On the Instability of Helmholtz Velocity Profiles in Stably Stratified Fluids When a Lower Boundary is Present

RICHARD S. LINDZEN AND ARTHUR J. ROSENTHAL

Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138

We investigate the stability of a Helmholtz velocity profile in a stably stratified fluid when a lower boundary is present. In addition to the traditional Kelvin-Helmholtz interfacial instability we find an infinitude of unstable internal gravity waves, the most unstable of which bear a close resemblance to disturbances observed in connection with clear air turbulence. It is shown that the gravity wave instabilities result from the ability of an unstable shear layer to overreflect such waves (reflected amplitudes are greater than incident amplitudes). Overreflection at the shear layer is constantly fed by total reflection from the lower boundary. The growth rate of gravity wave instabilities depends markedly on the distance between the shear layer and the ground. A mechanism is suggested whereby gravity wave instability can lead to long-lived almost neutral shear layers.

1. INTRODUCTION

In an earlier paper [*Lindzen*, 1974] it was shown that a Helmholtz velocity profile in an infinite fluid with constant stable stratification throughout would sustain not only interfacial instabilities (Kelvin-Helmholtz instabilities) but also neutral internal gravity waves radiating energy away from the interface. Although *Lindzen* [1974] only investigated solutions with zero real phase speed, neutral radiating solutions with nonzero phase speed also exist which are continuations of Kelvin-Helmholtz instabilities at low wave numbers. In section 2 of this paper we will review and generalize these earlier results.

It was also noted by *Lindzen* [1974] that the existence of neutral internal gravity wave solutions was equivalent to the existence of infinite overreflection (i.e., outgoing energy with no incoming energy). The question of the reflection of gravity waves incident from below the shear zone is dealt with in detail in section 3. We find that all gravity waves with phase speeds between $\pm U$ (see Figure 1) and wave numbers such that wave solutions are possible above and below the shear zone are overreflected (i.e., the amplitude of the reflected wave is greater than that of the incident wave). For phase speeds and wave numbers such that solutions are wavelike below the shear zone and exponential above we obtain total reflection (i.e., the amplitude of the reflected wave is equal to that of the incident wave), and for phase speeds outside $(-U, +U)$ and wave numbers such that wavelike solutions exist both above and below the shear layer the reflection coefficient is less than 1. It is found that the neutral solutions of section 2 are indeed associated with infinite overreflection.

Finally, in the work by *Lindzen* [1974] it was suggested that overreflection might lead to instability in the presence of a rigid lower boundary. The point was simply that a wave overreflected at the shear layer would be totally reflected at the lower boundary and returned to the shear layer to be further overreflected. It is the main purpose of this paper to confirm this last suggestion. In section 4 the stability properties of a Helmholtz velocity profile in a constantly stratified fluid possessing a rigid lower boundary are investigated. It is found that for each $k < N/U$ (where k is horizontal wave number, N is Brunt-Väisälä frequency, and U is one half the velocity jump in the shear layer; see Figure 1) there exist one or more unstable solutions, the number of unstable modes approaching infinity when $kU/N < \frac{1}{2}$. The real part of the vertical wave number is quantized by the distance between the shear layer and the ground. Indeed the presentation and analysis of our results are complicated by the fact that the stability properties depend not only on N, U, and k but also on the distance between the ground and shear layer. It is found that for each vertical mode number the greatest growth rate is closely associated with a neutral mode of the unbounded problem (except for the Kelvin-Helmholtz instabilities at large k). It is also found that (again apart from the Kelvin-Helmholtz instability) the vertical mode number associated with maximum instability is not necessarily the lowest one (i.e., about one-half vertical wavelength between the ground and the shear layer); rather it is that mode number whose vertical wavelength is closest to the vertical wavelength of the 'most efficient' neutral mode discussed by *Lindzen* [1974]. ('Efficiency' in the work by *Lindzen* [1974] referred to the ratio of energy flux away from the shear zone to amplitude of vertical displacement of the shear layer.) Consistent with this we find that those values of ground to shear layer distance for which the vertical quantization allows vertical wavelengths corresponding to the most efficient neutral waves are also associated with greater growth rates than other values of this distance. This behavior is remarkably suggestive of a laser.

In section 5 the relation between instability and overreflection is explored in detail. We show that for small growth rates (due either to small overreflection or a large distance between the shear layer and the ground) it is possible to accurately estimate growth rate on the basis of overreflection coefficient, vertical group velocity of the mode, and ground to shear layer distance. Such an approach, however, can greatly overestimate larger growth rates.

Section 6 deals with the relative importance of gravity wave and Kelvin-Helmholtz instabilities, while section 7 relates our results to some observed cases of clear air turbulence. Section 8 suggests how a relatively long-lived shear zone having Richardson number near 0.25 might be maintained through gravity waves. Section 9 briefly summarizes some of our results; it also discusses why previous studies may have missed the instabilities discussed in this paper and discusses some clear next steps in the investigation.

2. REVIEW OF RESULTS FOR AN UNBOUNDED FLUID

As was done by *Lindzen* [1974], we consider a Boussinesq fluid with a constant Brunt-Väisälä frequency N. The basic

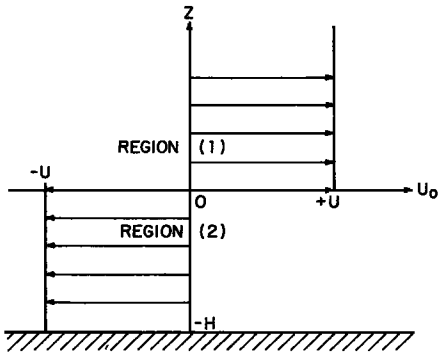


Fig. 1. Helmholtz velocity profile. For calculations in section 2 the lower boundary is removed.

velocity profile is shown in Figure 1: $U_0 = +U$ for $z > 0$ and $U_0 = -U$ for $z < 0$. Our equations are identical to those of *Lindzen* [1974] and will not for the most part be repeated here. Assuming solutions of the form

$$f' = f(z)e^{ik(x-ct)} \tag{1}$$

one obtains an equation for vertical velocity w applicable outside the shear zone

$$\frac{d^2w}{dz^2} + \left[\frac{N^2}{(U_0 - c)^2} - k^2\right]w = 0 \tag{2}$$

Recall, x is horizontal distance, t is time, z is vertical distance, and $c = c_r + ic_i$. Continuity of pressure and displacement at the shear interface requires continuity of the following quantities:

$$(U_0 - c)\, dw/dz \tag{2a}$$

and

$$w/(U_0 - c) \tag{2b}$$

We will refer to the region $z > 0$ as region 1 and $z < 0$ as region 2. For $z > 0$ the bracketed quantity in (2) is

$$l_1^2 = \frac{N^2}{(U - c)^2} - k^2 \tag{3}$$

In region 2 we have

$$l_2^2 = \frac{N^2}{(U + c)^2} - k^2 \tag{4}$$

If l_1^2 is negative or has an imaginary part, then our solution in region (1) corresponds to that l_1 which yields exponential decay with increasing z. If l_1^2 is positive, we require the radiation condition which, as shown by *Lindzen* [1974], requires that our solution behave as $\exp il_1 z$, where l_1 is the positive root of (3). Similarly, in region 2, in an unbounded fluid we also require either exponential decay with decreasing z or for $l_2^2 > 0$, the radiation condition. The latter requires solutions of the form $\exp il_2 z$, where l_2 is the positive root of (4).

We will first review neutral solutions for which $c_i = 0$. It will prove useful for our discussion to distinguish four regions in (k, c_r) space:

1. Region A is $l_1^2 > 0$, $l_2^2 > 0$, or equivalently, $k^2(U - c_r)^2 < N^2$ and $k^2(U + c_r)^2 < N^2$. In region A solutions propagate vertically for both $z > 0$ and $z < 0$.

2. Region B is $l_1^2 > 0$, $l_2^2 < 0$, or equivalently, $k^2(U - c_r)^2 < N^2$ and $k^2(U + c_r)^2 > N^2$. In region B solutions propagate vertically for $z > 0$ but decay for $z < 0$.

3. Region C is $l_1^2 < 0$, $l_2^2 > 0$, or equivalently, $k^2(U - c_r)^2 > N^2$ and $k^2(U + c_r)^2 < N^2$. In region C solutions decay for $z > 0$ but propagate for $z < 0$.

4. Region D is $l_1^2 < 0$, $l_2^2 < 0$, or equivalently, $k^2(U - c_r)^2 > N^2$ and $k^2(U + c_r)^2 > N^2$. In region D solutions decay for $z > 0$ and $z < 0$.

The above four regions are delineated in Figure 2.

We find neutral solutions satisfying continuity of quantities (2*a*) and (2*b*) only for region A. Two such solutions are found. For the first, $c_r = 0$ and $k \leq N/U$ (these were found by *Lindzen* [1974]); for the second, $c_r = \pm[N^2/(2k^2) - U^2]^{1/2}$ and $N/(2U) < k < N/(2^{1/2}U)$. Both solutions are indicated in Figure 2.

The unbounded configuration also has unstable solutions ($c_i > 0$) which exist only for $c_r = 0$. These unstable solutions are Kelvin-Helmholtz instabilities for which

$$c_i = \left(U^2 - \frac{N^2}{2k^2}\right)^{1/2} \qquad k \geq \frac{N}{2^{1/2}U}$$

Clearly, the second neutral solution is a continuation of the Kelvin-Helmholtz instability for $k \leq N/(2^{1/2}U)$. It was not described by *Lindzen* [1974] where $c_r \neq 0$ was not investigated. Otherwise, the reader is referred to *Lindzen* [1974] for further discussion of the unbounded case.

3. Overreflection

The existence of neutral radiating solutions in an unbounded fluid is, as we shall see, closely related to the matter of overreflection. To investigate this matter, we consider an upward propagating wave of the form

$$w_i = A_i \exp\{i[-l_2 z + k(x - c_r t)]\}$$

incident on the shear layer. In the terminology of section 2 we are restricted to regions A or C. The continuity of (2*a*) and (2*b*) require that there be a reflected wave

$$w_r = A_r \exp\{i[l_2 z + k(x - c_r t)]\}$$

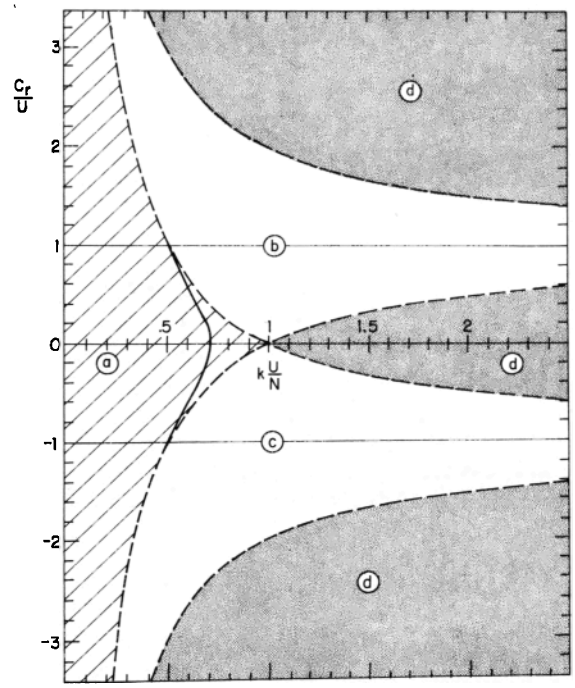


Fig. 2. Graph of c_r versus k for neutral radiating waves in unbounded fluid. Also shown are the various regions in k, c_r space discussed in section 2.

in region 2 and a transmitted disturbance in region 1

$$w_t = A_t \exp\{i[l_1 z + k(x - c_r t)]\}$$

If we are in region C of (k, c_r) space then l_1 is imaginary. It is readily shown that the reflection coefficient is given by

$$R_w = \left|\frac{A_r}{A_i}\right| = \left|\frac{(U + c_r)^2 l_2 + (U - c_r)^2 l_1}{(U + c_r)^2 l_2 - (U - c_r)^2 l_1}\right| \qquad (5)$$

From (5) we can immediately see the following:

1. When l_1 is imaginary, the numerator and denominator within the absolute value lines in (5) are complex conjugates of each other, and $R_w = 1$.
2. When l_1 is real and either $c_r < -U$ or $c_r > U$, the numerator in (5) is the difference of positive numbers, while the denominator is the sum of the same numbers. Hence $R_w \leq 1$.
3. When l_1 is real and $-U < c_r < U$, the denominator in (5) is the difference of positive numbers, while the numerator is the sum of the same numbers. Hence $R_w \geq 1$.

From (3), (4), and (5) it is easily confirmed that for case 3, R_w goes to infinity when $c_r = 0$ or $c_r = \pm[N^2/(2k^2) - U^2]^{1/2}$, i.e., for the neutral solutions in section 2. All solutions for case 3 are overreflected and, as we will see in sections 4 and 5, in the presence of a lower boundary all can be associated with instability.

4. Solutions for a Semi-Infinite Fluid

The equations for this case are identical to those in section 2, except that we now have a lower boundary at $z = -H$, where $w = 0$. As in section 2, quantities (2*a*) and (2*b*) must be continuous at $z = 0$.

Neutral solutions. We first investigate the nature of neutral solutions in the present configuration. In contrast to the situation in section 2 we find no neutral solutions in region A of (k, c_r) space. Only in region C do we find neutral solutions; these are associated with total reflection as discussed in section 3. Such solutions involve no net energy transmission away from the shear layer and hence are energetically neutral in addition to having $c_i = 0$.

The solutions in this case are of the form

$$w_1 = A_1 \exp(-n_1 z)$$

and

$$w_2 = -2iA_2 \exp(il_2 H) \sin[n_2(H + z)]$$

where

$$n_1 = +\left[k^2 - \left(\frac{N}{U - c_r}\right)^2\right]^{1/2} = il_1$$

and

$$l_2 = +\left[\left(\frac{N}{U + c_r}\right)^2 - k^2\right]^{1/2}$$

Continuity of quantities (2*a*) and (2*b*) at $z = 0$ leads to the dispersion relation

$$\tan l_2 H = \gamma \qquad (6a)$$

where

$$\gamma = -\frac{l_2}{n_1}\left(\frac{U + c_r}{U - c_r}\right)^2 \qquad (6b)$$

(Equation (6*a*) is what one would obtain from *Gossard*'s [1974] three-layer model by taking $\Delta H \to 0$. However, it should be noted that the limit $H/\Delta H \to \infty$ obtained by taking $H \to \infty$ is not the same as that obtained by taking $\Delta H \to 0$. The former limit can, for example, allow internal modes in the middle layer when the Brunt-Väisälä frequency there is large enough. Such a case was considered by *Gossard et al.* [1973].)

Equation (6*a*) may be rewritten

$$l_2 = 1/H(\arctan\gamma + m\pi) \qquad m = 0, 1, 2, 3, \cdots \qquad (6c)$$

Equation (6*c*) clearly leads to an infinite number of standing modes for $m = 1, 2, 3, 4, \cdots$ in region 2, each having approximately $m/2$ vertical wavelengths between $z = 0$ and $z = -H$. For the choice $H = 7.14U/N$ the resulting neutral curves are shown in Figure 3. For each mode number m there are two types of solutions, one with $-U < c_r < 0$ and the other with $c_r < -U$. As we shall show next in this section, the former continue into unstable solutions in region A (viz., section 2). This is consistent with the Miles-Howard theorem [*Howard*, 1961] which restricts c for unstable solutions to a semicircle of radius U. Finally, the solutions for other choices of H are qualitatively similar to those shown in Figure 3.

Unstable solutions. We now seek solutions for which $c_i \neq 0$. Our solutions are now

$$w_1 = A_1 \exp(-n_1 z) \qquad (7)$$

$$w_2 = -A_2 \exp(2n_2 H + n_2 z) + A_2 \exp(-n_2 z) \qquad (8)$$

where

$$n_1^2 = k^2 - \frac{N^2}{(U - c_r - ic_i)^2} \qquad (9a)$$

$$n_1^2 = k^2 - \frac{N^2[(U - c_r)^2 - c_i^2]}{[(U - c_r)^2 + c_i^2]} - i\frac{2N^2(U - c_r)c_i}{[(U - c_r)^2 + c_i^2]} = -l_1^2 \qquad (9b)$$

and

$$n_2^2 = k^2 - \frac{N^2}{(U + c_r + ic_i)^2} \qquad (10a)$$

$$n_2^2 = k^2 - \frac{N^2[(U + c_r)^2 - c_i^2]}{[(U + c_r)^2 + c_i^2]^2} + i\frac{2N^2(U + c_r)c_i}{[(U + c_r)^2 + c_i^2]^2} = -l_2^2 \qquad (10b)$$

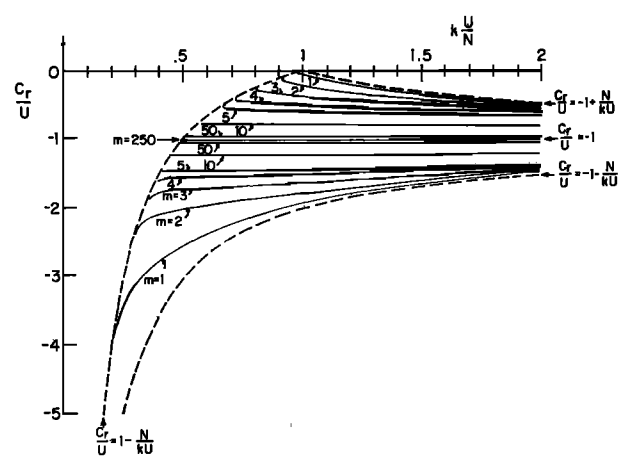


Fig. 3. Graph of c_r versus k for neutral modes in the presence of a lower boundary. Each curve is labeled by its mode number. Also shown are the boundaries of region C discussed in section 2. Note that all neutral solutions are now in region C rather than region A.

In general, n_1 and n_2 are complex if $c_i \neq 0$. The n_1 in (7) is chosen to be that root of (9*a*) with a positive real part. At the lower boundary we take $w_2(-H) = 0$. Continuity of (2*a*) and (2*b*) at $z = 0$ gives the following dispersion relation:

$$(U + c_r + ic_i)^2 n_2[1 + \exp(2n_2H)] = (U - c_r - ic_i)^2 n_1[1 - \exp(2n_2H)] \quad (11)$$

We wish to determine the behavior of c as a function of k for each N, U, and H. In general, at each k there will be several solutions corresponding to the internal modes described earlier in this section and the Kelvin-Helmholtz instability described by *Lindzen* [1974]. It is possible to exploit a measure of dynamic similarity by the following nondimensionalization:

$$\hat{n}_{1,2} = n_{1,2}U/N$$
$$\hat{k} = kU/N$$
$$\hat{c} = c/U$$
$$\hat{\omega} = \omega/N = kc/N$$
$$\hat{H} = HN/U$$

Equation (11) then becomes

$$(1 + \hat{c}_r + i\hat{c}_i)^2\hat{n}_2[1 + \exp(2\hat{n}_2\hat{H})] - (1 - \hat{c}_r - i\hat{c}_i)^2\hat{n}_1[1 - \exp(2\hat{n}_2\hat{H})] = 0 \quad (12)$$

where

$$\hat{n}_1 = \left[\hat{k}^2 - \frac{1}{(1 - \hat{c})^2}\right]^{1/2} \quad (13)$$

$$\hat{n}_2 = \left[\hat{k}^2 - \frac{1}{(1 + \hat{c})^2}\right]^{1/2} \quad (14)$$

Our solutions now depend on $\hat{H} = HN/U$ rather than on H, N, and U. To obtain our solutions, we treat the left-hand side of (12) as a function of $\hat{c}_r$ and $\hat{c}_i$ and search for its roots at each $\hat{k}$ for a given $\hat{H}$. The search is facilitated by our problem satisfying the conditions required by the Miles-Howard semicircle theorem (i.e., $w_1 \to 0$ as $z \to \infty$ and $w_2(-H) = 0$), and hence we may confine our search to the semicircle in the upper half of the complex $\hat{c}$ plane, where $\hat{c}_r^2 + \hat{c}_i^2 < 1$. The roots are determined iteratively by using the two-dimensional secant method [*Acton*, 1970]. Details are presented in the appendix.

Our results show both quantitative and qualitative variations with varying $\hat{H}$, and this makes a compact presentation difficult. We shall first present detailed results for two cases chosen for their relevance to observed situations which will be discussed in section 7. In Figures 4*a*–4*d* we show detailed results for $\hat{H} = 3.24$. In Figure 4*a* we show $\hat{c}_r$ versus $\hat{k}$ for the first few unstable modes. The curve associated with the Kelvin-Helmholtz instability behaves similarly to the Kelvin-Helmholtz instability in the unbounded case, including its gravity wave extension at low wave numbers. For $\hat{k}$ larger than 1 the Kelvin-Helmholtz instabilities have $\hat{c}_i^2 \approx 1 - 1/(2\hat{k}^2)$, $\hat{c}_r \approx 0$, and $|\hat{n}|^2 \approx |\hat{k}|^2$, as they do in the unbounded case. The remaining curves represent the continuations into region A of the neutral solutions in region C which were discussed in section 4. In Figure 4*b* we show $\hat{c}_i$ versus $\hat{k}$ for the same modes. The Kelvin-Helmholtz instability remains; however, now its gravity wave extension at small $\hat{k}$ is also unstable. Similarly, the extension of the neutral gravity wave modes into region A are unstable as well. For this case, growth rates diminish as mode numbers (appropriate to region 2) increase. It should also be noted that for all but mode numbers less than the most unstable mode number, the most unstable values of $\hat{k}$ are associated with these modes being close to the negative gravity wave continuation of the Kelvin-Helmholtz instability in the unbounded case (viz., Figure 2). At sufficiently small values of kU/N ($=\hat{k}$) the growth rates of the gravity modes are greater than the growth rate for the Kelvin-Helmholtz mode. In this particular case (as well as the second case we will discuss) the growth rate of the most rapidly growing gravity mode at that value of $\hat{k}$ where its growth rate is a maximum is still somewhat less than that of the Kelvin-Helmholtz mode. However, we have found many cases where this is not so. Figure 4*c* (part 1) shows Im $(n_1)U/N$ versus $\hat{k}$; neutral continuations of unstable solutions are not shown because Im $(n_1) = 0$ for these continuations. Figure 4*c* (part 2) shows Im $(n_2)H/\pi$ versus $\hat{k}$; we see that this quantity tends to be approximately an integer (which is henceforth called the mode number of the wave), an asymptotic result which can readily be established analytically. Finally, in Figure 4*d* (parts 1 and 2) we show $-\text{Re}\,(\hat{n})U/N$ versus $\hat{k}$ for regions 1 and 2. In general, there is a tendency for more rapidly growing modes to be most tightly trapped near $z = 0$, though there are some obvious exceptions. Figures 5*a*–5*d* show the same quantities for $\hat{H} = 7.14$. The most rapidly growing gravity mode is then associated with mode number 2 rather than 1. Also, in this case the behavior of the Kelvin-Helmholtz mode where $\hat{k} < 1$ appears closer to the behavior of a gravity mode; while the $m = 1$ gravity mode behaves like the Kelvin-Helmholtz mode of the preceding case. To a certain extent, this is a semantic problem; all instabilities at small $\hat{k}$ are associated with internal gravity waves. The identification of a given mode is based on continuity of the dispersion relation in $\hat{c}_r$, $\hat{c}_i$ space with $\hat{k}$ varying. Finally it should be noted that growth rates for the gravity modes are smaller in this case than in the preceding one.

The above two cases should offer the reader some idea of the nature of the instabilities that result from the presence of a lower boundary. Basically, the presence of the lower boundary quantizes Im $(n_2)H/\pi$. As a result of this there is generally only one value of $\hat{k}$ for each mode where that mode has $\hat{c}_r$ and $\hat{n}$ corresponding approximately to a neutral wave in the unbounded case (or equivalently, to a mode for which there is infinite overreflection). It is at approximately this value that we find the greatest growth rate for that mode. (This matter is explored in greater detail in section 5.) Moreover, *Lindzen* [1974] found that the neutral gravity wave for which $c_r = 0$ and $k = (1/2^{1/2})(N/U) = \text{Im}\,(n)$ was in some sense most efficient. The present quantization will allow this value of Im (n) only for certain values of $\hat{H}$, and intuitively, we expect that at these values of $\hat{H}$ the maximum growth rate associated with a gravity wave instability will be greater than for neighboring values of $\hat{H}$; i.e., if m is the mode number, we expect the greatest gravity wave instability when

$$\hat{H} = \frac{m\pi}{1/2^{1/2}} \cong \frac{m\pi}{0.707}$$

By actual calculation we find maximum growth rates when

$$\hat{H} = \frac{[m - (0.125 \pm 0.009)]\pi}{1/2^{1/2}} \quad m = 1, 2, 3, \cdots \quad (15)$$

The relevant results are shown in Figure 6*a*, where we show maximum growth rate associated with gravity wave instabilities versus $\hat{H}$. Several features are worth noting:

1. Sharp maxima occur for those values of $\hat{H}$ satisfying (15).

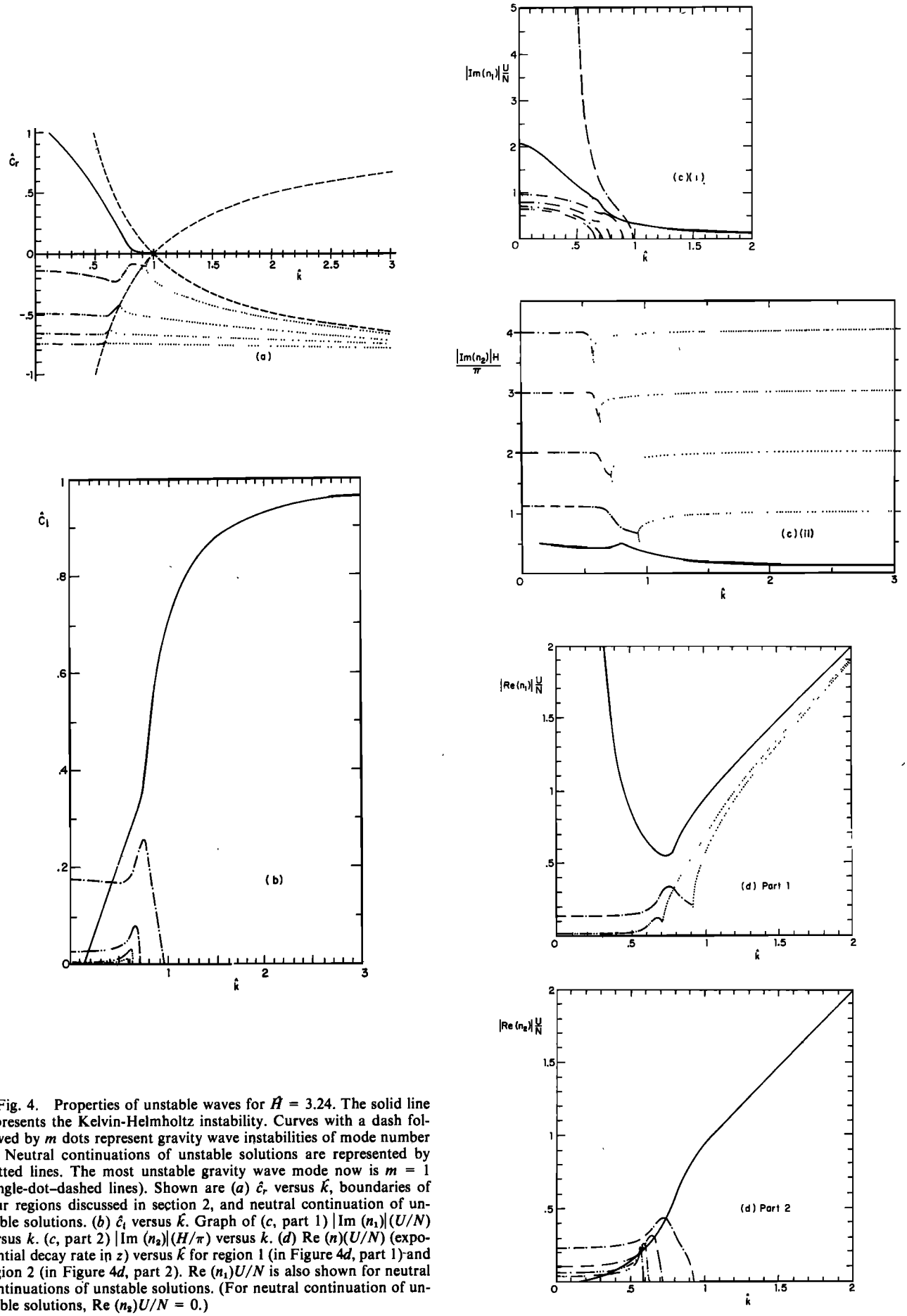
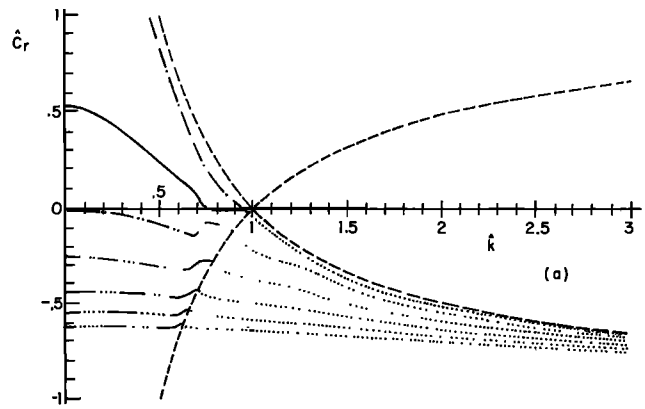
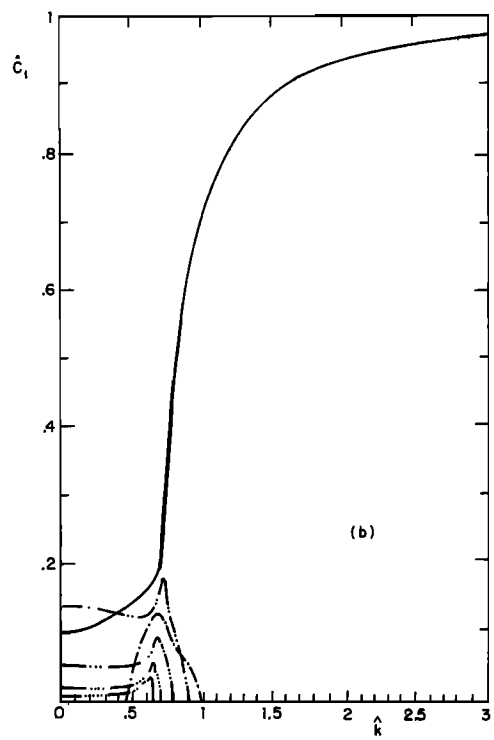
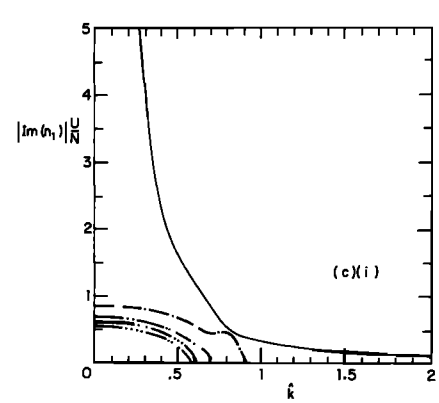
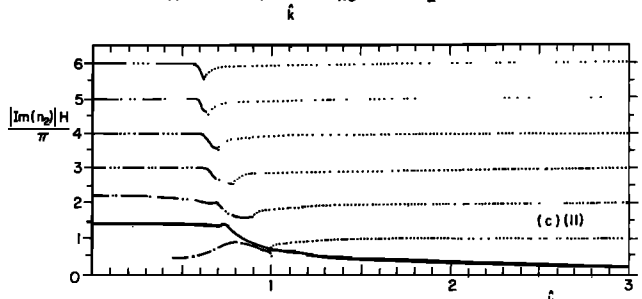


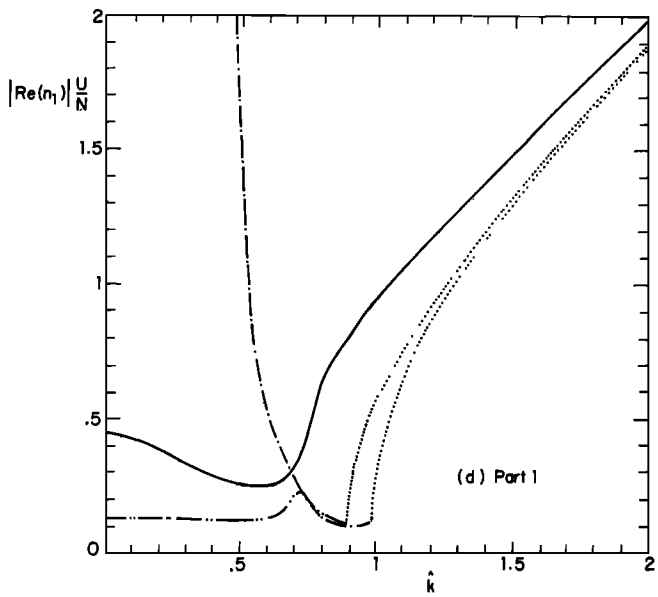
Fig. 4. Properties of unstable waves for $\hat{H} = 3.24$. The solid line represents the Kelvin-Helmholtz instability. Curves with a dash followed by m dots represent gravity wave instabilities of mode number m. Neutral continuations of unstable solutions are represented by dotted lines. The most unstable gravity wave mode now is $m = 1$ (single-dot–dashed lines). Shown are (*a*) $\hat{c}_r$ versus $\hat{k}$, boundaries of four regions discussed in section 2, and neutral continuation of unstable solutions. (*b*) $\hat{c}_i$ versus $\hat{k}$. Graph of (*c*, part 1) $|\text{Im}\,(n_1)|(U/N)$ versus k. (*c*, part 2) $|\text{Im}\,(n_2)|(H/\pi)$ versus k. (*d*) Re $(n)(U/N)$ (exponential decay rate in z) versus $\hat{k}$ for region 1 (in Figure 4*d*, part 1) and region 2 (in Figure 4*d*, part 2). Re $(n_1)U/N$ is also shown for neutral continuations of unstable solutions. (For neutral continuation of unstable solutions, Re $(n_2)U/N = 0$.)











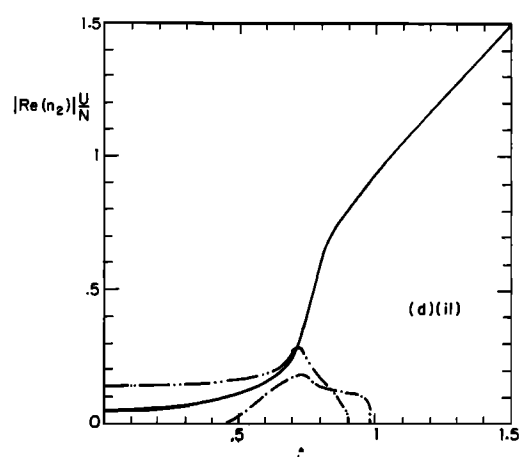


Fig. 5. Same as Figure 4 but for $\hat{H} = 7.14$. Now the most unstable gravity wave mode is $m = 2$ (double-dot–dashed line).

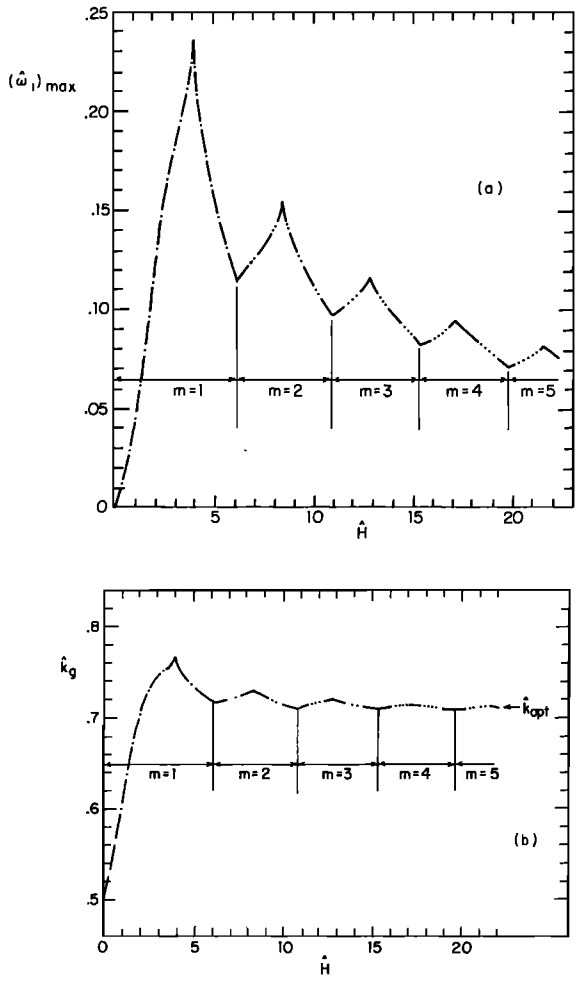


Fig. 6. (*a*) Maximum growth rate (divided by N) for gravity wave instabilities versus $\hat{H}$. Also shown are the mode numbers associated with the most unstable waves. (*b*) Wave number $\hat{k}_g$ for the gravity wave instability that has the most rapid growth at fixed $\hat{H}$. Also shown are the mode numbers associated with the most unstable gravity wave modes. The wave number most efficient at carrying energy away from the shear zone for an infinite fluid has been labeled $\hat{k}_{opt}$ on the graph. See Figure 4 for curve-labeling convention.

2. The maximum growth rate is associated with increasing mode number m as $\hat{H}$ increases. This is simply because maximum growth rate tends to be associated with a particular horizontal and vertical wavelength rather than a mode number, and as $\hat{H}$ varies, that mode number which comes closest to yielding the required wavelength is most unstable.

3. A general trend in Figure 6*a* is for maximum growth rate to decrease with increasing $\hat{H}$. This feature will be explained in detail in section 5.

Figure 6*b* shows the values of $\hat{k}$ associated with the instabilities of Figure 6*a*.

It should be remarked that the mode number (Im $(n_2)H/\pi$) for the gravity wave extension of the Kelvin-Helmholtz instability is generally close to that of the most unstable gravity wave mode when $\hat{k}$ is such that this gravity wave mode has maximal growth rate.

5. Instability and Overreflection

As we saw in section 4, at values of k for which only neutral gravity waves existed in the unbounded case (as well as at values of k where both gravity waves and Kelvin-Helmholtz instabilities coexisted) there are, in the presence of a lower boundary, instabilities. It has been suggested that these instabilities are associated with the phenomenon of overreflection. Supporting this suggestion is the fact that no neutral solution can be found in region A of Figure 2 (where there is overreflection). The purpose of this section is to explore this suggestion in detail.

Let us consider for the moment an instability for which c_i/U is sufficiently small that the main properties of the unstable mode (like n, for example) are determined by c_r (and k). From (5) of section 3 we can calculate R_w. The vertical group velocity in region 2 is given by $V_g = \partial\omega_r/\partial l_2$ where $\omega_r = kc_r = -kU + kN/(l_2^2 + k^2)^{1/2}$. More explicitly,

$$V_g = -\frac{kl_2N}{(l_2^2 + k^2)^{3/2}} = -\frac{kl_2(U + c_r)^3}{N^2} \tag{16}$$

The time τ it takes such a wave to go from the shear layer ($z = 0$) to the ground ($z = -H$) and back to $z = 0$ is

$$\tau = \frac{2H}{|V_g|} = \frac{2HN^2}{kl_2(U + c_r)^3} \tag{17}$$

If instability is due to overreflection, we would plausibly expect

$$R_w \approx e^{kc_i\tau} \tag{18a}$$

or

$$c_i \approx (\log R_w)/(k\tau) \tag{18b}$$

Equation (18*b*) could be used to estimate c_i, and this estimate could be compared with results from section 4. We find in practice that it is more convenient to take an instability calculated in section 4, calculate τ from (17) by using c_r and k from the unstable solution, use equation (18*a*) to determine an R_{eff} based on the calculated instability, and finally to compare R_w and R_{eff}. In general, it is possible, given any $\hat{k} < 1$ and mode number m, to have c_r take on a range of values by varying $\hat{H}$. Doing so leads to the results shown in Figures 7*a* and 7*b*. Figure 7*a* shows, for $\hat{k} = 0.6$, R_w versus $\hat{c}_r$, as well as R_{eff} versus $\hat{c}_r$ for various mode numbers (designated by the closest integer to Im $(n_2)H/\pi$). Note that R_w has three poles, as discussed in sections 3 and 2. For $\hat{k} < 0.5$, two of these disappear. Figure 7*b* shows results for $\hat{k} = 0.48$. In both cases we see that $R_{eff} \approx R_w$ away from the poles of R_w. It follows that in these regions, c_i as estimated from (18*b*) approximates the calculated c_i. Moreover, as the mode number increases, R_{eff} and R_w remain in agreement ever closer to the poles of R_w. The agreement of R_{eff} and R_w for smaller values of R_w as well as for higher mode numbers are both associated with c_i being small. Clearly, τ tends to increase with increasing mode number, and from (18*b*) both small R_w and large τ are associated with small c_i. At relatively large values of c_i it is no longer quantitatively correct to calculate properties of the unstable mode as though it were a neutral wave. It is also clear that the Miles-Howard theorem limits agreement near the poles of R_w. It should be noted that although R_{eff}/R_w tends to be smallest for mode number 1, this mode tends to have the largest growth rates (largely due to τ being small). We also find that any value of R_w in excess of 1 is associated with instability. The qualitative and quantitative agreement of R_{eff} and R_w in Figure 7 convincingly identifies overreflection with the instabilities discussed in section 4. Figure 7, which shows results for fixed values of $\hat{k}$, does not show how R_{eff} behaves for the most unstable gravity modes at each $\hat{H}$. Figure 8 shows R_{eff} versus $\hat{c}_r$ for the most unstable modes shown in Figure 6. The max-

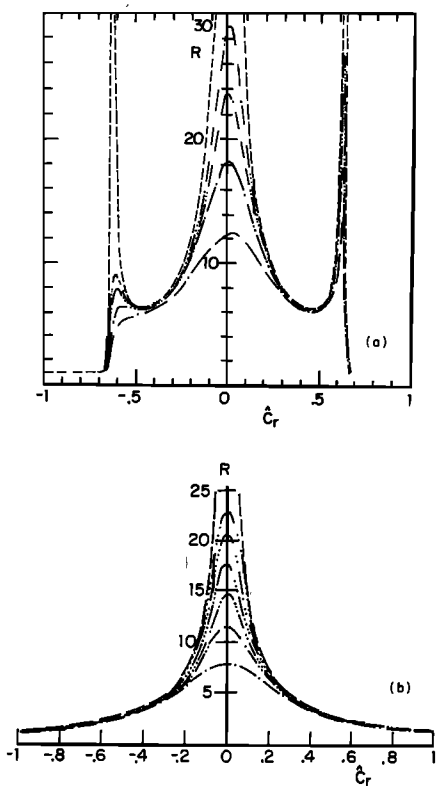


Fig. 7. Overreflection coefficient versus $\hat{c}_r$ (dashed line) for (*a*) $\hat{k} = 0.6$ and (*b*) $\hat{k} = 0.48$. Also shown are the effective overreflection coefficients based on growth rates for various mode numbers (see section 5 for details). See Figure 4 for curve-labeling convention for gravity waves.

imum values of R_{eff} are, not surprisingly, greater than those in Figure 7. Even so, for the most rapidly growing gravity wave instability (which occurs for $\hat{k} = 0.7684$ and $\hat{H} = 3.96504$ and for which $\hat{c}_i = 0.306$), R_{eff} is still only about 45. This suggests that a gravity wave instability might have a growth rate comparable to those found in this paper at Richardson numbers greater than zero, provided that the maximum overreflection coefficient is somewhat greater than 45 rather than infinity as in the present case.

6. Gravity Wave Instability Versus Kelvin-Helmholtz Instability

It is evident from the results in section 4 that for the present geometry the largest growth rates are still associated with Kelvin-Helmholtz instabilities with $\hat{k}\hat{c}_i$ ($= \hat{\omega}_i$) approaching infinity as $\hat{k} \to \infty$. Indeed, for $\hat{k} > 1$ the presence of a lower boundary barely affects this instability. *Lindzen* [1974] noted that Kelvin-Helmholtz instabilities would tend to destroy unstable shears on a more rapid time scale than would wave radiation. It can also be shown that the Kelvin-Helmholtz instabilites, in the present case, would destroy shears more rapidly than the gravity wave instabilities. Although we shall return to this point in sections 7 and 8, we will omit details which for the most part can be found in the work by *Lindzen* [1974]. The main reason the gravity wave instabilities are of interest is that observed shear instabilities in the atmosphere appear to have wave numbers very close to those associated with the most unstable gravity waves, and layered structures extending far from the shear levels. For the present geometry, Kelvin-Helmholtz instabilities would tend to emphasize larger horizontal wave numbers for which exponential decay in z is much faster than for the gravity wave instabilities of lower k, as shown by Figures 4*d* and 5*d*.

The only feature of the present results which would suggest the prominence of the gravity wave instabilities is the fact that they decay away from the shear layer more slowly than the Kelvin-Helmholtz instability and hence would be observable over a larger region. Of more likely importance, however, is the fact that the present shear profile is highly unrealistic. In most observed situations the Richardson number is closer to 0.25 than to zero, the shear being distributed over a finite region. The existence of shear zones of finite thickness suppresses Kelvin-Helmholtz instabilities with large wave numbers (associated with horizontal scales small compared with the thickness of the shear zone). It was shown by *Jones* [1968] that overreflection exists for all Richardson numbers under 0.25; in section 5 we showed that the largest growth rates for gravity wave instabilities were associated with overreflection coefficients, R_{eff}, which were less than infinite and which might therefore be associated with Richardson numbers greater than zero (zero corresponds to the Helmholtz velocity profile). It is conceivable that as the Richardson number (Ri) approaches 0.25 from below, the growth rate for Kelvin-Helmholtz instabilities might become less than that for the most unstable gravity wave, at least for certain values of $\hat{H}$ (viz., Figure 6). We are currently investigating this possibility.

There is a related possibility. *Miles and Howard* [1964] showed (for a somewhat different geometry) that the most unstable Kelvin-Helmholtz instability has a wavelength which is $\sim 7.57(2d)$ where d is essentially half the width of the shear zone. For $Ri = 0.25$, $d \approx U/2N$ and the wavelength is approximately $7.57U/N$, which is quite close to the wavelength we find to be associated with the most unstable gravity wave ($\approx 8.5U/N$). For the wavelength found by Miles and Howard (which would be associated with a neutral disturbance at $Ri = 0.25$) the disturbance will in fact be a gravity wave away from the shear zone. It is therefore conceivable that at $Ri = \frac{1}{4}$ the two types of instability merge. Some evidence for this possibility has in fact been described in section 4, where we found that in some cases the most unstable gravity wave was almost indistinguishable from the gravity wave continuation of the Kelvin-Helmholtz instability. It is interesting to note that although the wavelength of the dominant Kelvin-Helmholtz instability increases in proportion to d, the wavelength of the dominant gravity wave instability is likely to remain close to

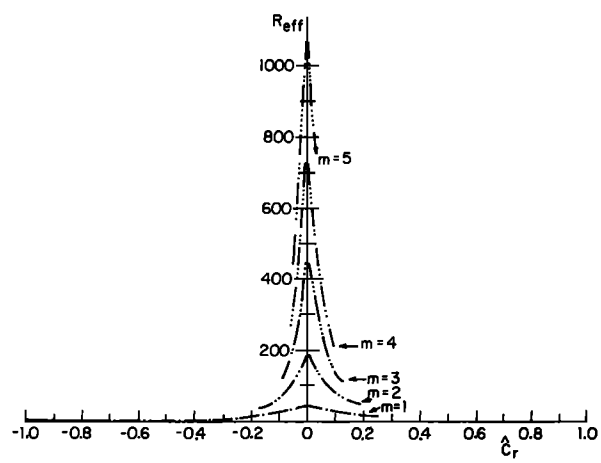


Fig. 8. Effective overreflection coefficient versus $\hat{c}_r$ for the most unstable gravity waves as given in Figure 6. See Figure 4 for curve-labeling convention.

TABLE 1. Parameters for Observed Cases of Clear Air Turbulence

	Location	Date	Reference	H, km	U, m s^{-1}	N, 10^{-2} s^{-1}	$\hat{H}$	Observed Wavelength, km	Lifetime, min	Observed vertical displacement, m
Case 1	Wallops Island, Va.	Feb. 7, 1968	*Ottersten et al.* [1973]	11	10	2	22	6	NR	250
Case 2	Haswell, Colo.	Nov. 12, 1971	*Hooke et al.* [1973]	0.08–0.15	1	2.7	3.24	0.35	10	15
Case 3	Wallops Island, Va.	Feb. 19, 1970	*Hardy et al.* [1973]	2–3	7	2	7.14	2.7	10	250
Case 4	Wallops Island, Va.		*Reed and Hardy* [1972]	9	28	1.3	4.18	15–20	NR	10[a]

NR, not reported.

the observed wavelength independent of d. In this respect, it may prove correct to regard the observed instability as a continuation of the gravity wave instability, even though the gravity wave instability has a smaller growth rate at $Ri = 0$.

7. Specific Cases

We return in this section to the four specific observed cases of shear instability discussed by *Lindzen* [1974]. In the work by *Lindzen* [1974] it was found that observed wavelengths were close to those of the most efficient waves. In the light of the results of section 4 the observed wavelengths must also be close to those of the most unstable gravity waves. In Table 1 we repeat the salient characteristics of the observed cases. For each of the four cases we have calculated detailed stability properties based on a Helmholtz velocity profile. The results for cases 2 and 3 were shown in section 4. In Table 2 we show results for all four cases pertaining to the most unstable ($\hat{\omega}_i$ ($= \hat{k}\hat{c}_i$) maximum) gravity wave instability. Also shown for reference purposes are some results from the unbounded case reviewed in section 2. The agreement between calculated and observed horizontal wavelengths has already been noted. We also see that the crudely estimated lifetime for an unstable shear layer is reduced in the presence of a lower boundary. These estimated lifetimes are, it should be noted, much shorter than observed lifetimes. We shall return to this point in section 8. Data are not available for comparison with the remaining quantities in Table 2. However, in connection with case 3, *Hardy et al.* [1973] did report that the observed waves traveled somewhat slower than the mean velocity in the shear layer, consistent with our calculated value for c_r/U of -0.104.

8. Possible Equilibration Mechanism

It was noted in section 7 that lifetimes for unstable shear zones estimated by the crude approach of *Lindzen* [1974] were much shorter than observed lifetimes. In this section we will briefly suggest an explanation for the relatively long observed lifetimes. *Booker and Bretherton* [1967] showed that for $Ri > \frac{1}{4}$, internal gravity waves are essentially absorbed at critical levels. *Lindzen* [1968] showed that such absorption led to a sharpening of shears and suggested that this mechanism might produce unstable shear zones associated with clear air turbulence, at least in some cases. However, as *Jones* [1968] showed, when $Ri < \frac{1}{4}$, overreflection rather than absorption occurs at a critical level; and as we have shown in section 5, overreflection is always associated with instability when a lower boundary is present. Also, instability is generally associated with reduction of unstable shears. We may now envisage the following sequence of events:

1. An unstable shear zone is established, either by critical level absorption or by any other mechanism (upper level frontogenesis for example).
2. A gravity wave instability sets in involving overreflection at the shear layer and reflection at the lower boundary.
3. The Richardson number of the shear zone rises above 0.25.
4. That portion of the gravity wave instability reflected from the lower boundary is now absorbed at the shear layer, leading to a resharpening of the shear.

We suggest that relatively long lifetimes of unstable shear zones can result from the sequential repetition of items 1–4 above, with the initially unstable gravity wave serving to resharpen shears when the flow becomes slightly stable.

9. Conclusions and Suggestions

We have shown in this paper that the overreflection of internal gravity waves from an unstable shear layer leads to the existence of a set (generally infinite) of unstable modes when a lower boundary is included. We refer to these as gravity wave instabilities. Although our explicit results are for a Helmholtz velocity profile, *Jones*'s [1968] finding that overreflection exists whenever $Ri < \frac{1}{4}$ implies the existence of such instabilities for all unstable shear flows in stably stratified flows. Although these gravity wave instabilities bear close resemblance to observed instabilities, they tend, for the present geometry, to have smaller growth rates than the traditional Kelvin-Helmholtz instabilities. We are currently investigating whether this remains the case for smooth velocity profiles as $Ri \to \frac{1}{4}$ from below. In this connection one might ask why the infinitude of gravity wave instabilities have not been previously found. For

TABLE 2. Characteristics of the Most Unstable Gravity Wave Instabilities for Helmholtz Velocity Profiles Using Parameters From Table 1.

Case	λ_i, km	λ_{opt},* km	τ_i, s	τ_w,* s	c_r/U	c_i/U	$(kc_i)^{-1}$, s	$\lvert 2\pi/\mathrm{Im}\ n_2\rvert$, km	$\lvert 1/\mathrm{Re}\ n_2\rvert$, km	$\lvert 2\pi/\mathrm{Im}\ n_1\rvert$, km	$\lvert 1/\mathrm{Re}\ n_1\rvert$, km
1	4.41	4.44	110	133	0.026	0.109	647	4.77	3.35	4.31	3.19
2	0.309	0.329	82	150	−0.150	0.257	192	0.283	0.087	0.518	0.109
3	3.04	3.1	54	65	−0.104	0.178	387	2.70	1.24	4.08	1.50
4	17.9	19.1	141	237	0.108	0.273	373	27.6	5.93	18.0	5.04

Here λ_i, wavelength of instability; λ_{opt}, wavelength of 'most efficient' wave; τ_i, lifetime of unstable shear due to gravity wave instability; τ_w, lifetime of unstable shear due to wave radiation.

*These quantities (appropriate to an unbounded fluid) are taken from *Lindzen* [1974] and are shown for reference purposes.

most studies the answer is clear: both shear and stability were confined to a layer, and hence internal gravity waves cannot exist away from the shear layer. For the geometry used in this paper it is easily shown that gravity wave instabilities disappear when $N = 0$ in region 2. There have, however, been studies where gravity wave instabilities would have been expected [e.g., *Jones*, 1968]. In these studies we can only suspect that the numerical search was incomplete. It has been our experience that if one searches by seeking instabilities which are the continuation of instabilities at large $\hat{k}$ (these are inevitably traditional Kelvin-Helmholtz instabilities), one will generally miss gravity wave instabilities, even at values of $\hat{k}$ for which the latter have greater growth rates.

Despite the unrealistic aspects of the present geometry we have obtained one result which we believe has observational implications which could be explored. From Figure 6, discussed in section 4, we see that growth rates for gravity wave instabilities depend markedly on $\hat{H}$, especially when $\hat{H}$ is between 0 and 6.05, where the most unstable mode number is 1. We do not expect this dependence to be significantly different for other shear profiles, since the physical nature of gravity wave instabilities should remain the same (including the need for quantization, etc.). Hence we expect that a correlation should exist between such quantities as frequency and intensity of clear air turbulence and the closeness of $\hat{H}$ to the values in Figure 6 (or its modification for more realistic shears) for which maximum growth rates are obtained. This feature may aid in predictions of such turbulence. It may be noted that the values for $\hat{H}$ for cases 2 and 4 in section 7 are, in fact, reasonably close to the most unstable value in Figure 6.

APPENDIX: TWO-DIMENSIONAL SECANT METHOD

For fixed $\hat{H}$ and $\hat{k}$ we want to find the intersection of the zeros of the real and imaginary parts of the left-hand side of (12), henceforth denoted by $R(\hat{c}_r, \hat{c}_i)$ and $I(\hat{c}_r, \hat{c}_i)$, respectively. Having obtained initial estimates $(\hat{c}_{r,1}, \hat{c}_{i,1})$, $(\hat{c}_{r,2}, \hat{c}_{i,2})$, and $(\hat{c}_{r,3}, \hat{c}_{i,3})$ for a root of $R + (-l)^{1/2}I$ by, for example, a search-and-step method, our new estimate according to the secant method for the root is the point $(\hat{c}_{r,4}, \hat{c}_{i,4})$ given by

$$\det \begin{vmatrix} \hat{c}_{r,4} - \hat{c}_{r,1} & \hat{c}_{r,4} - \hat{c}_{r,2} & \hat{c}_{r,4} - \hat{c}_{r,3} \\ R_1 & R_2 & R_3 \\ I_1 & I_2 & I_3 \end{vmatrix} = 0$$

$$\det \begin{vmatrix} \hat{c}_{i,4} - \hat{c}_{i,1} & \hat{c}_{i,4} - \hat{c}_{i,2} & \hat{c}_{i,4} - \hat{c}_{i,3} \\ R_1 & R_2 & R_3 \\ I_1 & I_2 & I_3 \end{vmatrix} = 0 \qquad \text{(A1)}$$

where $R_j = R(\hat{c}_{r,j}, \hat{c}_{i,j})$ and $I_j = I(\hat{c}_{r,j}, \hat{c}_{i,j})$. Geometrically, the secant method is equivalent to finding the point of intersection $(\hat{c}_{r,4}, \hat{c}_{i,4}, 0)$ of three planes: the plane through the points $(\hat{c}_{r,j}, \hat{c}_{i,j}, R_j)$ $(j = 1, 2, 3)$; the plane through $(\hat{c}_{r,j}, \hat{c}_{i,j}, I_j)$ $(j = 1, 2, 3)$; and the $(\hat{c}_r, \hat{c}_i)$ plane. The relation (A1) defining $(\hat{c}_{r,4}, \hat{c}_{i,4})$ is the same as that derived by *Acton* [1970] for the two-dimensional false position method.

In practice, we calculate our new root estimate $(\hat{c}_{r,4}, \hat{c}_{i,4})$ from

$$\hat{c}_{r,4} = \frac{a_1\hat{c}_{r,1} + a_2\hat{c}_{r,2} + a_3\hat{c}_{r,3}}{a_1 + a_2 + a_3}$$

and

$$\hat{c}_{i,4} = \frac{a_1\hat{c}_{i,1} + a_2\hat{c}_{i,2} + a_3\hat{c}_{i,3}}{a_1 + a_2 + a_3}$$

where $a_1 = R_2I_3 - R_3I_2$, $a_2 = R_3I_1 - R_1I_3$, and $a_3 = R_1I_2 - R_2I_1$. Next, we discard $(\hat{c}_{r,1}, \hat{c}_{i,1})$ and repeat the secant method with $(\hat{c}_{r,2}, \hat{c}_{i,2})$, $(\hat{c}_{r,3}, \hat{c}_{i,3})$, and $(\hat{c}_{r,4}, \hat{c}_{i,4})$ as our initial estimates for the root. We repeat this procedure until the satisfaction of the convergence criterion

$$\frac{||(\hat{c}_{r,N+3}, \hat{c}_{i,N+3}) - (\hat{c}_{r,N}, \hat{c}_{i,N})||_2}{||(\hat{c}_{r,N}, \hat{c}_{i,N})||_2} < 10^{-6} \qquad \text{(A2)}$$

shows that $(\hat{c}_{r,N+3}, \hat{c}_{i,N+3})$ is a sufficiently good estimate for a root.

For example, using single precision arithmetic on the CDC 6600 with $\hat{H} = 4$ and $\hat{k} = 0.7$, we find, with initial estimates, three corners of a 0.05×0.05 rectangle in $(\hat{c}_r, \hat{c}_i)$ space for which neither R nor I have values all of the same sign on the four corners, that after usually eight iterations, $(\hat{c}_{r,11}, \hat{c}_{i,11})$ satisfies condition (A2). Then the residuals are such that $R(\hat{c}_{r,11}, \hat{c}_{i,11}) = O(10^{-14})$ and $I(\hat{c}_{r,11}, \hat{c}_{i,11}) = O(10^{-14})$ (where R and I typically have values $O(1)$ away from roots). It turns out that if we take Re $(\hat{n}_2) > 0$, then R and I can become very large (i.e., $> 10^{99}$) away from roots because of the term exp $(2\hat{n}_2\hat{H})$ in (12). Even though the same set of roots is obtained no matter what square root of $\hat{n}_2^2$ we take, convergence of the secant method is much faster if we choose Re $(\hat{n}_2) < 0$ rather than Re $(\hat{n}_2) > 0$.

Even if the initial estimates $(\hat{c}_{r,1}, \hat{c}_{i,1})$, $(\hat{c}_{r,2}, \hat{c}_{i,2})$, and $(\hat{c}_{r,3}, \hat{c}_{i,3})$ do not bound a root, the secant method usually converges to a root somewhere else. This is an advantage of the secant method over the false position method.

Acknowledgment. This work was supported by grant GA 33990X from the National Science Foundation. R. Lindzen acknowledges support from an Alfred P. Sloan Foundation fellowship. A. Rosenthal acknowledges support from an Imperial Oil graduate research fellowship.

REFERENCES

Acton, F. S., *Numerical Methods that Work*, Harper and Row, New York, 1970.

Booker, J. R., and F. P. Bretherton, The critical layer for internal gravity waves in a shear flow, *J. Fluid Mech., 27*, 513–539, 1967.

Gossard, E. E., Dynamic stability of an isentropic shear layer in a statically stable medium, *J. Atmos. Sci., 31*, 483–492, 1974.

Gossard, E. E., J. H. Richter, and D. R. Jensen, Effect of wind shear on atmospheric wave instabilities revealed by FM/CW radar observations, *Boundary Layer Meteorol., 4*, 113–131, 1973.

Hardy, K. N., R. J. Reed, and G. K. Mather, Observation of Kelvin-Helmholtz billows and their mesoscale environment by radar, instrumented aircraft, and a dense radiosonde network, *Quart. J. Roy. Meteorol. Soc., 99*, 279–293, 1973.

Hooke, W. H., F. F. Hall, Jr., and E. E. Gossard, Observed generation of an atmospheric gravity wave by shear instability in the mean flow of the planetary boundary layer, *Boundary Layer Meteorol., 5*, 29–41, 1973.

Howard, L. N., Note on a paper of John W. Miles, *J. Fluid Mech., 10*, 509–512, 1961.

Jones, W. L., Reflexion and stability of waves in stably stratified fluids with shear flow: A numerical study, *J. Fluid Mech., 34*, 609–624, 1968.

Lindzen, R. S., Some speculations on the roles of critical level interactions between internal gravity waves and mean flows, in *Acoustic Gravity Waves in the Atmosphere*, edited by F. M. Georges, U.S. Government Printing Office, Washington, D. C., 1968.

Lindzen, R. S., Stability of a Helmholtz velocity profile in a continuously stratified, infinite Boussinesq fluid—Applications to clear air turbulence, *J. Atmos. Sci., 31*, 1507–1514, 1974.

Miles, J. W., and L. N. Howard, Note on a heterogeneous shear flow, *J. Fluid Mech., 20,* 331–336, 1964.
Ottersten, H., K. R. Hardy, and C. G. Little, Radar and sodar probing of waves and turbulence in statically stable clear-air layers, *Boundary Layer Meteorol., 4,* 47–89, 1973.
Reed, R. J., and K. R. Hardy, A case study of persistent, intense clear air turbulence in an upper level frontal zone, *J. Appl. Meteorol., 11,* 541–549, 1972.

(Received August 22, 1975;
revised November 24, 1975;
accepted December 11, 1975.)