \sin\phi_b$. For our choice of parameters the right-hand side of (20) is 0.0071, giving $\phi_b = 83.2°$. The wave packet then reaches the seafloor 6.8° away from the pole, at a distance of 760 km. The corresponding value of ϵ is cosec $\phi_b - 1 = 0.007$; near-inertial waves interact with bottom topography before they have acquired a frequency much greater than the local f.

Near the equator, on the other hand, there is a latitude from which waves generated at the local inertial frequency reach the seafloor right at the equator. Because this latitude is small, the left-hand side of (20) may be approximated by

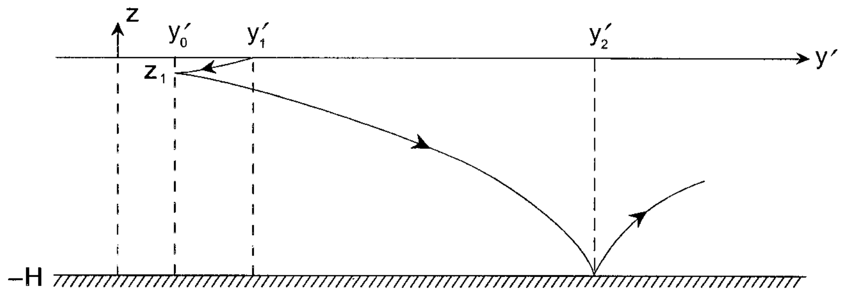

FIG. 3. Schematic of the ray for near-inertial waves generated at the surface with finite initial horizontal wavenumbers.

$$\int_{\phi_b}^{\phi_0} (\phi_0^2 - \phi^2)^{1/2}\, d\phi = \left(\frac{\pi}{4}\right)\phi_0^2,$$

giving $\phi_0 = 5.4°$. Waves originating closer to the equator cross it but still eventually reach the seafloor, possibly after internal reflections at the turning latitude in the other hemisphere and also at the original latitude (e.g., Eriksen 1993, 1999). Within $\pm 5.4°$ of the equator there is really no simple definition of a near-inertial frequency band along the lines suggested in this paper.

b. Effects of finite initial wavenumber

The above calculation has assumed that the initial wavenumbers are both zero. The zonal wavenumber k then remains zero, while the meridional wavenumber l becomes increasingly negative as the wave group propagates equatorward. It is straightforward to allow for finite k and for finite initial l.

A ray path for the general situation is shown in Fig. 3. Waves of frequency ω are generated a distance y_1' equatorward of the latitude where $f = \omega$. Waves with a negative initial meridional wavenumber l_0 propagate equatorward and reach the seafloor poleward of the location calculated with $k = l_0 = 0$. For a given value of $|l_0|$, however, greater changes are achieved by choosing $l_0 > 0$, in which case the wave packet initially propagates poleward, reaching a depth z_1 at a turning latitude where $l = 0$ and $y' = y_0'$. If $k \neq 0$, then $y_0' > 0$. After internal reflection the wave packet propagates equatorward and downward, reaching the seafloor at y_2'. We seek the correction to the value of Y given by (17).

The governing equation is (15), with $+$ for the first part of the ray and $-$ for the second. It integrates to

$$\left[\frac{2}{3}(y' - y_0')^{3/2} + 2y_0'(y' - y_0')^{1/2}\right] = \pm(2\omega\beta)^{-1/2}N_0 b[e^{z/b}], \tag{21}$$

where brackets denote the change between end points of a ray section and $2\omega\beta y_0' = (N_0 bk/j\pi)^2$. The ray originates where $2\omega\beta y_1' = (N_0 b/j\pi)^2(k^2 + l_0^2)$. For the first part of the ray, from $(y_1', 0)$ to (y_0', z_1), using these formulas and (17) for Y, we obtain

$$1 - e^{z_1/b} = (8/9)(Y/j\pi)^3(l_0^3/3 + l_0k^2). \tag{22}$$

For the second part of the ray, from (y_0', z_1) to $(y_2', -H)$, assuming $y_2' \gg y_0'$, $H \gg b$, and using $y_2' \simeq Y$ in the expansion,

$$y_2' \simeq Y(e^{z_1/b})^{2/3}\left[1 - \frac{4}{9}\left(\frac{kY}{j\pi}\right)^2\right]. \tag{23}$$

Combining (22) and (23) leads to

$$\frac{y_2'}{Y} \simeq 1 - \frac{4}{9}\left(\frac{kY}{j\pi}\right)^2 - \frac{16}{27}\left(\frac{kY}{j\pi}\right)^2\left(\frac{l_0Y}{j\pi}\right) - \frac{16}{81}\left(\frac{l_0Y}{j\pi}\right)^3. \tag{24}$$

For k and l_0 of comparable magnitude, the first correction in (24) is dominant, though for a zonal track of the storm generating the waves we might expect $k \ll l_0$ so that the last term of (24) dominates the reduction in the meridional distance traveled before the waves reach the seafloor. If we take k^{-1} or $l_0^{-1} \simeq 100$ km, then the factors $kY/j\pi$ and $l_0Y/j\pi$ are approximately j^{-1} at latitude 30°. For low mode numbers this is not really small enough for the approximations used, but does suggest a significant reduction in the estimate of ϵ for the definition of the near-inertial band. The correction becomes much less for smaller k and l_0 and for high modes. We also note that with a finite initial meridional wavenumber l_0, but with $k = 0$, the waves reach their turning latitude at $y' = 0$ at a depth $z_1 = -(N_0^2/3\omega\beta)(l_0b/j\pi)^3$, which is $0.45j^{-3}$ km for $l_0^{-1} = 100$ km at latitude 30°.

The overall conclusion is that the bandwidth measure ϵ given in (18), which was independent of the vertical wavenumber of the waves, will be reduced by an amount that might be some tens of percent for waves with an equivalent mode number of 1 but much less for higher modes.

c. Evidence?

Fu (1981) examined the frequency of the peak of the internal wave spectrum using data from a variety of current meter moorings in the North Atlantic, finding a considerable scatter but with a definite tendency for the peak frequency to increase with depth. Fu attributed this to the fact that only waves with superinertial frequency can propagate downward. It seems quite likely, however, that the ''blue shift'' is a consequence of the β plane, as discussed above. At a depth of 4 km the average peak frequency for several datasets from moorings over smooth topography is about 1.05 times the local inertial frequency (see Fu's Fig. 5, upper panel). The moorings were at latitudes between 28° and 38°N; at 33°N the value of ϵ from (19), with the correction factor 0.97, is 0.076. Comparing this with 0.05 assumes that the model stratification is reasonably appropriate, as seems to be the case. Quite possibly the reduced average blue shift is a consequence of waves being generated with non-negligible zonal and meridional wavenumbers, as discussed in the previous section.

Over rough topography, Fu (1981) found much more scatter in the peak frequency, and less of a tendency for a blue shift with depth. It seems possible that the near-inertial waves are partially back-reflected or backscattered from the topography, rather than being reflected equatorward, and we may ask whether this could give rise to deep energy closer to the local inertial frequency. For a start, it is worth remarking that the bottom slope that would give back-reflection is $\tan^{-1}(2\epsilon)^{1/2}(f/N)$. At latitude 33°N again, and with a typical bottom value of $5 \times 10^{-4}\ \mathrm{s}^{-1}$ for N, this critical slope is 0.06, smaller than the root mean square value of 0.2 calculated by Müller and Xu (1992) for Bell's (1975) model spectrum of seafloor topography truncated at a horizontal wavelength of 400 m. The net effect of rough topography on forward- or backscattering is complicated, but Longuet-Higgins (1969) found that typically 50% of the energy is backscattered from a modulated sinusoid with a steepest slope greater than that of the wave ray, and this is probably a good guide even though his approach, based on characteristics, fails to satisfy the radiation condition exactly (Baines 1971). Simple back-reflection or backscattering along the same ray would certainly give upward- as well as downward-propagating waves near the inertial frequency, but the expected peak frequency would remain as for smooth topography and forward reflection at the seafloor. Getting a frequency peak close to the local f might call for a more complicated scenario, for example, if the waves are reflected poleward and downward from steep topography and then continue poleward but upward after reflection from a flatter region of the seafloor. This may be inadequate, though, to explain situations where the peak is very close to the local inertial frequency. Inertial peaks in the deep ocean over rough topography could be associated with inertial waves generated as part of the geostrophic adjustment of mixed regions, and it is interesting that Fu (1981) found that the height of the near-inertial peak was negatively correlated with the departure of the peak frequency from the local f, as might happen if more vigorous mixing events give rise to energetic motions very close the the local inertial frequency. This calls for investigation beyond the scope of the present paper.

The most compelling evidence for the main suggestion of this paper comes from the analysis by D'Asaro and Perkins (1984) of data obtained in the Sargasso Sea from a current meter mooring at 30°29.7′N, 71°45.6′W and from surrounding vertical profiles of horizontal currents. The region is over the smooth Hatteras abyssal plain, so there should be no backscattered upward-propagating near-inertial waves. We should expect to see only downward-propagating energy in the abyss between f and about $1.085f$; near the surface where the data were obtained a further factor 1.6 expands the bandwidth to about $1.14f$. Any wave with a frequency between the local f and 1.14 times it must, if generated at the inertial frequency farther poleward, have originated at a smaller distance poleward than a wave with frequency 1.14 times the local f; it could not have reached the surface again even if it had been reflected. The factor 1.14 will, as discussed earlier, be reduced for low modes by the effect of a finite initial wavenumber.

Figure 4 here reproduces panels from Figs. 7 and 8 of D'Asaro and Perkins (1984). The data show a conspicuous lack of upward-propagating energy within a band from f to about $1.1f$. If anything, the bandwidth of this gap is a little wider for the first two modes than for modes 3–5, contrary to the expectations here of the effect of finite initial bandwidth, but this is a detail at this stage given that the authors estimate the frequency using an average of the fairly wide range of horizontal wavenumbers determined from the data.

d. Travel time versus interaction time

It begins to seem that much of the behavior of the near-inertial band is determined by propagation and bottom reflection rather than by wave–wave interactions, even though the latter are sometimes thought to be important for near-inertial waves that have a very small vertical group velocity. In fact, waves propagating equatorward ''escape'' the inertial frequency on a spherical earth and so pick up a greater vertical speed with time. For our simple example of waves starting with $k = l = 0$, the travel time to the seafloor may easily be calculated, since (8), (11), and (13) give

$$\frac{dy'}{dt} = \left(\frac{N_0 b}{j\pi}\right)^2 \frac{\beta t}{\omega}. \tag{25}$$

The travel time T to the latitude of bottom reflection, at an equatorward distance given by (17), is

$$T = 3^{1/3} j\pi \left(\frac{R^2 \sin\phi}{2\Omega N_0^2 b^2 \cos^2\phi}\right)^{1/3}. \tag{26}$$

For our choice of parameters this is $9.5j(\sin\phi/\cos^2\phi)^{2/3}$ days (Fig. 5), or $8.3j$ days at latitude 30°. For finite initial wavenumbers it will be reduced slightly by the reduction in meridional distance traveled from the turning latitude to the bottom, though also increased a little

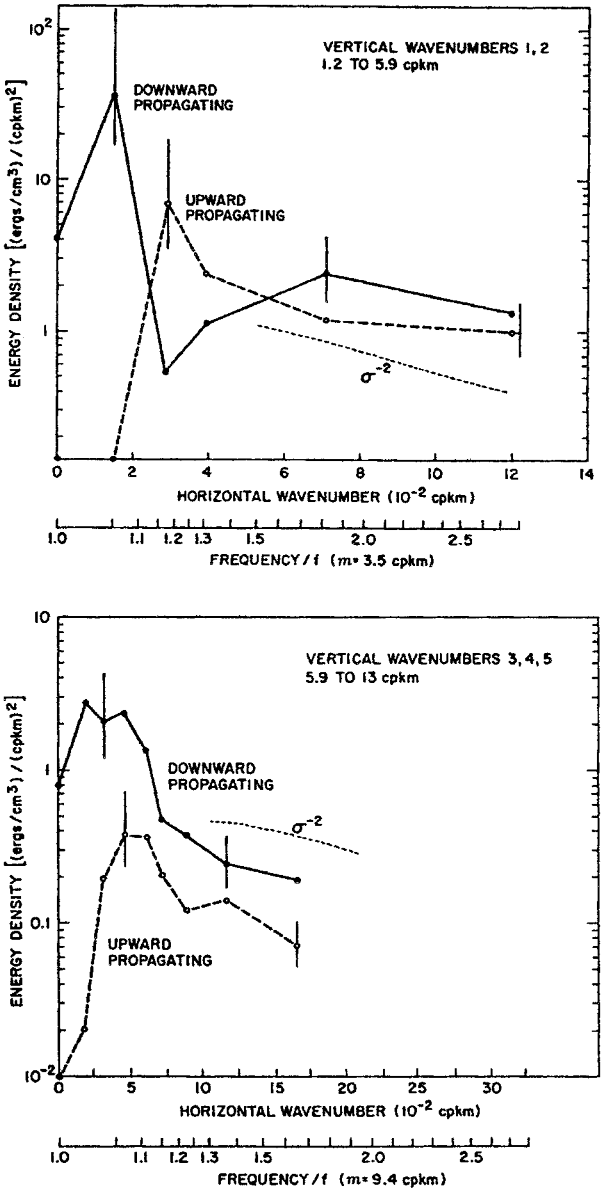

FIG. 4. Energy spectra for upward and downward propagating internal waves as a function of horizontal wavenumber. The abscissa has also been presented as frequency, using an average of the vertical wavenumber over the range of vertical modes considered. From D'Asaro and Perkins (1984).

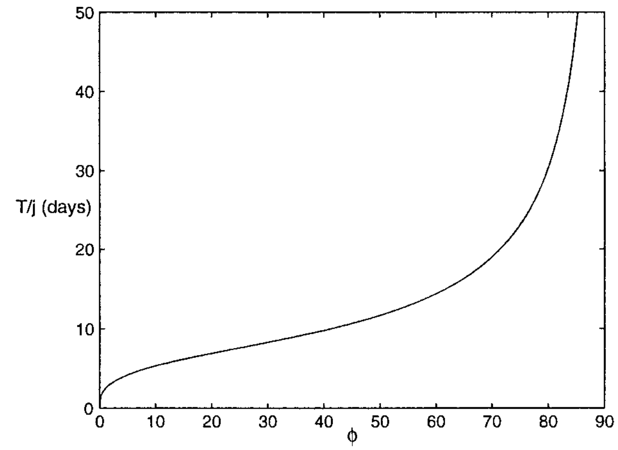

FIG. 5. The time T in days, divided by equivalent vertical mode number j, taken for waves generated at the inertial frequency at latitude ϕ to reach the seafloor.

by the time taken to propagate poleward to the turning latitude.

This time does not seem particularly long for waves with a small equivalent mode number j, but should be compared with the time for significant changes caused by nonlinear wave–wave interactions. As mentioned earlier, the details of these interactions need reexamination for a spectrum in which the vertical wavenumber behavior is different at near-inertial and higher frequencies. The main mechanism near f seems to be induced diffusion in which waves with a large vertical wavenumber are refracted by waves with a small vertical wavenumber; the effect on the lowest modes themselves is unclear. Nonetheless, we use the interaction times for a typical spectrum, as calculated by McComas and Müller (1981), discussed by Müller et al. (1986), and summarized by Hirst (1991). The time is a function of the vertical wavenumber, which changes during the propagation of a wave group because N varies with depth, but it seems that the interaction time is much longer than the propagation time for waves with a low mode number, becoming less than the propagation time at latitude 30° for mode number 5 and higher. The low-mode near-inertial waves, which contain most of the energy in the internal wave spectrum, are thus affected more by bottom interactions than by wave–wave interactions. Higher modes may be affected by wave–wave interactions and by interaction with mesoscale eddies.

This discusssion breaks down near the pole. At midlatitudes an important aspect of the waves is that they have a finite acceleration, even if their initial group speed is zero. This is not true at the pole, where dl/dt from (9) may be written as $-2\Omega y'/R^2$. Combined with $dy'/dt = (N_0 b/j\pi)^2(l/\omega)$ from (8), and with $\omega = 2\Omega$, this gives

$$\frac{d^2y'}{dt^2} = \left(\frac{N_0 b}{j\pi R}\right)^2 y'. \tag{27}$$

Ray theory is not, of course, valid for long waves, but (27) does show that a finite wavenumber is required right at the generation site, and the distance from the pole can then increase exponentially, though with the very long timescale of $j\pi R/(N_0 b)$, which is $34j$ days for our choice of parameters.

4. Discussion

The main theme of this short paper is the need to pay more attention to the near-inertial part of the internal wave spectrum, as this is what seems to be primarily generated and to contain most of the energy. Moreover, near-inertial waves seem to have a different vertical wavenumber content from the rest of the spectrum and have a predominantly downward energy flux (Sanford 1991; Sherman and Pinkel 1991; D'Asaro and Perkins 1984). In discussing near-inertial waves, the key factor to recognize is that they propagate on a β plane, not an f plane, as is, of course, well recognized in the literature (e.g., Gill 1984). The suggestion here is that the main factor, which defines the ''near-inertial'' band and accounts for its different characteristics, is that by the time the waves reach the seafloor their frequency is greater than the local f. The excess varies with latitude, and with the initial horizontal wavenumbers, but is of the order of 10% of f.

The apparent success of this idea in accounting for some observed spectral features, as a consequence of propagation rather than nonlinear interaction, is encouraging and suggests the need for a reappraisal of internal wave theories with an increased emphasis on the near-inertial band. Obvious questions include:

1) Can the steep vertical wavenumber spectrum of near-inertial waves reported by Sanford (1991) be attributed largely to generation and propagation? If not, at what vertical wavenumber do nonlinear interactions start to play a dominant role?
2) Is the whiter vertical wavenumber spectrum for higher-frequency waves reported by Sanford (1991) and Sherman and Pinkel (1991), even in a semi-Lagrangian frame, a consequence of reflection and scattering off a rough seafloor of waves that have a frequency that is ever greater than the local f as they propagate southward? Reexamination of the scattering results of Müller and Xu (1992) in the near-inertial limit would be interesting, though their model will need to be extended, since their assumption that the bottom slope is much less than the wave slope is not valid near the inertial frequency. In fact, a single incident wavenumber reflected from curved topography with a critical slope at one point will produce a vertical wavenumber spectrum proportional to m^{-1} at high wavenumber (Gilbert and Garrett 1989). The nonconvergence of an integral of this means that some energy will be lost, but the general whitening of the spectrum is interesting.
3) What can we say about the general behavior in wave–wave interactions of the near-inertial band, given that it is propagating on a β plane? Much of the previous theoretical work seems to take the inertial waves as a given background on an f plane, rather than the primary source on a β plane.

In a sense, the view expressed here is just reinforcing the suggestion of D'Asaro (1991) that in future internal wave research an emphasis be placed on the large-scale ''propagating'' modes, though he was concerned that much of the near-inertial energy would be subject to refraction by weak ocean currents, as discussed by Kunze (1985), D'Asaro (1995), and Young and Jelloul (1997). This is a major issue that has been ignored in the present paper.

We are just beginning to address the task of combining generation, propagation, nonlinear interactions, and topographic interactions of internal waves. A preliminary concentration on the dominant near-inertial band may help to focus this discussion and complement recent interest in the role of internal tides in ocean mixing (e.g., Munk and Wunsch 1998; Ledwell et al. 2000). The discovery by Mihaly et al. (1998) of apparent interactions between the M_2 tide and inertial waves suggests that an important way in which the internal tide may enhance ocean mixing is by acting as a catalyst for the dissipation of wind-generated internal waves. This need not show up in the energy budget for the tides themselves.

Acknowledgments. I thank Eric Kunze, Bill Young, and Rob Pinkel for useful comments and references, and two anonymous reviewers for further valuable comments and questions. A first draft of this paper was presented at the 1999 'Aha Huliko'a (Müller and Henderson 1999). Further work and the preparation of this paper were carried out during a visit to Scripps Institution of Oceanography, with the support of the Green Foundation, and facilitated by Walter Munk. This hospitality and the support of the Office of Naval Research are acknowledged with gratitude.

REFERENCES

Baines, P. G., 1971: The reflexion of inertial/internal waves from bumpy surfaces. *J. Fluid Mech.,* **46,** 273–291.

Bell, T. H., 1975: Statistical features of sea-floor topography. *Deep-Sea Res.,* **22,** 883–892.

D'Asaro, E. A., 1989: The decay of wind-forced mixed layer inertial oscillations due to the beta effect. *J. Geophys. Res.,* **94,** 2045–2056.

——, 1991: A strategy for investigating and modeling internal wave sources and sinks. *Dynamics of Oceanic Internal Gravity Waves, II: Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop,* P. Müller and D. Henderson, Eds., SOEST, 451–465.

——, 1995: Upper-ocean inertial currents forced by a strong storm. Part III: Interaction of inertial currents and mesoscale eddies. *J. Phys. Oceanogr.,* **25,** 2953–2958.

——, and H. Perkins, 1984: A near-inertial internal wave spectrum for the Sargasso Sea in late summer. *J. Phys. Oceanogr.,* **14,** 489–505.

Eriksen, C. C., 1993: Equatorial ocean response to rapidly translating wind bursts. *J. Phys. Oceanogr.,* **23,** 1208–1230.

——, 1999: Observations of low-latitude, near-inertial gravity waves forced by westerly wind bursts. *Dynamics of Oceanic Internal Gravity Waves, II: Proceedings of the 'Aha Huliko' a Hawaiian Winter Workshop,* P. Müller and D. Henderson, Eds., SOEST, 173–179.

Fu, L.-L., 1981: Observations and models of inertial waves in the deep ocean. *Rev. Geophys. Space Phys.,* **19,** 141–170.

Garrett, C. J. R., and W. H. Munk, 1972: Space–time scales of internal waves. *Geophys. Fluid Dyn.,* **2,** 225–264.

Gilbert, D., and C. Garrett, 1989: Implications for ocean mixing of

internal wave scattering off irregular topography. *J. Phys. Oceanogr.*, **19,** 1716–1729.
Gill, A. E., 1984: On the behavior of internal waves in the wakes of storms. *J. Phys. Oceanogr.,* **14,** 1129–1151.
Henyey, F. S., J. Wright, and S. M. Flatté, 1986: Energy and action flow through the internal wave field: An eikonal approach. *J. Geophys. Res.,* **91,** 8487–8495.
Hibiya, T., M. Nagasawa, and Y. Niwa, 1999: Model predicted distribution of internal wave energy for diapycnal mixing processes in the deep waters of the North Pacific. *Dynamics of Oceanic Internal Gravity Waves, II: Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop,* P. Müller and D. Henderson, Eds., SOEST, 205–213.
Hirst, E., 1991: Internal wave–wave resonance theory: Fundamentals and limitations. *Dynamics of Oceanic Internal Gravity Waves, II:Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop,* P. Müller and D. Henderson, Eds., SOEST, 211–226.
Kroll, J., 1975: The propagation of wind-generated inertial oscillations from the surface into the deep ocean. *J. Mar. Res.,* **33,** 15–51.
Kundu, P., 1993: On internal waves generated by travelling wind. *J. Fluid Mech.,* **254,** 529–559.
Kunze, E., 1985: Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.,* **15,** 544–565.
Leaman, K. D., and T. B. Sanford, 1975: Vertical propagation of inertial waves: A vector spectral analysis of velocity profiles. *J. Geophys. Res.,* **80,** 1975–1978.
Ledwell, J. R., E. T. Montgomery, K. L. Polzin, L. C. St. Laurent, R. W. Schmitt, and J. M. Toole, 2000: Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature,* **403,** 179–182.
Longuet-Higgins, M. S., 1969: On the reflexion of wave characteristics from rough surfaces. *J. Fluid Mech.,* **37,** 231–250.
McComas, C. H., and P. Müller, 1981: Time scales of interaction among oceanic internal waves. *J. Phys. Oceanogr.,* **11,** 139–147.
Mihaly, S. F., R. E. Thomson, and A. B. Rabinovich, 1998: Evidence for nonlinear interaction between internal waves of inertial and semidiurnal frequency. *Geophys. Res. Lett.,* **25,** 1205–1208.
Müller, P., and N. Xu, 1992: Scattering of oceanic internal gravity waves off random bottom topography. *J. Phys. Oceanogr.,* **22,** 474–488.
——, and D. Henderson, Eds., 1999: *Dynamics of Oceanic Internal Gravity Waves,* II: *Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop,* SOEST, 294 pp.
——, G. Holloway, F. Henyey, and N. Pomphrey, 1986: Nonlinear interactions among internal gravity waves. *Rev. Geophys.,* **24,** 493–536.
Munk, W. H., 1980: Internal wave spectra at the buoyant and inertial frequencies. *J. Phys. Oceanogr.,* **10,** 1718–1728.
——, 1981: Internal waves and small-scale processes. *Evolution of Physical Oceanography,* B. A. Warren and C. Wunsch, Eds., The MIT Press, 264–291.
——, and C. Wunsch, 1998: The Moon and mixing: Abyssal recipes II. *Deep-Sea Res.,* **45,** 1997–2010.
Nagasawa, M., Y. Niwa, and T. Hibiya, 2000: Spatial and temporal distribution of the wind-induced internal wave energy available for deep water mixing in the North Pacific. *J. Geophys. Res.,* **105,** 13 933–13 943.
Pinkel, R., 1984: Doppler sonar observations of internal waves: The wavenumber frequency spectrum. *J. Phys. Oceanogr.,* **14,** 1249–1270.
Polzin, K., J. M. Toole, and R. W. Schmitt, 1995: Finescale parameterizations of turbulent dissipation. *J. Phys. Oceanogr.,* **25,** 306–328.
Sanford, T. B., 1991: Spatial structure of thermocline and abyssal internal waves. *Dynamics of Oceanic Internal Gravity Waves, II: Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop,* P. Müller and D. Henderson, Eds., SOEST, 109–141.
Sherman, J. T., and R. Pinkel, 1991: Estimates of the vertical wavenumber–frequency spectra of vertical shear and strain. *J. Phys. Oceanogr.,* **21,** 292–303.
Sun, H., and E. Kunze, 1999a: Internal wave–wave interactions. Part I: The role of internal wave vertical divergence. *J. Phys. Oceanogr.,* **29,** 2886–2904.
——, and ——, 1999b: Internal wave–wave interactions. Part II: Spectral energy transfer and turbulence production. *J. Phys. Oceanogr.,* **29,** 2905–2919.
Young, W. R., and M. B. Jelloul, 1997: Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.,* **55,** 735–766.
Zervakis, V., and M. D. Levine, 1995: Near-inertial energy propagation from the mixed layer: Theoretical considerations. *J. Phys. Oceanogr.,* **25,** 2872–2889.