The Influence of Horizontal Inhomogeneities on the Propagation of High-Frequency Linear Internal Gravity Waves across a Baroclinic Flow

KEVIN G. LAMB

Department of Physics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada

JENNIFER A. SHORE*

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada

(Manuscript received 27 December 1990, in final form 2 August 1991)

ABSTRACT

The effect of the inclusion of the horizontal gradients of background density and baroclinic flow on the propagation of linear internal gravity wave packets across the flow is investigated using ray-tracing techniques. The focus is on the behavior of wave packets at the boundaries of regions supporting wave propagation as well as on identifying the locations of critical layers. For the case of isopycnals with constant slope but arbitrary spacing, it is shown that the critical-layer valves (critical layers for which one wave may pass through while another encounters a critical layer) found if the horizontal gradients are not included do not exist. An analytic example of a quite different basic state that has a critical-layer valve is presented. It has the property that for the wave encountering a critical layer, the group velocity goes to infinity as it is approached and the wavenumber also goes to infinity. The analytic and numerical results indicate that critical layers occur in only a few special cases. One case where they do occur is an important one; namely, that of high-frequency waves in a tidally mixed front.

1. Introduction

It is well known that, due to their low damping, internal gravity waves are capable of propagating hundreds of kilometers in the ocean. Over these length scales horizontal inhomogeneities of the background flow can significantly effect their propagation. Important horizontal variation of the background flow can occur on length scales comparable to the wavelength of the waves, particularly for low-frequency waves and in the vicinity of fronts and baroclinic jets. The effect of these variations is an important question in determining the energetics and mixing that may result. In particular, are there mechanisms for wave absorption such as critical layers, or are waves reflected?

During the past two decades a number of authors have investigated the propagation of linear internal gravity waves in a horizontally and vertically inhomogeneous ocean. These studies have largely been restricted to special cases. Samodurov (1974), Miropol'skiy (1974), and Miropol'skiy et al. (1976) considered the case of a horizontally varying but vertically constant buoyancy frequency N. In this case the places where the buoyancy frequency and the wave frequency are equal act as critical layers. This can potentially result in wave absorption and mixing. In contrast, it is well known that for a vertically varying buoyancy frequency these locations act as reflectors. See Olbers (1981) for a review of these cases.

Mooers (1975a) generalized the problem by taking more effects into account by including all linear terms in a perturbation expansion, including not only the vertical density gradient $\bar{\rho}_z$ (i.e., N^2), but $\bar{\rho}_x$ and $\bar{V}_x$ as well. (Here $\bar{\rho}$ and $\bar{V}$ are the background density and geostrophic flow, assumed to depend on x and z only.) For waves propagating across a baroclinic current he identified regions of wave propagation in which the wave frequency is larger than N as well as regions where it is smaller than f, the vertical component of the Coriolis parameter.

In a strict application of WKB methods, in which the wave amplitude α and the ratio of the wavelength to the scale of variation of the basic state, β, are given by the same small parameter, Olbers (1981) argued that the linear terms involving the gradients of the basic flow (i.e., the effects of advection of the basic flow by the wave) are negligible. In doing so it is possible to consider the propagation of waves in three dimensions. He considered in detail the two-dimensional case in which the isopycnals all have the same constant slope

* Present affiliation: Department of Oceanography, University of British Columbia.

Corresponding author address: Dr. Kevin G. Lamb, Memorial University of Newfoundland, Department of Physics, St. John's, A1B 3X7 NFDL, Canada.

and found that the waves are trapped in a region where the intrinsic wave frequency σ is less than N and is greater than a critical frequency $\sigma_c > f$. Between these two bounds he also found a critical layer exhibiting a valve effect, in that one set of waves passed through unaffected, while another set of waves encountered a critical layer. These types of critical layers were first described by Acheson (1973) in the context of waves propagating in shear flows.

Kunze (1985) retained the terms Olbers ignored, arguing that WKB methods are more robust than a strict application of the theory might imply. Being primarily interested in near-inertial (low frequency) waves, he simplified the exceedingly complex equation for 3D wave propagation to arrive at a more tractable problem. He showed that near-inertial waves are restricted to regions where $\sigma > f_{\text{eff}}$, where f_{eff} is equal to f plus half the relative vorticity of the background flow. Thus, regions of negative relative vorticity can trap waves with frequencies less than f, while regions of positive relative vorticity can act as barriers to waves with low frequencies. Consideration of a baroclinic jet showed that horizontal and vertical trapping in a negative vorticity region results in critical-layer phenomena. These results qualitatively bear out observations of enhanced energy levels on the negative vorticity sides of fronts (for example, see Kunze and Sanford 1984).

In this paper the problem considered by Mooers (1975a), with all gradients of the background velocity field and density retained, is reconsidered. The focus of attention is high-frequency waves, thus complementing the work of Kunze (1985). Observations of high-frequency waves propagating across a density front toward an unstratified region (Loder et al. 1992; Brickman and Loder 1992) on top of Georges Bank provided the initial motivation for this work. Our aim is to investigate the significance of the terms neglected by Olbers (1981). Because of the complexity of the general case only cross-stream propagation is considered; that is, the component of the wave vector in the direction of the background flow is taken to be zero. This restriction allows a strict application of WKB methods to determine paths followed by wave packets and to identify regions of reflection and critical behavior. The behavior of wave amplitudes along the rays is not discussed. The reader is referred to Lighthill (1978) or Olbers (1981), among others, for a discussion of this.

By including the horizontal gradients of the background flow in the linearized equations, we find that the critical layer with a valve effect (henceforth called critical-layer valves after Acheson 1973) found by Olbers for the two-dimensional case of isopycnals with constant slope is no longer present. The physical reason for the presence of these critical-layer valves in the absence of these terms is discussed. It is argued that such behavior is possible if these terms are retained and, indeed, an analytic example of it is presented. An important case with critical layers, the one that motivated this work, is that of a tidally mixed front. It is reminiscent of the trapping of near-inertial waves in regions of negative vorticity described by Kunze (1985). In this case, high-frequency waves trapped vertically encounter critical layers if they propagate horizontally toward unstratified water. The third example in section 4 discusses this case.

The plan of the paper is as follows. In section 2 the model is introduced, in section 3 the theory is developed, and in section 4 three examples are discussed. Finally, section 5 summarizes our findings.

2. The model

The governing equations are those for an incompressible, inviscid Boussinesq fluid on an f plane. In making the Boussinesq approximation the density is assumed to be constant everywhere in the momentum equations except in the buoyancy term. We let ρ_0 be this constant and define P as the normalized pressure p/ρ_0 and b as the buoyancy $\rho g/\rho_0$. The velocities are (u, v, w) in the (x, y, z) directions, x and y being the horizontal coordinates and z the vertical coordinate. The equations are nondimensionalized using a length scale L, a velocity scale U, and the convective time scale L/U. Taking typical values at Georges Bank (Loder et al. 1992) we have $L = 100$ m (based on water depth at bank edge) and $U = 0.1$ m s^{-1} (based on along-bank velocities). This gives a nondimensional value $f = O(0.1)$. We expect these scalings to be reasonable for other tidally mixed fronts. Typically N is about an order of magnitude or more larger than f.

The nondimensional governing equations are

$$
\begin{aligned}
u_t + uu_x + vu_y + wu_z - fv &= -P_x \\
v_t + uv_x + vv_y + wv_z + fu &= -P_y \\
w_t + uw_x + vw_y + ww_z &= -P_z - b \\
b_t + ub_x + vb_y + wb_z &= 0 \\
u_x + v_y + w_z &= 0.
\end{aligned}
\tag{2.1}
$$

The basic state considered is a steady geostrophic flow of the form

$$
(u, v, w, b, P) = (0, \bar{V}(x, z), 0, \bar{b}(x, z), \bar{P}(x, z)). \tag{2.2}
$$

From the governing equations we see that the basic state must satisfy

$$
f\bar{V} = \bar{P}_x \quad \text{and} \quad \bar{b} = -\bar{P}_z, \tag{2.3}
$$

from which we get the "thermal wind" equation

$$
\bar{b}_x = -f\bar{V}_z \tag{2.4}
$$

relating vertical velocity shear to the horizontal density gradient.

A perturbation is now introduced via

$$
(u, v, w, b, P) = (0, \bar{V}, 0, \bar{b}, \bar{P}) + \epsilon(\tilde{u}, \tilde{v}, \tilde{w}, \tilde{b}, \tilde{P}),
$$

where ϵ is a small parameter measuring the wave amplitude. Linearizing about the basic state gives

$$\tilde{u}_t + \bar{V}\tilde{u}_y - f\tilde{v} = -\tilde{P}_x$$
$$\tilde{v}_t + \bar{V}\tilde{v}_y + \tilde{u}\bar{V}_x + \tilde{w}\bar{V}_z + f\tilde{u} = -\tilde{P}_y$$
$$\tilde{w}_t + \bar{V}\tilde{w}_y = -\tilde{P}_z - \tilde{b}$$
$$\tilde{b}_t + \bar{V}\tilde{b}_y = \tilde{u}M^2 + \tilde{w}N^2$$
$$\tilde{u}_x + \tilde{v}_y + \tilde{w}_z = 0,$$

where N, given by

$$N^2 = -\bar{b}_z = -\frac{g\bar{\rho}_z}{\rho_0} \quad (2.5)$$

is the buoyancy frequency and M^2 is defined as

$$M^2 = -\bar{b}_x = -\frac{g\bar{\rho}_x}{\rho_0}. \quad (2.6)$$

Note that while $N^2 > 0$ is required for stability of the basic state there is no such constraint on M^2. Following Mooers (1975a) the squared notation is used solely for its analogy with N^2.

Kunze (1985) argued that WKB methods are more robust than a strict mathematical application would imply. Putting aside the question of the validity of this for the moment, we proceed following Kunze and express all variables in the form of a plane wave that is,

$$\phi = \phi_0 e^{i(kx+ly+mz-\sigma t)}; \quad \phi_0 \quad \text{const},$$

where ϕ can be any one of $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, $\tilde{P}$, or $\tilde{b}$. This leads directly to the dispersion relation, which (from Kunze 1985; also Mooers 1975b) can be written as

$$i\sigma_0^3(k^2 + l^2 + m^2) + \sigma_0^2[\bar{V}_x k + \bar{V}_z m]$$
$$- i\sigma_0[N^2(k^2 + l^2) + \sigma_f^2 m^2 - 2M^2 km]$$
$$+ [(-N^2\bar{V}_x + M^2\bar{V}_z)kl - fM^2 m]l = 0, \quad (2.7)$$

where

$$\sigma_f^2 = f(f + \bar{V}_x)$$

and

$$\sigma_0 = \sigma - \mathbf{k}\cdot\mathbf{V} = \sigma - l\bar{V}$$

is the intrinsic frequency. This equation for σ has three roots, two for the internal gravity modes of the form $\sigma = \pm q + ir$ with q real or imaginary and r real and an imaginary root $\sigma = is$ corresponding to the vortical mode. The region in which q is real is the region of wave propagation, the hyperbolic region. The waves are evanescent in the elliptic region where q is imaginary. If $l = 0$ then r and s are zero.

This dispersion relation is difficult to deal with analytically. Kunze (1985) was interested in low-frequency motion allowing him to suitably simplify a more complicated dispersion relation involving a more general basic state. Here we wish to include high-frequency motions in our investigation so that a simplification is not so straightforward. Thus, we consider the 2D case $l = 0$. This has a further mathematical advantage in that the preceding dispersion relation can be arrived at in a mathematically rigorous manner. Suppose that the wavelength is O(1) and that the basic state varies on the slow scale, that is, are functions of βx and βz where $\beta \ll 1$. Then one can have $\bar{V}_x$, $\bar{V}_z$, N^2, and M^2, all O(1). Note that this means that the baroclinic current V must be large, of $O(\beta^{-1})$, so that one would anticipate the results to be most useful in the vicinity of strong baroclinic fronts and jets.

Note that Eq. (2.7) can also be arrived at rigorously if one considers solutions with a weak along-current variation, that is, the case of $l = O(\beta)$.

3. Analytical results

This section contains the main analytical results. As previously discussed, they are restricted to the 2D case for which all terms are independent of y. Thus, $l = 0$ and the intrinsic frequency σ_0 and the wave frequency σ are identical. The three roots of (2.7) become $\sigma = 0$ for the vortical mode and $\sigma = \pm\Omega$ with

$$\Omega = \left(\frac{N^2k^2 + \sigma_f^2 m^2 - 2M^2 km}{k^2 + m^2}\right)^{1/2} \quad (3.1)$$

for the two internal gravity wave modes. The ray equations governing the motion of a wave packet are (Lighthill 1978)

$$\frac{dx}{dt} = \Omega_k = \frac{m}{\Omega}\left(\frac{km(N^2 - \sigma_f^2) + M^2(k^2 - m^2)}{(k^2 + m^2)^2}\right) \quad (3.2a)$$

$$\frac{dz}{dt} = \Omega_m = -\frac{k}{\Omega}\left(\frac{km(N^2 - \sigma_f^2) + M^2(k^2 - m^2)}{(k^2 + m^2)^2}\right) \quad (3.2b)$$

$$\frac{dk}{dt} = -\Omega_x = \frac{-1}{2\Omega(k^2 + m^2)} \times \left(k^2 \frac{\partial}{\partial x} N^2 + m^2 \frac{\partial}{\partial x}\sigma_f^2 - 2km\frac{\partial}{\partial x} M^2\right) \quad (3.2c)$$

$$\frac{dm}{dt} = -\Omega_z = \frac{-1}{2\Omega(k^2 + m^2)} \times \left(k^2 \frac{\partial}{\partial z} N^2 + m^2 \frac{\partial}{\partial z}\sigma_f^2 - 2km\frac{\partial}{\partial z} M^2\right). \quad (3.2d)$$

Here, the total derivative d/dt is the derivative moving with the local group velocity (Ω_k, Ω_m). Thus,

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \Omega_k\frac{\partial}{\partial x} + \Omega_m\frac{\partial}{\partial z}.$$

Since the basic state is independent of t, the frequency σ is constant and will always be assumed positive.

From (3.1) we have

$$\frac{k}{m} = \frac{M^2 \pm I^{1/2}}{N^2 - \sigma^2}, \tag{3.3}$$

where I is defined by

$$I = M^4 + (N^2 - \sigma^2)(\sigma^2 - \sigma_f^2). \tag{3.4}$$

Waves exist only in the hyperbolic regions where $I \geqslant 0$ given by

$$\frac{N^2 + \sigma_f^2}{2} - \left[\left(\frac{N^2 + \sigma_f^2}{2}\right)^2 + M^4\right]^{1/2} \leqslant \sigma^2 \leqslant \frac{N^2 + \sigma_f^2}{2} + \left[\left(\frac{N^2 - \sigma_f^2}{2}\right)^2 + M^4\right]^{1/2}. \tag{3.5}$$

The boundaries of the hyperbolic regions are the curves $I = 0$. From (3.5), as pointed out by Mooers (1975a), two consequences of including $\sigma_f^2 = f(f + \bar{V}_x)$ and M^2 in the dispersion relation are that waves with anomalously low frequencies $\sigma < f$ as well as waves with anomalously high frequencies $\sigma > N$ are possible (assuming of course that $N > f$). Note that the right side of (3.5) is always greater than N^2 so that the region in which waves of a given frequency can exist is made larger at the upper end; that is, they can pass through the boundary $N = \sigma$. The region can be decreased or enlarged at the lower end depending on the sign of $f\bar{V}_x$. The anomalously low-frequency waves have been investigated by Kunze (1985).

In addition to determining the behavior of wave packets at the boundaries of the hyperbolic regions, one of the purposes of this paper is to investigate the possibility of critical layers. The two well-known examples of critical layers in the context of internal gravity waves (M^2 and $f\bar{V}_x$ terms ignored) are the vertical critical layer occurring where $N(x) = \sigma$ (Samodurov 1974; Miropol'skiy 1974 or Miropol'skiy et al. 1976) and the horizontal layer where $\sigma - \bar{U}(z)k = f$, where $\bar{U}(z)$ is a horizontal mean flow in the x direction (Bretherton 1966; Olbers 1981). In both of these cases the critical curve is a curve that rays approach asymptotically while the group velocity goes to zero (or rather to the velocity of the mean flow) so rapidly that the curve is never reached. This is accompanied by unbounded growth of the wave amplitude and the wavenumber

$$K = (k^2 + m^2)^{1/2}. \tag{3.6}$$

Whether the critical layer is considered a viscous critical layer or a nonlinear critical layer depends on which of the neglected terms, nonlinear or viscous, become important first (Benney and Bergeron 1969). In example 3 of section 4 we find a curve that has the property that rays approach it asymptotically with the wavenumber going to infinity but with the group velocity becoming infinite! Whether the wave amplitude grows or not has not been determined. It would depend on which effect dominates: the convergence of the rays or the stretching. In either case, because the wavelength goes to zero we would expect viscous effects, which are present in any physical situation, to lead to absorption of wave energy into the mean flow unless the wave amplitude becomes very large, in which case a more complex nonlinear behavior, including reflection, may occur (for example, see Killworth and McIntyre 1985, who consider Rossby wave critical layers). In this paper we will define a critical curve to be a curve that has the property that the wavenumber K goes to infinity along a ray as the curve is approached asymptotically. No restriction on the behavior of the group velocity of the wave amplitude is imposed.

Curves of reflection are also curves along which the group velocity is zero. As rays approach these curves the group velocity goes to zero in such a way that the curve is reached in a finite time. The group velocity normal to the curve passes through zero resulting in reflection. There is one caveat. Although linear theory remains valid, the WKB solution breaks down near the reflecting curve and must be modified. For simple 1D problems the solution is modified through the use of Airy functions (Desaubies 1973). The behavior of the curve as a reflector is unchanged and we assume this to be the case for the more complicated geometries considered here.

With this in mind we now consider the group velocity G along a ray. Defining G as

$$G = \frac{ds}{dt} = (\Omega_k^2 + \Omega_m^2)^{1/2}, \tag{3.7}$$

where s is distance along the ray, we find that G can be written as

$$G = \frac{I^{1/2}}{\sigma K} \tag{3.8}$$

and that the total wavenumber K changes along a ray according to

$$\frac{dK}{dt} = -\frac{(N^2 - \sigma^2)^2}{2II\sigma}\left[\left(\frac{k}{m}\right)^2 \frac{\partial N^2}{\partial n} - 2\frac{k}{m}\frac{\partial M^2}{\partial n} + \frac{\partial \sigma_f^2}{\partial n}\right], \tag{3.9}$$

where

$$II = (N^2 - \sigma^2)^2 + (M^2 + pI^{1/2})^2 \tag{3.10}$$

with $p = \pm 1$ given by

$$\frac{k}{m} = \frac{M^2 + pI^{1/2}}{N^2 - \sigma^2} \tag{3.11}$$

and

$$\frac{\partial}{\partial n} = \frac{\mathbf{k} \cdot \nabla}{K} \tag{3.12}$$

is the derivative normal to the ray in the direction of the wave vector $\mathbf{k} = (k, m)$.

It is apparent that the group velocity can be zero only if $K = \infty$ or at the boundaries of the hyperbolic region where $I = 0$ as long as K is nonzero (as always appears to be the case). Unless $K \rightarrow \infty$ at the boundaries we expect wave reflection. Thus, we are led to investigating the behavior of K.

Because of the form of (3.9) it seems unlikely that anything can be said in general. The curve $z(x)$ followed by the ray is given by

$$\frac{dz}{dx} = -\frac{k}{m} = -\frac{M^2 + pI^{1/2}}{N^2 - \sigma^2} \quad (3.13)$$

and depends on the values of M^2, N^2, and σ_f^2, while the rate of change of K along the ray depends on the derivatives of these quantities normal to the curve.

There are, however, a few things that can be said. First consider the curve $I = 0$. From (3.8) and (3.9) we have

$$\frac{dK}{ds} = \frac{dK}{dt}\Big/G = -\frac{(N^2 - \sigma^2)^2}{2I^{1/2}II}\left[\left(\frac{k}{m}\right)^2 \frac{\partial N^2}{\partial n} - 2\frac{k}{m}\frac{\partial M^2}{\partial n} + \frac{\partial \sigma_f^2}{\partial n}\right]K \quad (3.14)$$

and as $I \rightarrow 0$

$$\frac{k}{m} \rightarrow \frac{M^2}{N^2 - \sigma^2}$$

$$II \rightarrow M^4 + (N^2 - \sigma^2)^2. \quad (3.15)$$

The slope of the ray $z(x)$ is given by (3.13) so that in general the ray is not parallel to the boundary $I = 0$. Hence, as the group velocity goes to zero the wave packet travels a finite distance. Assuming $M^2 \neq 0$ the behavior of K is determined by how I goes to zero because all the other terms in the coefficient of K on the right side of (3.14) can be regarded as constant, equal to their values at the point where the ray reaches the boundary.

Let

$$I \sim a(s - s_0)^\alpha \quad \text{as} \quad s \rightarrow s_0,$$

where $I = 0$ at $s = s_0$. Thus, from (3.14)

$$K \sim ce^{\gamma(s-s_0)^{1-\alpha/2}} \quad \text{as} \quad s \rightarrow s_0$$

and remains finite if $\alpha < 2$. Assuming that M^2, N^2, and σ_f^2 are smooth we expect that

$$I = M^4 + (N^2 - \sigma^2)(\sigma^2 - \sigma_f^2) \sim a_1(s - s_0) + a_2(s - s_0)^2 \cdots \quad \text{as} \quad s \rightarrow s_0$$

implying that in general $\alpha = 1$.

Thus, in general K remains finite and we expect reflection at the boundaries of the hyperbolic regions. The exception to this rule is the case when the ray approaches the boundary asymptotically. In this case the foregoing analysis is invalid because the distance traveled along the ray becomes infinite and the point s_0 is undefined. In addition, there is no reason that the terms in the coefficient of K, other than I, should approach constant values. Numerical results verify that the boundary acts as a critical layer if the rays approach it asymptotically. This is easily seen to be the case if M^2 and $f\bar{V}_x$ are ignored. Briefly reviewing this case, we have

$$I = (N^2 - \sigma^2)(\sigma^2 - f^2).$$

The boundary (for high-frequency waves) is $N^2 - \sigma^2 = 0$. Since

$$\frac{k}{m} = p\left(\frac{\sigma^2 - f^2}{N^2 - \sigma^2}\right)^{1/2}$$

we see that the ray becomes vertical as it approaches $I = 0$. Thus, the ray travels a finite distance if the curve $I = 0$ is not vertical and an analysis similar to the preceding shows that

$$K \sim Ae^{a(\sigma^2-f^2)^{-3/2}(s-s_0)^{1/2}}.$$

Thus, K approaches a finite value as $s \rightarrow s_0$, $G \sim (s - s_0)^{1/2}$ and the ray reaches the curve in a finite time and reflects. If the curve $N^2 - \sigma^2 = 0$ (i.e., $I = 0$) is vertical the ray approaches the boundary asymptotically, never reaching it. As is well known, $K \rightarrow \infty$ so that the boundary is a critical curve (Samodurov 1974; Miropol'skiy 1974; Miropol'skiy et al. 1976).

In conclusion, then, the inclusion of the M^2 and $f\bar{V}_x$ terms in the dispersion relation displaces the boundary between the hyperbolic and elliptic regions. The behavior of the boundary is unchanged in that it acts as a reflector unless the slope of the boundary is equal to the slope of the rays at the boundary, which, using $I = 0$ in (3.13) and (3.4), can be written as $(\sigma^2 - \sigma_f^2)/M^2$.

In the general case with M^2 and $f\bar{V}_x$ included the curve $N^2 - \sigma^2 = 0$ is of special interest. From (3.4), (3.10), and (3.11) we see that as $N^2 - \sigma^2 \rightarrow 0$

$$I \sim M^4, \quad (3.16a)$$

and that if $pM^2 > 0$

$$II \sim 4M^4, \qquad \frac{k}{m} \sim \frac{2M^2}{N^2 - \sigma^2}, \quad (3.16b)$$

while if $pM^2 < 0$

$$II \sim \left[1 + \left(\frac{\sigma^2 - \sigma_f^2}{2M^2}\right)^2\right](N^2 - \sigma^2)^2, \qquad \frac{k}{m} \sim -\frac{\sigma^2 - \sigma_f^2}{2M^2}. \quad (3.16c)$$

Note that if $pM^2 > 0$ the ray becomes vertical as the curve $N^2 - \sigma^2 = 0$ is approached.

Consider a ray approaching $N^2 - \sigma^2 = 0$ and assume that it does not approach the curve asymptotically. Then the ray reaches the curve after traveling a finite distance. Invoking an analysis similar to the preceding, we find that if $pM^2 > 0$

$$\frac{dK}{ds} \sim -\left(\frac{4M^4\partial N^2/\partial n - 4M^2(N^2-\sigma^2)\partial M^2/\partial n + (N^2-\sigma^2)^2\partial\sigma_f^2/\partial n}{8|M^2|M^4}\right)K, \tag{3.17a}$$

while if $pM^2 < 0$

$$\frac{dK}{ds} \sim -\left(\frac{\left(\frac{\sigma^2-\sigma_f^2}{2M^2}\right)^2\frac{\partial N^2}{\partial n} + 2\frac{\sigma^2-\sigma_f^2}{2M^2}\frac{\partial M^2}{\partial n} + \frac{\partial \sigma_f^2}{\partial n}}{2|M^2|\left[1+\left(\frac{\sigma^2-\sigma_f^2}{2M^2}\right)^2\right]}\right)K. \tag{3.17b}$$

Thus, K remains finite and from (3.8) the group velocity does not go to zero. Nothing special happens. Note that if the curve $N^2 - \sigma^2 = 0$ is not vertical then both rays pass through the curve with no effect. The ray with $pM^2 > 0$ always has infinite slope along the curve $N^2 - \sigma^2 = 0$, making this curve easy to spot on a plot of a family of rays.

The result is, of course, different if the curve $N^2 - \sigma^2 = 0$ is vertical and we consider the ray with $pM^2 > 0$. In this case the ray approaches the curve asymptotically. Equations (3.17) are still valid. For the case in question we see that

$$\frac{dK}{ds} \sim -\lambda^2\frac{\partial N^2}{\partial n}K, \tag{3.18}$$

where λ is a constant (assuming M^2 is bounded away from zero we can take M^2 to be constant). Thus, the behavior of K is determined by the sign of $\partial N^2/\partial n$. Since $k/m \to \infty$ we have

$$\frac{\partial N^2}{\partial n} \approx \frac{k}{|k|}\frac{\partial N^2}{\partial x}. \tag{3.19}$$

From (3.2a), as $k/m \to \infty$ the horizontal group velocity behaves like

$$\frac{dx}{dt} \approx \frac{mM^2}{\sigma k^2} = \left(\frac{k}{m}M^2\right)\frac{m^2}{k^3\sigma^2}.$$

From (3.16b) $(k/m)M^2$ has the same sign as $N^2 - \sigma^2$. Hence, we can write $dx/dt \approx \Lambda^2 k(N^2 - \sigma^2)$ for some real Λ so that

$$\frac{dx}{dt}\frac{\partial N^2}{\partial x}(N^2-\sigma^2) \approx \Lambda^2(N^2-\sigma^2)^2 k\frac{\partial N^2}{\partial x}. \tag{3.20}$$

Since we wish to approach the curve $N^2 - \sigma^2 = 0$, which minimizes $(N^2 - \sigma^2)^2$, it follows that

$$\frac{dx}{dt}\frac{\partial(N^2-\sigma^2)^2}{\partial x} < 0.$$

Hence, the left side of (3.20) is negative. Thus, (3.19) and (3.20) give

$$\mathrm{sgn}\left(\frac{\partial N^2}{\partial n}\right) = \mathrm{sgn}\left(k\frac{\partial N^2}{\partial x}\right) = -1.$$

Hence, from (3.18) K increases exponentially with distance implying that if the curve $N^2 - \sigma^2 = 0$ is vertical it is a critical curve. If $M^2 \neq 0$ on the curve then it is a critical-layer valve since the ray with $pM^2 < 0$ passes through it. Equation (3.8) shows that if M^2 is bounded, or increases more slowly than K, the group velocity goes to zero as the curve $N^2 - \sigma^2 = 0$ is approached.

4. Examples

a. Example 1

The first example is one for which the basic pressure is given by

$$\bar{P} = \Theta(Ax + Bz) = \Theta(\xi), \tag{4.1}$$

where A and B are real constants satisfying $A^2 + B^2 = 1$ and Θ is an arbitrary function of $\xi = Ax + Bz$. This example was considered by Olbers (1981) who, through neglect of the M^2 and $f\bar{V}_x$ terms in the dispersion relation, deduced that there was a critical valve at $\sigma^2 = N^2A^2 + f^2B^2$.

The pertinent functions in the dispersion relation are

$$N^2 = B^2\Theta'' \tag{4.2a}$$

$$M^2 = AB\Theta'' \tag{4.2b}$$

$$\sigma_f^2 = f^2 + A^2\Theta'', \tag{4.2c}$$

where primes denote differentiation w.r.t. ξ. Because N^2, M^2, and σ_f^2 are functions of the single variable ξ, the component of the wave vector parallel to lines of constant ξ, $Bk - Am$, is constant along a ray. Set

$$l = Bk - Am$$
$$q = Ak + Bm. \tag{4.3}$$

Then using (3.3) for k/m and (4.2) we find that

$$\frac{q}{l} = \frac{f^2AB - pI^{1/2}}{B^2f^2 - \sigma^2}. \tag{4.4}$$

Note the denominator is constant. This shows that q, and hence $K = (q^2 + l^2)^{1/2}$, is never infinite so that the group velocity G can only be zero at the boundaries $I = 0$ where there is wave reflection. In particular, there are no critical curves. Olbers, on the other hand, by simplifying the dispersion relation obtained a different form for q/l, with the denominator becoming

$$-A(N^2 - \sigma^2) + pB[(N^2 - \sigma^2)(\sigma^2 - f^2)]^{1/2},$$

which is zero when

$$\sigma^2 = A^2N^2 + f^2B^2 \tag{4.5}$$

provided $AB(N^2 - \sigma^2)p > 0$. Hence, this curve is a critical valve. Equations (4.2) show that M^2 and $f\bar{V}_x$ are negligible if $A \ll B$. Hence, in order to be consistent, one should ignore the A^2N^2 term in (4.5). It is the inclusion of the "small" M^2 and $f\bar{V}_x$ terms that cancels this term to give the denominator in (4.4).

Olbers' simplified case is, however, an enlightening one, particularly if we further assume that the B^2f^2 term is negligible. If we do so, the slope of the rays are

$$-\frac{k}{m} = -p\left(\frac{\sigma^2}{N^2 - \sigma^2}\right)^{1/2},$$

which, along $\sigma^2 = A^2N^2$, is equal to $-p|A|/|B|$. For the ray encountering a critical curve along $\sigma^2 = A^2N^2$, we have the slope as $-A/B$, that is, the ray is tangent to the isopycnals. Now for internal gravity waves, if the M^2 and σ_f^2 terms are ignored, fluid parcels oscillate along the ray. If the ray is parallel to the isopycnals the gravitational restoring force disappears. It is not surprising that there is a valve effect, since along $\sigma^2 = A^2N^2$ the gravitational restoring force is only missing for one of the rays. Of course with M^2 and σ_f^2 included the picture is more complicated, yet it seems reasonable to expect cases with critical valves. Example 2 is such a case.

Some numerical calculations were done to illustrate the nonexistence of critical valves if the M^2 and $f\bar{V}_x$ terms are retained. In order to determine how stable our result is to perturbations of the basic state, a number of cases of the form

$$\bar{P} = \Theta(Ax + Bz) + \epsilon Y(x, z)$$

were investigated numerically. Figure 1a shows the results for three values of ϵ for which $\Theta(\xi) = 0.1(\xi^2/2 + \sin\xi)$ with $A = B = 1/\sqrt{2}$, and $Y(x, z) = \sin[(x - z)/\sqrt{2}]$. For $\epsilon = 0$ the basic state is periodic, with waveguides running parallel to lines of constant ξ. For large values of ϵ (not shown) the perturbation connects the waveguides of the unperturbed state at regular intervals. Wave packets then appear to propagate up the unperturbed waveguide for a while and then appear to "jump" to another waveguide along which they travel for a considerable distance before jumping again. Several other perturbation functions $Y(x, z)$ were tried. In none of the cases for small perturbations was a critical valve found. Figure 1b shows the critical valve that is present if M^2 and $f\bar{V}_x$ are ignored.

b. Example 2

Consider the basic pressure field

$$\bar{P} = Ax\frac{z^2}{2} + B\frac{z^3}{6} \tag{4.6}$$

for which

$$N^2 = Ax + Bz \tag{4.7a}$$

$$M^2 = Az \tag{4.7b}$$

$$\sigma_f^2 = f^2. \tag{4.7c}$$

Although not physically realistic for the whole domain, since N^2 is negative if $Ax + Bz < 0$, this example is useful analytically because

$$R = k^2(3Am - Bk) \tag{4.8}$$

is constant along a ray.

Critical curves are located by considering locations where $k = \infty$ or $m = \infty$. Using the aforementioned we find that

$$\frac{k}{m} = \frac{3Ak^3}{R + Bk^3} = \frac{M^2 + pI^{1/2}}{N^2 - \sigma^2}$$

from which we obtain

$$k^3 = \frac{R(M^2 + pI^{1/2})}{3A(N^2 - \sigma^2) - B(M^2 + pI^{1/2})}. \tag{4.9}$$

From (4.8) $m = \infty$ when $k = 0$ or when $k = \infty$. Thus, there are two cases to consider.

1) CASE (i): $k = 0$

Consider the numerator of (4.9). For it to be zero, we need

$$M^4 = I = M^4 + (N^2 - \sigma^2)(\sigma^2 - f^2).$$

Thus, we need $N^2 = \sigma^2$ or $\sigma^2 = f^2$. For the first possibility we find that k is in fact nonzero: the denominator of (4.9) is zero and one can show that $k^3 = -R/B$ along $N^2 - \sigma^2 = 0$. This confirms what we already know; the curve $N^2 - \sigma^2 = 0$ in this case is not vertical and hence nothing special happens here. For the second possibility, $\sigma^2 = f^2$, we find that $k = 0$ everywhere and m is constant. This case is of no interest.

2) CASE (ii): $k = \infty$

From (4.9) we see that k is infinite along

$$3A(N^2 - \sigma^2) - BM^2 - pBI^{1/2} = 0,$$

which, after squaring, gives

$$Ax + \frac{B}{3}z = \left[1 + \left(\frac{B}{3A}\right)^2\right]\sigma^2 - \left(\frac{B}{3A}\right)^2 f^2. \tag{4.10}$$

Thus, along this curve k and m are infinite, provided p has the correct sign,

$$p = \text{sgn}\left(\frac{3A(N^2 - \sigma^2) - BM^2}{B}\right). \tag{4.11}$$

Here, then, is a critical valve. Some further analysis shows that I increases rapidly enough so that the group velocity $G \to \infty$ as $z \to +\infty$. Figure 2 shows some typical rays for this case. Note that the rays passing

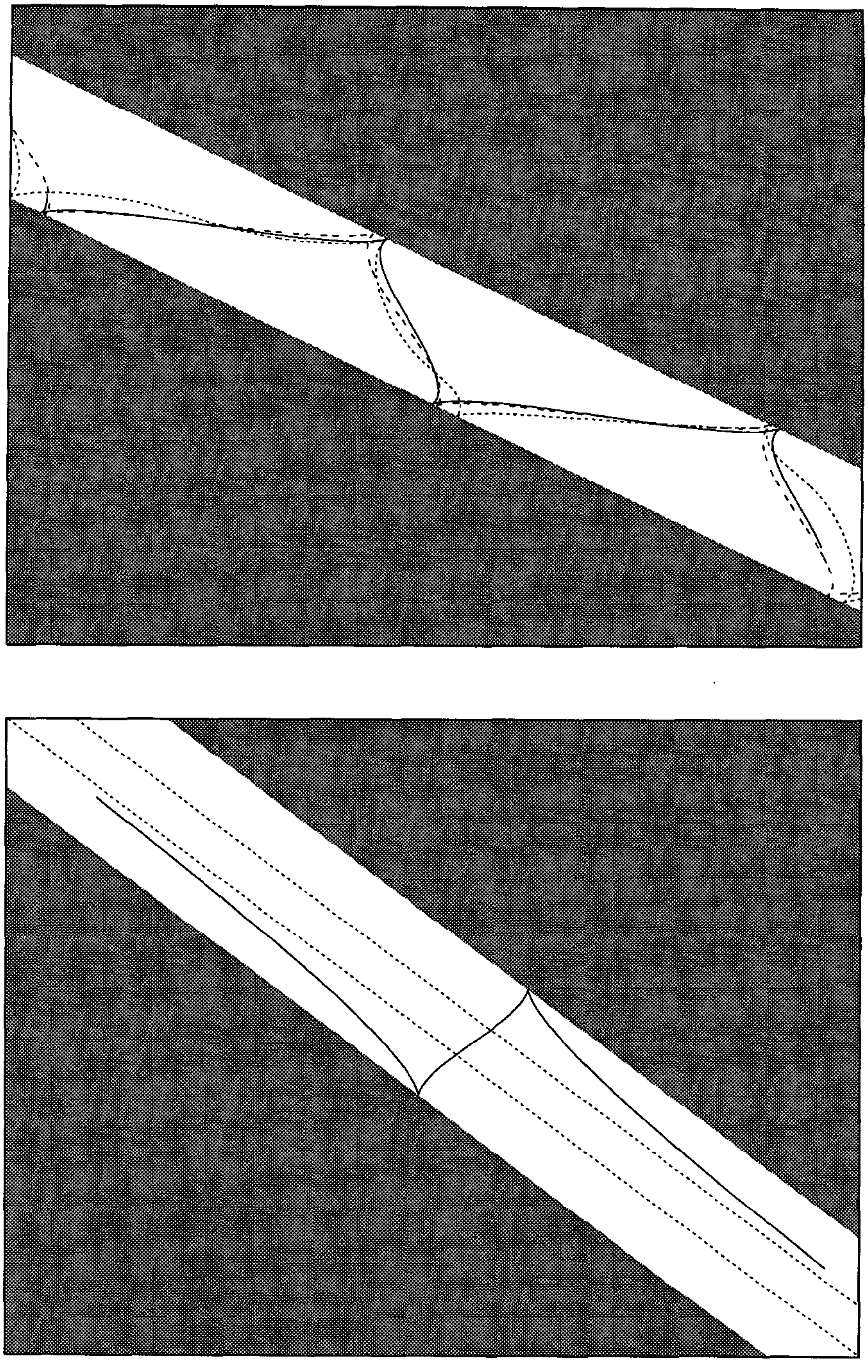

FIG. 1. Basic states of the form $\bar{P} = 0.1(\xi^2/2 + \sin\xi) + \epsilon \sin[(x - z)/\sqrt{2}]$ with $\xi = (x + z)/\sqrt{2}$ are considered. A single waveguide for the case $\epsilon = 0$ is left unshaded. Panel (a) shows rays for $\epsilon = 0$, 0.5, and 0.75 (solid curve, long dashed curve, and short dashed curve, respectively). For larger values of ϵ the shape of the waveguide changes dramatically, connecting waveguides that are separate at lower values of ϵ. These allow wave packets to "jump" from one waveguide to another. A critical valve was never observed. Panel (b) shows the results ($\epsilon = 0$) when the M^2 and $f\bar{V}_x$ terms are excluded. Here the ray encounters a critical valve (dashed line) along $\sigma^2 = (N^2 + f^2)/2$.

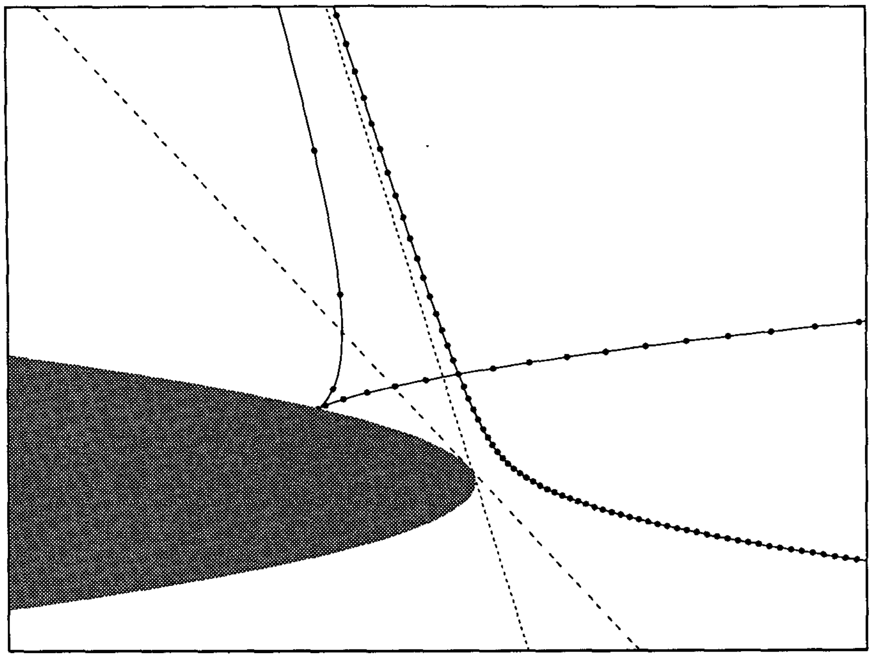

FIG. 2. Here the basic state has $\bar{P} = xz^2/2 + z^3/6$. The shaded region is the elliptic region where wave propagation is not possible (at the given wave frequency). The short dashed line is a critical valve and the long dashed line is the curve $N = \sigma$. The solid curves represent the paths of two rays, one passing through the critical valve and reflecting off the boundary of the hyperbolic region, while the other does not pass through the critical valve. The solid circles along the ray paths give locations of the wave packet at equal time intervals. It is clear that as the critical valve is approached the group velocity is increasing.

through the critical valve ultimately get caught there after being reflected off $I = 0$.

c. Example 3

Consider the basic pressure field

$$\bar{P} = -z + \left(\frac{1 + \tanh x}{2}\right) \ln(\cosh z) \qquad (4.12)$$

for which

$$\bar{b} = 1 - \left(\frac{1 + \tanh x}{2}\right) \tanh z \qquad (4.13a)$$

$$N^2 = \left(\frac{1 + \tanh x}{2}\right) \mathrm{sech}^2 z \qquad (4.13b)$$

$$M^2 = \frac{\mathrm{sech}^2 x \tanh z}{2} \qquad (4.13c)$$

$$\sigma_f^2 = f^2 - \mathrm{sech}^2 x \tanh x \ln(\cosh z). \qquad (4.13d)$$

This example has a few desirable features. First, $\bar{b} \in (1, 0)$ and $N^2 \in (0, 1)$. The latter corresponds to a dimensional buoyancy frequency of about 10^{-3} s^{-1} using the scalings mentioned in section 2. Contour plots of the buoyancy and of the buoyancy frequency are shown in Fig. 3a and Fig. 3b, respectively. We see a buoyancy field qualitatively similar to the time-averaged fields observed at Georges Bank (Loder et al. 1991), which is indeed qualitatively similar to a tidally mixed front (Garrett et al. 1981). To the left the fluid is unstratified, while on the right a pool of light warm water sits above denser cold bottom water. The contour plot of N suggests vertical trapping and the inability of wave packets to penetrate past a certain point to the left. The undesirable features of this example are of course that $\bar{V}$ is linear with z as $z \to \pm\infty$, resulting in very large values of σ_f^2 away from $z = 0$.

Figure 4 shows the paths followed by two wave packets of different frequencies, both propagating toward the unstratified region. The high-frequency wave packet ($\sigma = 0.8$, $f = 0.1$) encounters a critical layer as it tries to propagate out of the region of high N. The behavior of this wave packet is reminiscent of the vertically and horizontally trapped near-inertial wave packets discussed by Kunze (1985). Here the role of the horizontal and vertical trapping is reversed. For the high-frequency wave it is the horizontal trapping by an N contour passing through the vertical that results in the critical layer, whereas for the near-inertial wave the critical-layer behavior results from a σ_f contour passing through the horizontal.

The other ray shown in Fig. 4 is for a lower-frequency wave ($\sigma = 5f = 0.5$). This wave escapes from the region due to the influence of the M^2 and σ_f^2 terms.

Thus, it would appear that in such a density stratification most high-frequency (close to the maximum value of N) waves propagating onto a bank will encounter critical layers. Lower-frequency waves (perhaps not too low) may be more likely to escape past the curve $N^2 = \sigma^2$ and get far enough to be reflected at the bottom.

5. Summary

The propagation of internal gravity waves in an inhomogeneous background field, varying with x and z only, has been investigated. Only waves with no component in the alongflow direction were considered; thus, effects of Doppler shifting are not included. In addition, the domain was assumed infinite so that reflection at the sea surface or at the sea bottom was not treated.

One of the main goals of this work was to consider the effects of retaining the horizontal derivatives of the background density and baroclinic flow. Waves can exist in the hyperbolic region where

$$I = M^4 + (N^2 - \sigma^2)(\sigma^2 - \sigma_f^2)$$

is positive. Waves with frequencies higher than the local value of N or lower than f may exist. In particular, provided that M^2 is nonzero, wave packets can pass through the contour $N = \sigma$ to enter regions where $\sigma > N$. At the other end of the frequency range, the hyperbolic region is made larger if $f\bar{V}_x < 0$ and smaller if $f\bar{V}_x > 0$.

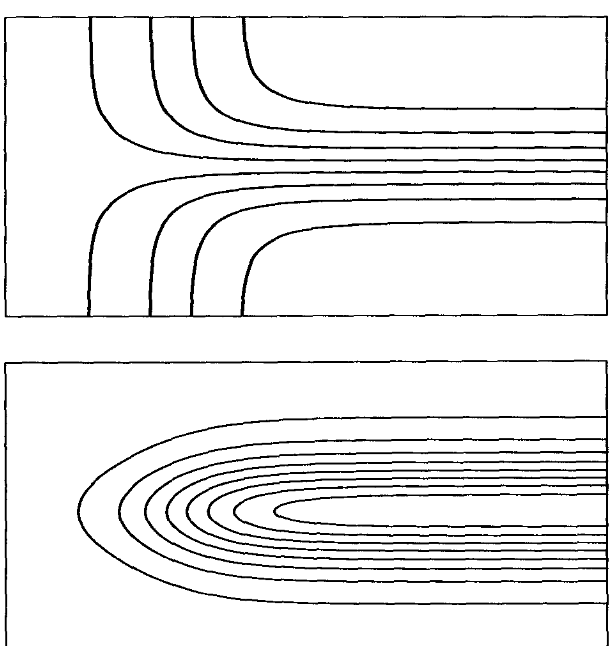

FIG. 3. The upper panel shows a contour plot of the density field $b = 1.0 - (1 + \tanh x)(\tanh z)/2$ that is qualitatively similar to ones observed at tidally mixed fronts. The lower panel is the contour plot of N^2. The plotted regions are the same in both cases. The density decreases with height and N increases from left to right.

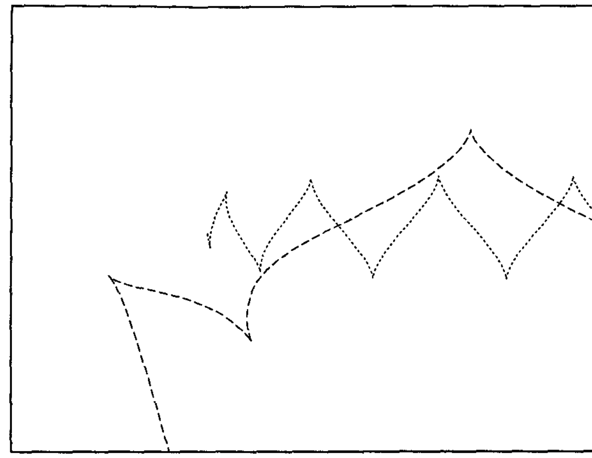

FIG. 4. This figure shows two rays for wave packets of different frequencies. For both cases $f = 0.1$. Note that $N \in (0, 1)$. The plotted region is identical to that in Fig. 3. The high-frequency wave with $\sigma = 0.8$ (short dashes) gets trapped in the region of high N. The lower-frequency wave, $\sigma = 0.5$, (long dashes) escapes because of the retention of the M^2 and $\bar{V}_x$ terms.

Critical curves are defined as curves for which the total wavenumber $K = (k^2 + m^2)^{1/2}$ becomes infinite as the curve is approached. As illustrated by the second example discussed, the group velocity does not necessarily go to zero. The boundaries of the hyperbolic region generally reflect wave packets. They act as critical curves only if they are tangent to the rays. This is also true if the M^2 and $f\bar{V}_x$ terms are ignored. In that case the curve $N = \sigma$ is a boundary of the hyperbolic region and is a critical curve only if it is vertical. With all the terms included, the curve $N = \sigma$ is still a critical curve only if it is vertical—but now for just one family of rays. Hence, in this sense using the complete set of terms has little effect, merely displacing the boundary of the hyperbolic region slightly.

The example of a basic flow field with isopycnals having constant slope considered by Olbers (1981), who neglected the horizontal gradients of the background flow in the dispersion relation, was reconsidered. It is shown that the critical-layer valves found by Olbers do not exist when the horizontal gradients are retained. Through numerical calculations this new result appears to be stable to small perturbations of the background field. Olbers had neglected these terms, arguing that they are small. The radically different behavior found when the terms are retained raises the question of how the neglected nonlinear terms might affect the behavior. It would be interesting to do a numerical simulation of the full nonlinear equations to investigate this.

It was argued that with the complete set of terms one should still have cases with critical-layer valves (apart from the case when the curve $N = \sigma$ is vertical). The second example discussed is such a case. One interesting feature of the critical curve for this example is that the group velocity becomes infinite as it is approached.

The final example considered is a case with a density transect qualitatively similar to those observed at Georges Bank and at other tidally mixed fronts. It traps high-frequency waves both vertically and horizontally if they are propagating onto a bank toward the unstratified water. For high-frequency waves, critical layers reminiscent of those for trapped low-frequency waves (see Kunze 1985) are obtained, although it is now the horizontal trapping that results in the critical-layer effects. For this particular example lower-frequency waves can escape from the region due to the M^2 and $f\bar{V}_x$ terms (the latter in particular). Thus, there is the possibility that high-frequency waves propagating onto a bank encounter critical layers, while low-frequency waves (which may not need to be too low, i.e., $\sigma = 5f$) may propagate farther past the curve $N = \sigma$, making them more likely to be reflected at solid boundaries.

Acknowledgments. This work was supported by grants to the first author from the National Sciences and Engineering Research Council of Canada and from the Canadian Department of Fisheries and Oceans. The second author was partially supported by an Undergraduate Summer Research Award by the National Sciences and Engineering Research Council of Canada.

REFERENCES

Acheson, D. J., 1973: Valve effect of inhomogeneities on anisotropic wave propagation. *J. Fluid Mech.* **58,** 27–37.

Benney, D. J., and R. F. Bergeron, Jr., 1969: A new class of nonlinear waves in parallel flows. *Studies in Appl. Math.,* **48,** 181–204.

Bretherton, F. P., 1966: The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Meteor. Soc.,* **92,** 466–480.

Brickman, D., and J. W. Loder, 1992: The energetics of the internal tide on Georges Bank. *J. Phys. Oceanogr.,* in press.

Desaubies, Y. J. F., 1973: Internal waves near the turning point. *Geophys. Fluid Dyn.,* **5,** 143–154.

Garrett, C. J. R., and J. W. Loder, 1981: Dynamical aspects of shallow sea fronts. *Phil. Trans. Roy. Soc. London,* **A302,** 563–581.

Killworth, P., and McIntyre, M. E., 1985: Do Rossby wave critical levels absorb, reflect or over-reflect? *J. Fluid Mech.,* **161,** 449–492.

Kunze, E., 1985: Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.,* **15,** 544–565.

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

Loder, J. W., D. Brickman, and E. P. W. Horne, 1992: Detailed structure of currents and hydrography on the northern side of Georges Bank. *J. Geophys. Res.,* in press.

Miropol'skiy, Yu. Z., 1974: Propagation of internal waves in an ocean with horizontal density field inhomogeneities. *Izv. Atmos. Oceanic Phys.,* **10,** 312–318.

——, N. I. Solntseva, and B. N. Filyushkin, 1976: On horizontal variability of the Brunt–Väisälä frequency in the ocean. *Oceanology,* **15,** 15–20.

Mooers, C. N. K., 1975a: Several effects of a baroclinic current on the cross-stream propagation of inertial–internal waves. *Geophys. Fluid Dyn.,* **6,** 245–275.

——, 1975b: Several effects of baroclinic currents on the three-dimensional propagation of inertial–internal waves. *Geophys. Fluid Dyn.,* **6,** 277–284.

Olbers, D. J., 1981: The propagation of internal waves in a geostrophic current. *J. Phys. Oceanogr.,* **11,** 1224–1233.

Samodurov, A. S., 1974: Internal waves in a medium with a horizontally-varying Brunt–Väisälä frequency. *Izv. Atmos. Oceanic Phys.,* **10,** 185–191.