

Wave Transience in a Compressible Atmosphere. Part III: The Saturation of Internal Gravity Waves in the Mesosphere¹

TIMOTHY J. DUNKERTON²

Department of Atmospheric Sciences, University of Washington, Seattle 98195

(Manuscript received 1 September 1981, in final form 17 November 1981)

ABSTRACT

It is shown that the analytic transient internal gravity wave solutions derived by Dunkerton (1981a,b) remain qualitatively unchanged when a "saturation hypothesis" is included in the analysis. Furthermore, the wave action flux in the saturated region is *not* constant in height, and experiences considerable falloff as the critical level is approached. Wave transience would appear to lower the level of wavebreaking on the order of a scale height.

It is also shown that these analytic solutions allow shock formation at the trailing edge of the wave packet, for both Boussinesq and atmospheric cases. An "equal-area" rule may be used to determine the position of both internal and trailing shocks. Saturation leads to a net mean flow change slightly different from that of the "nonsaturated" solution.

1. Introduction

In a recent two-part series (Dunkerton, 1981a,b) the author presented some analytic solutions for transient, conservative internal gravity wave, mean-flow interaction. In the first paper (hereafter I) a problem was posed involving a vertically propagating, linear, slowly-varying locally Boussinesq wave of discrete frequency and zonal wavenumber forced at a corrugated lower boundary. After switch-on of the forcing, this wave was found to induce a mean flow in the fluid above the boundary, despite the motion being formally completely conservative. These accelerations were due to wave transience, described by the nondimensional wave action equation

$$\frac{\partial A}{\partial t} + \left(\frac{\partial}{\partial z} - 1 \right) A(1 - A)^2 = 0, \quad (1.1)$$

in which dimensional height and time scales are H and NH/kc^2 , respectively, where k , N , H , and c are zonal wavenumber, static stability, density scale height, and wave phase speed; the scale for the wave action and mean flow is taken to be c , the wave phase speed. The form of Eq. (1.1) is itself a "traditional" result, in which wave action is conserved in terms of a local rate of change in wave action density A (i.e., wave transience) and a mass-weighted vertical derivative of wave action flux B , which in the slowly-varying approximation is equal to the product of A

and the vertical group velocity $(1 - \bar{u})^2$. However, (1.1) contains an essential modification of the traditional equation, viz., the mean flow \bar{u} is set equal to A as demanded by Kelvin's circulation theorem applied to this problem. In other words, mean flow changes are incorporated in an otherwise linear model, thereby producing what we refer to as a "quasi-linear" model.

The method of characteristics yields solutions to (1.1) as described in I. In all cases investigated there the solution took the form of an equilibrated region at low levels, a wavefront region propagating diagonally upward in the time-height cross section, and a nonunique region in which these solutions tended to overlap. An example of solution is shown in Fig. 1. The physical cause of the nonunique region due to the overlapping of characteristic curves arose from the inability of the atmosphere to absorb, as it were, momentum at the same rate it is being supplied at the lower boundary by the radiation stress, or "wave drag," there. This inability in turn arose from the combined effect of the density stratification and the self-reduction in the wave's vertical group velocity as the mean flow is accelerated. In I it was suggested that a "shock condition" may provide a unique solution in the overlapping region; this condition would satisfy the overall momentum budget in which the vertically-integrated mean momentum changes at a rate equal to the lower boundary momentum flux.

While (1.1) provides a powerful quasi-linear means of investigating gravity wave, mean-flow interaction, it has recently become clear that at least two modifications of this result may be necessary in order to describe this interaction more completely. First, in-

¹ Contribution No. 622, Department of Atmospheric Sciences, University of Washington, Seattle.

² Present affiliation: National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.

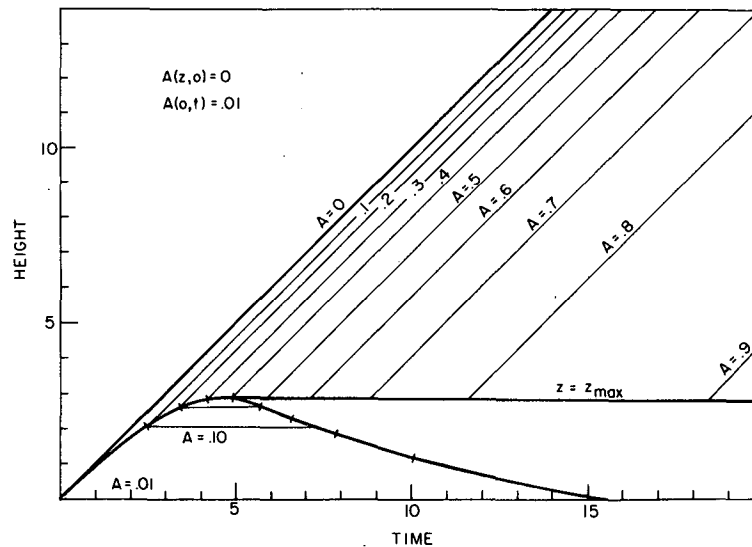


FIG. 1. Analytic solution for the wave action density and mean flow in the "non-saturated" $A_B = 0.01$ case as in Dunkerton (1981a).

spection of the radiating solution in examples such as Fig. 1 reveals that after some time it becomes convectively unstable. In other words, unstable lapse rates are implied in the temperature field, together with a tendency for the perturbation horizontal velocity to exceed the phase speed. Consequently, one would like to ascertain how the elementary saturation hypothesis of Lindzen (1981) would alter these analytic solutions. This hypothesis is consistent with previous notions of "breaking gravity waves" in the upper atmosphere, some of the relevant literature being Pitteway and Hines (1963); Leovy (1964); Lindzen (1967, 1968, 1971, 1981); Hodges (1969); Lindzen and Blake (1971); and Weinstock (1976). The saturation hypothesis proposed by Lindzen (1981) states, quite simply, that above the level of implied convective instability further wave growth (due to *either* density or mean wind shear) is exactly canceled by eddy diffusion arising from the convective instabilities themselves. This is sometimes referred to as the "amplitude balance model." Lindzen's hypothesis suggests that in the "saturated" region the perturbation lapse rate amplitude may *a priori* be set equal to the basic state lapse rate. In Section 2, we will find that this hypothesis, when applied to the analytic solutions of I, leads to a new analytic solution in the saturated region, provided that the effect of wave-induced eddy diffusion *on the mean flow* is neglected. Because the qualitative discussion of I, Section 2a, remains relevant, within the approximations used here, the saturated solutions will qualitatively resemble their "nonsaturated" counterparts (e.g., Fig. 1), and, in particular, the internal shock remains a feature of these solutions.

A second modification of Eq. (1.1), which will not be further explored in this paper (except qualita-

tively) arises from a recent study by Coy (1982) who has simulated some examples of gravity wave, mean-flow interaction in an atmosphere. Coy's results indicate the presence of an "accelerated" wave phase speed arising from what is sometimes referred to as conservation of wavenumber (Whitham, 1974). This effect was included, for example, in Grimshaw's (1975) analytic study; it was not, however, incorporated in I. As a result, the latter paper was able to derive the internal shock result *via* the more analytically tractable (1.1). While Coy's simulations prove conclusively that the "acceleration effect" is not negligible in these pure-transience examples, it is still apparent that internal shock formation is a correct feature even in the "accelerated" cases. The reason for this is that the argument laid out in I, Section 2a, remains relevant, because vertical group velocity is relevant; the difference is that the wave phase speed c is now a variable in the equations of motion.

The purpose of this article is to investigate the effect of the saturation hypothesis on the analytic "non-accelerated" solutions of I. Next, in Section 3 it is shown that for a realistic winter middle atmosphere wind profile, mean flow accelerations described by the quasi-linear model might lead to wavebreaking at significantly lower levels than predicted by Lindzen's (1981) linear model. The effect of wave-induced diffusion *on the mean flow* generally appears relatively small (Lindzen, 1981).

Finally, this paper seeks to further explore the temporal development of the mean flow by allowing the lower boundary forcing to be turned off at some reasonably long time after switch-on. In Section 4 it is shown that formation of a "trailing shock" is implied by (1.1) at the rear of the vertically-propagating

wave packet. Conservation of momentum is used to determine the position of the internal and trailing shocks by an "equal-area" rule analogous to a similar rule in the theory of weak acoustic shocks (Lighthill, 1978, Fig. 42).

A noteworthy result of Section 4 is that in the "nonsaturated" initially-at-rest atmospheric case the trailing shock appears to asymptote to a finite height. This indicates that while transience effects are formally temporary, even in nonlinear wave, mean-flow interaction (Andrews and McIntyre, 1978), transience can nevertheless lead to permanent mean flows if a restriction is placed on the vertical group velocity. In the present case, the vertical group velocity vanishes when the mean flow equals the forced wave phase speed [whereas in Coy's (1982) problem it is the accelerated, internal phase speed which is relevant]. In both cases, one might say that "effective critical layer absorption" brings about the permanent mean flow change (Bretherton, 1966).

This *mathematical* result does not depend on saturation, although the latter becomes physically important above the level of wavebreaking. Because saturation is assumed irreversible, and thus implicitly includes a departure from conservative motion, it is evident that saturation *per se* can also lead to permanent mean flow changes apart from the effective critical layer absorption. This causes the final, permanent mean flow to be somewhat different in the saturated case than in the "nonsaturated" solution, although the latter turns out to be a fair approximation for reasonably long forcings. In the saturated case, for such forcings, the final mean flow includes a "residual" shock left at the lowest level of saturation, while the trailing shock asymptotes to infinity, decreasing in amplitude so as to trace out a continuous final mean flow above this level. This satisfies the constraint that the final mean flow in both the "nonsaturated" and saturated solutions has the same total momentum.

2. The saturation hypothesis

Obviously the amplitude of vertically propagating internal gravity waves cannot become indefinitely large. Actual limits on the growth of the associated velocity and temperature fields are governed by several considerations. For example, the zonal perturbation velocity cannot realistically be expected to ever exceed the phase speed of the wave. Also, the occurrence of a Richardson number below $\frac{1}{4}$ may be expected with large wave amplitude, thereby implying the possibility of shear instability. Simplest of all is that the internal wave growth is limited by the occurrence of an unstable lapse rate. When there is exponential growth of the wave with height, without significant mean wind shear in between, these three limits may occur at or very close to one another;

hence for simplicity we may confine attention to the third, convective instability constraint following Lindzen (1981).

To determine the point of this unstable convective breakdown the wave action density (Dunkerton, 1981a) and vertical wavenumber are

$$A = \frac{1}{2} \frac{\overline{u'^2} + \overline{\phi_z'^2} N^{-2}}{c - \bar{u}} = \frac{\overline{\phi_z'^2} N^{-2}}{c - \bar{u}}, \quad (2.1)$$

$$m = \frac{N}{c - \bar{u}}, \quad (2.2)$$

where u' and ϕ' are zonal velocity and geopotential, the overbar is a zonal average, equipartition is used in (2.1), and we neglect any transverse wavenumber component. These relations are valid at lowest-order for linear, slowly-varying waves; the breaking wave-induced eddy diffusion is formally first order in the inverse square-root Richardson number $\mu \equiv |\bar{u}_z|/N$ as it is required to balance density changes and mean wind shear, both of which are $O(\mu)$. The transverse wavenumber component would increase the vertical wavenumber causing instability somewhat earlier, but again an exponential variation in height implies only a small vertical distance between these points. In the following the transverse wavenumber component is neglected although it may be significant (Lindzen, 1981). The point of unstable breakdown is where the *total* lapse rate becomes unstable somewhere within the wave, where for hydrostatic waves

$$|\phi_{zz}| = N^2 \quad (\phi' = \text{Re} \phi \exp i k x). \quad (2.3)$$

Noting that

$$\overline{\phi'^2} = \frac{1}{2} |\phi|^2, \quad (2.4)$$

the wave action density at the point of unstable breakdown is

$$A = \frac{1}{2} \frac{m^{-2} N^{-2}}{c - \bar{u}} |\phi_{zz}|^2 = \frac{1}{2} (c - \bar{u}). \quad (2.5)$$

The nondimensional saturated wave action density is therefore

$$A_s = \frac{1}{2} (1 - \bar{u}). \quad (2.6)$$

From now on, unless otherwise indicated, all quantities are nondimensional in accordance with the scalings quoted earlier.

The saturation hypothesis states that above the level of wave saturation, further wave growth is exactly canceled by the diffusion generated by the convective instabilities. Expressed in terms of the wave action density, then, the wave amplitude in the saturated region is just A_s . We see that there is no one universal value of A_s as it depends upon the mean flow in the saturated region.

Remarkably, the occurrence of saturation in the quasi-linear model coincides with the appearance of the nonunique region when there is no transverse

wavenumber component. When there is no initial mean flow, $\bar{u} = A$ before saturation and

$$A_s = 1/3, \quad (2.7)$$

which clearly marks the onset of the nonunique region in those solutions having no initial mean flow (e.g., Fig. 1). A similar result is implied even with a mean flow initially. Had we not taken into account the wave-induced mean flow, the saturated wave action density would have been

$$A_s = \frac{1}{2}(1 - \bar{u}_0), \quad (2.8)$$

being equal to $\frac{1}{2}$ without an initial mean flow. Hereafter this will be referred to as the "linear" model.

Now because the analytic solutions of I predict values of A in excess of A_s , this would seem to invalidate those solutions if saturation is assumed. However, it can be shown that *the qualitative features of the mean flow evolution remain the same even if saturation is assumed*, provided that we first reinterpret the contours in the time-height cross sections of I as denoting mean flow, and not wave action density. This demonstration rests on the quasi-linear result at lowest order in the inverse square-root Richardson number that

$$B = \frac{1}{2}(1 - \bar{u})^3, \quad (2.9)$$

viz., that the wave action flux in the saturated region may be assumed equal to the product of the wave action density A , and the usual vertical group velocity (Dunkerton, 1981a). Importantly, this flux is not constant in height (Weinstock, 1976), and is not to be set equal to its value at the level of wavebreaking. The latter was suggested by Lindzen (1981) as a first approximation. Eq. (2.9) actually indicates a considerable falloff of B as the critical level is approached. This is due to two factors: first, the vertical wavenumber becomes infinite at the critical level and hence convective instability is prevented only by an ever-decreasing wave action amplitude; second, the vertical group velocity is decreasing towards the critical level.

Curiously, (2.9) may be used to form a mean flow equation in the saturated region

$$\frac{\partial \bar{u}}{\partial t} + \left(\frac{\partial}{\partial z} - 1 \right) \frac{1}{2} (1 - \bar{u})^3 = 0, \quad (2.10)$$

in which the acceleration is equal to the mass-weighted convergence of wave action flux, the latter being equal to the mass-weighted convergence of momentum flux at lowest order in μ (Dunkerton, 1981a). It is important to note that this result neglects the diffusive component of the acceleration although the latter might become significant at the level of wavebreaking (Lindzen, 1981). For the mo-

ment we neglect this diffusion, but will attempt to assess its importance later.

Equation (2.10) has characteristic curves

$$t + c_1 = \frac{1}{(1 - \bar{u})^2}, \quad (2.11a)$$

$$z + c_2 = \ln(1 - \bar{u})^3. \quad (2.11b)$$

As in I, a universal curve exists apart from two constants used to apply initial and boundary conditions. This curve is entirely different from that found for the wave action density in I, however, and is a negative logarithm, with the mean flow \bar{u} proceeding from large negative values on the positive part of the curve to unity as t becomes infinite. Comparison to Eq. (18a) of I reveals that the approach to critical level conditions is slower in this case.

Keeping matters as simple as possible, let us apply the saturation hypothesis to the example in Fig. 1. We first note that the $A (= \bar{u})$ solution is valid outside $A = \frac{1}{3}$. Both the equilibrated solution and the radiating solution for $A \leq \frac{1}{3}$ remain.

Next, the new characteristic (2.11a,b) may be used to trace out \bar{u} in the saturated region, matching the point $\bar{u} = \frac{1}{3}$ on this curve to the diagonal line of the radiating solution having $A = \frac{1}{3}$. This immediately implies a new set of diagonal lines for \bar{u} (Fig. 2). Also, the nonunique region at first begins to descend along the new characteristic emanating from the first nonunique point. The qualitative feature of the radiating solution would therefore remain in the saturated wave case. But there would be a quantitative difference; the saturated wave solution appears to evolve more slowly towards the critical level. This may be partly attributed to the decelerating influence of the wave transience in the saturated region.

The descending shear zone would also remain, as indeed it must, for the qualitative discussion in I, Section 2a, remains relevant. Judging from the slower approach to critical level conditions the discontinuity in \bar{u} at the shock would need to descend slightly faster to satisfy the momentum budget. (The remark on page 288 of I indicating shear zone descent along the primary characteristic should be disregarded.) A graphical method for determination of the shock position is suggested in Section 4.

Comparison of Figs. 1 and 2 indicates that saturation has a quantitative, but not qualitative, effect on the mean flow evolution. It will be noted, however, that contours of A are very different in the two cases. This result, incidentally, helps explain the numerical standing wave solution in I, Section 3c, in which a crude saturation hypothesis was applied; this solution evolved much like its nonsaturated counterpart but with a slightly weaker amplitude.

Before attempting to assess the role of eddy diffusion in the mean flow acceleration, some additional remarks should be made. First, the variable wave

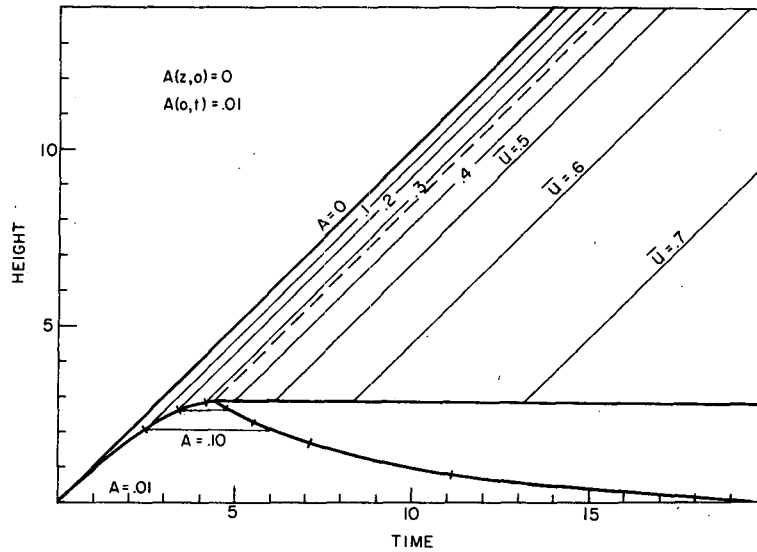


FIG. 2. Analytic solution for the mean flow in the saturated wave case with $A_g = 0.01$.

action flux is crucial to this calculation involving transient waves if for no other reason than wave signals require a finite time to propagate vertically; obviously one cannot set the wave action flux instantaneously equal to its value at the level of wavebreaking. Second, the saturation hypothesis resolves a dilemma in I concerning what to do if the forcing exceeds $1/3$; the answer is that saturation occurs immediately. This does *not* prevent further vertical propagation of the wave, however. This result may be relevant to the radiation of gravity waves from “saturated” Kelvin-Helmholtz instabilities in the troposphere (Davis and Peltier, 1979; Fritts, 1979, 1982a; McIntyre and Weissman, 1978). Finally, it is our belief that saturation may help rationalize the quasi-linear solutions of I to some extent, insofar as the beginning of the nonunique region corresponds to a physical event (*viz.*, the unstable breakdown of the wave). Further growth of the wave is prevented by saturation and this may help keep the linearization approximation viable.

In comparing the non-saturated and saturated solutions of Figs. 1 and 2 it may be objected that eddy diffusion is strongest in the saturated region and is therefore not negligible. We conclude this section by examining whether or not this is so. Now the expression for the eddy diffusion coefficient in the saturated region is

$$D_{\text{eddy}}^* = \Lambda \nu^* [(1 - \bar{u})^3 (1 - \bar{u} + 3\bar{u}_z) + \bar{u}_z (1 - \bar{u})]. \quad (2.12)$$

Here, Lindzen’s (1981) formula is slightly modified to include \bar{u}_z , and

$$\Lambda = \frac{1}{2}(c/NH)^2, \quad (2.13a)$$

$$\nu^* = kc^2H/N, \quad (2.13b)$$

the latter being a transient wave diffusion scale constructed from the height and time scales in this case. This scale forms a non-dimensional eddy diffusion coefficient D_{eddy} when ν^* is divided out of (2.12). From (2.12) we observe that eddy diffusion is likely to be important only in “adverse” circumstances where the waves are able to break at some point well away from the critical level ($\bar{u} < 0$), which is where they are least likely to break (as seen in the next section). As the mean flow evolves towards unity, the diffusion coefficients markedly decrease (Lindzen, 1981). The parameter in (2.13a) is $O(\mu^2)$ in the WKB approximation.

We conclude that eddy diffusion’s effect on the mean flow will be significant only in relatively special circumstances. Even in such cases the net effect will be merely to redistribute momentum over the area where the diffusion is significant (Lindzen, 1981).

An explicit example considered in the next section will help illustrate these remarks.

3. Wavebreaking in linear and quasi-linear models

The saturated-waves model discussed in Lindzen’s (1981) timely article corresponds to what was referred to above as the linear model, as wave-induced mean flow changes are not considered. Here we will make an attempt to discuss how the quasi-linear model differs from the linear wave model insofar as the level of wavebreaking is concerned.

Because the wave transience term in the wave action equation determines the *time* of unstable breakdown, but not the *height*, it is sufficient for the present purpose to confine attention to the steady equation

$$\left(\frac{\partial}{\partial z} - 1\right)A(1 - \bar{u})^2 = 0, \quad (3.1)$$

having the solution

$$A(1 - \bar{u})^2 = B_0 \exp z, \quad (3.2)$$

with the subscript denoting the lower boundary forcing.

The linear model with no transverse wavenumber component indicates breakdown at the height determined implicitly by

$$z_s = \ln \left[\frac{1}{2} \frac{(1 - \bar{u}_0)^3}{B_0} \right], \quad (3.3a)$$

while the corresponding quasi-linear breakdown point is

$$z_s = \ln \left[\frac{4}{27} \frac{(1 - \bar{u}_0)^3}{B_0} \right]. \quad (3.3b)$$

(Note that transience *per se* has not been ignored in deriving this result, only the transience term in the wave action equation.) When there is no initial mean flow, the quasi-linear model predicts breakdown at $\ln 27/8 (\approx 1.22)$ scale heights below the point of steady wave breakdown. In general, for waves propagating toward the critical level the difference in breaking levels is less than this amount, and *vice versa*.

In the middle atmosphere this difference in breaking levels is complicated by the fact that the waves may initially propagate away from the critical level as they enter the stratosphere while returning towards it in the mesosphere.³ As an example of this behavior consider an initial mean flow of the form

$$\bar{u} = \sum_{i=1}^3 u_i \exp - \left(\frac{z - z_i}{a_i} \right)^2, \quad (3.4)$$

where

$$u_i = (-1.1, -3.8, 1.76), \quad (3.5a)$$

$$z_i = (1.7, 8.0, 11.7), \quad (3.5b)$$

$$a_i = (0.8, 2.9, 1.6). \quad (3.5c)$$

This profile is a fit to the vertical wind profile in Fig. 2 of Lindzen (1981) at 45° in January, where we have taken the wave to be a topographically-generated disturbance having an intrinsic phase speed in the source region of -15 m s^{-1} . We also take $H = 7 \text{ km}$ for simplicity. Because the mean flow is normalized by the phase speed it should be remembered that smaller phase speeds would amplify the mean flow in Fig. 3a, and *vice versa*.

In Fig. 3c the height of unstable breakdown is indicated for the linear and quasi-linear models. We note, first, that the difference in breaking levels is quite small in the "mesosphere" ($z > 8$). Second, it is possible for a quasi-linear wave to initially break at a much lower level than the linear wave over two

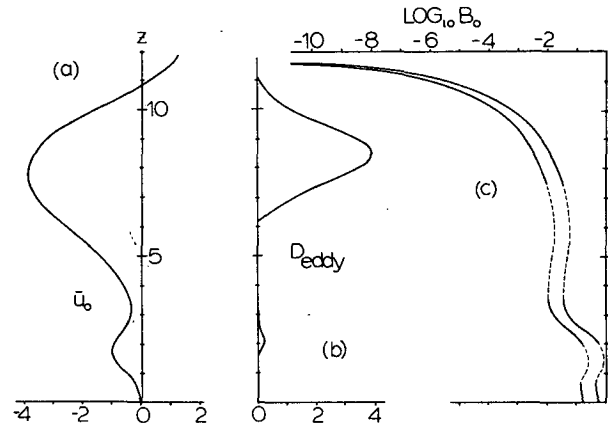


FIG. 3. A simple example of vertical propagation of gravity waves into the mesosphere (see text for details). (a) Mean flow having a critical level at $z = 11.66$ scale heights; (b) maximum possible eddy diffusion for this case (note that the diffusion is zero below the wavebreaking level and this figure does not account for accelerations toward the critical level); (c) wavebreaking level as a function of the wave action flux at the lower boundary (B_0), for linear (right) and quasi-linear (left) cases.

narrow ranges of B_0 . Neither of these waves is able to initially break beneath both westerly jets, where they are propagating away from the critical level at such a rate that their exponential density-induced growth is more than canceled. For adequately small forcing in this wind profile, gravity waves might literally disappear in the "stratosphere" ($3 < z < 8$) before reappearing in the "mesosphere."

In Fig. 3b is shown the maximum possible eddy diffusion coefficient (excluding near a shock) implied by (2.12) with $N = 0.02 \text{ s}^{-1}$. (D_{eddy} is zero below the wavebreaking level though all possible values are shown here.) The eddy diffusion may be large, but only where the waves are least likely to break. For values of B_0 below 10^{-4} the diffusion will be a small effect above the initial level of wavebreaking. Actual values of the diffusion will be considerably less depending on the acceleration of the mean flow towards the critical level. These remarks agree with Lindzen's (1981) statement that saturated wave stresses cannot be balanced by the eddy diffusion.

The temporal evolution of the quasi-linear solution (not shown) indicates that the front of the wave packet follows the linear group velocity trajectory

$$\left. \frac{dz}{dt} \right|_{A=0} = (1 - \bar{u}_0)^2, \quad (3.6)$$

slanting up at an angle greater than 45° in the time-height cross section when $\bar{u}_0 < 0$. The "primary characteristic" associated with the switched-on forcing behaves similarly but gradually falls behind the wavefront until peaking at the point of breakdown. As the critical level is approached in height the wavefront flattens out as in Appendix A of I. The deposition of momentum occurs mostly between z_s and the critical level, with some initial acceleration form-

³ This is reminiscent of the dynamical effect in the quasi-biennial oscillation (Matsuno, 1982).

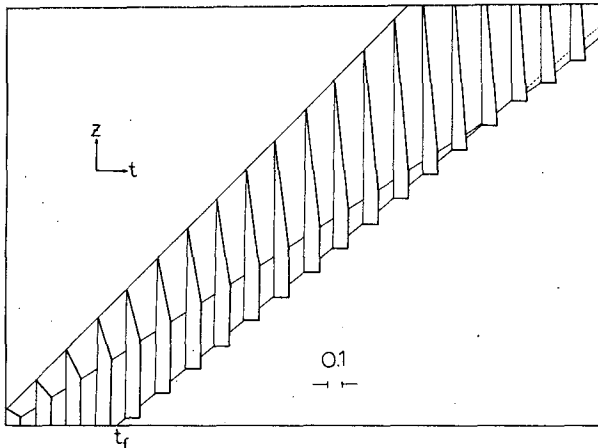


FIG. 4. Vertical propagation in the Boussinesq case with forcing turned off at time t_f . Features include the tilted wavefront amplitude, trailing shock, and a point where the rear of the wavefront and trailing shock merge.

ing the equilibrated solution beneath z_s . If the forcing remains on long enough, z_s may descend in the form of a “shock,” or strong shear zone in \bar{u} .

While this example is intended to be illustrative, it may also be of general importance to the middle atmosphere vertical propagation of gravity waves at the solstices. In conclusion it should be pointed out that an aspect of this problem not treated here concerns the effect of an external forcing of the mean flow, such as is associated with the mean Coriolis accelerations at the solstices due to the diabatic heating. Because these accelerations may be comparable to the gravity wave accelerations (Lindzen, 1981), it would be of interest to solve the initial value problem in which such an external forcing (as expressed by a relaxation to radiative equilibrium) has equal *a priori* emphasis (Holton, 1982).

4. Trailing shock formation, the “equal-area” rule, and final mean flows

The transient wave solutions of I, together with the saturated version of these solutions presented here, have another interesting property when the wave forcing ceases at some time t_f after switch-on. Because small-amplitude signals overcome large-amplitude ones there is a nonunique region produced once the forcing is turned off. In fact this is also a feature of the initially-at-rest Boussinesq case for which characteristics are (Dunkerton, 1981a)

$$\left. \frac{dz}{dt} \right|_A = (1 - A)(1 - 3A). \quad (4.1)$$

Obviously the primary characteristic associated with the constant forcing (A_B) slopes up less than 45° in the time–height cross section implying a nonunique region once the forcing ceases, i.e., when an $A = 0$

characteristic is generated (see Fig. 4). A unique solution, however, is found after time t_f having a shock at $z = (1 - A_B)^2(t - t_f)$. This is because the total transient acceleration in the wavefront is constant in time and is equal to the previously constant forcing. To identically cancel this acceleration we need merely to construct the trailing shock as a slanting line beginning at t_f . Eventually, however, this line intersects the characteristic associated with the initial switch-on; here, the shock departs from its diagonal course, ascending somewhat more slowly.

The “nonsaturated” atmospheric version of this solution exhibits similar features, together with the appearance of the internal shock at the rear of the wavefront if the forcing is left on long enough. A similar remark applies to Boussinesq and atmospheric solutions with the wave propagating towards its critical level as in Appendix A of I. There might appear to be some ambiguity concerning where to place both shocks since the overall momentum budget provides only one constraint. However, there is a graphical method which may be used to determine the shock position which is analogous to the “equal-area rule” employed in the theory of weak acoustic shocks (Lighthill, 1978, Fig. 42). This rule may be used to determine the position of either the internal or trailing shock or both, provided that it is employed separately for each. For example, Fig. 5 shows the *density-weighted* nonsaturated, “non-accelerated,” atmospheric solution of Fig. 1 at $t = 8$ (heavy curve). The area enclosed by this curve equals the total momentum fluxed across the lower boundary up to this time, the latter being $B_0 t = 0.0784$. The internal shock cuts out equal areas in the regions shown in the figure so as to conserve momentum (which, it should be remembered, is always weighted by the mean density e^{-z}). Nicely, this method gives the trailing shock position also, and for both shocks does not require previous knowledge of the solution (except in the saturated case for the asymptotic behavior of the trailing shock, Fig. 6). Supposing that at time t_f the forcing is turned off instantaneously on the slow time scale, the characteristics emanating from $(0, t_f)$ are identical to those of the initial switch-on, and there is now a “wavefront-like” solution beneath the previous equilibrated solution (i.e., there is a new nonunique region). This overlapping can be uniquely resolved by the equal-area rule; in fact, we used this rule implicitly in the simpler Boussinesq problem shown in Fig. 4.

There is, however, a fundamental quantitative difference between the initially-at-rest Boussinesq case (Fig. 4) and its “nonsaturated” atmospheric counterpart. While in the latter case the internal and trailing shocks merge, after this time the trailing shock ascends very slowly. In fact the asymptotic height attained by this shock appears finite, since M , the vertically-integrated mean momentum, is restricted

in this case by the slowly-varying requirement that $\bar{u} < 1$, i.e.,

$$M = \int_{z_f}^{\infty} \bar{u}(z)e^{-z} dz < e^{-z_f}, \quad (4.2)$$

where z_f is the asymptotic height of the trailing shock. But

$$M = \int_0^{t_f} B_0(t) dt = B_0 t_f, \quad (4.3)$$

if B_0 is constant in this interval; hence

$$z_f = -\ln B_0 t_f \quad (4.4)$$

in this case.

This remarkable result illustrates the “effective critical layer absorption” discussed by Bretherton (1966). Because vertical group velocity vanishes at $\bar{u} = 1$ there is an effective absorption of the wave insofar as ray paths become horizontal there. Interestingly this mathematical result does not depend on any *real* wave absorption (e.g., saturation) although the latter becomes important above the wave-breaking level. While effects of transience are formally temporary, even in nonlinear wave, mean-flow interaction (Andrews and McIntyre, 1978), as for example was previously illustrated by Uryu (1974) in

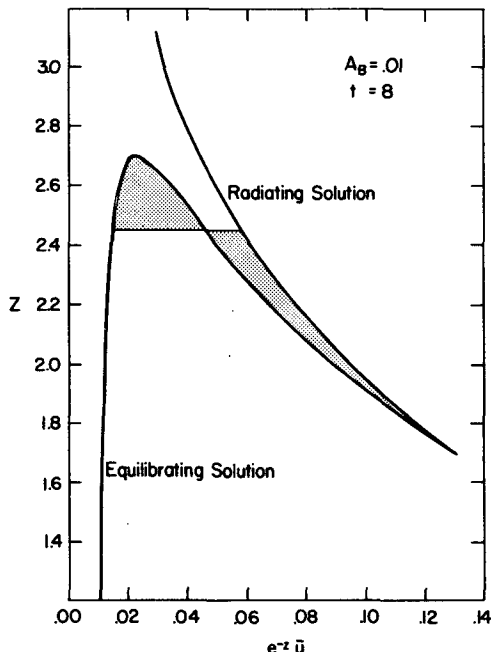


FIG. 5. The equal-area rule, nonsaturated case. The area enclosed by the heavy line equals the total momentum fluxed across the lower boundary up to time $t = 8$. A unique solution with discontinuous shock cuts out equal areas on both sides of the shock. This is analogous to the weak shock in acoustic waves (Lighthill, 1978, Fig. 42) although we are here discussing momentum and not mass, and (formally at least) z is a “slow” height. The saturated version of this curve is similar, and the equal-area rule again applies. The trailing shock position is determined in the same way.

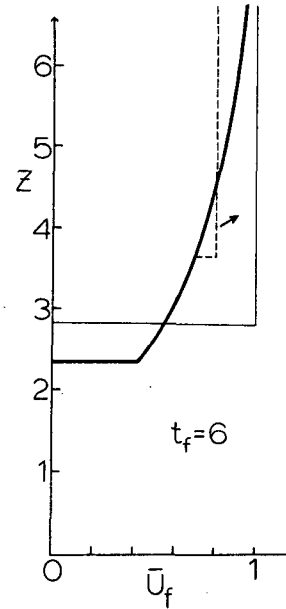


FIG. 6. Approximate final mean flow for the saturated initially-at-rest atmospheric case (heavy line). Also shown is the “non-saturated” step function with $z_f = 2.83$ (light line). The trailing shock traces out the final mean flow asymptotically (dashed line).

a linear Rossby wave context, our example indicates that the concept of “transience” may itself be governed by restrictions placed on the vertical group velocity.

The equal-area rule also applies in the saturated wave case, the difference being that the heavy curve in Fig. 5 descends more steeply below the initial level of saturation (where it turns over), as it is saturated here (and in the radiating solution above, up to $A = 1/3$), as was previously illustrated in Fig. 2. Because there is some real absorption in the saturated case, it is possible to attain a permanent mean flow below the level predicted by (4.4). Above the lowest level of saturation this final mean flow implied by (2.6) is

$$\bar{u}_f = \frac{3}{2}\bar{u}_- - \frac{1}{2}, \quad (4.5)$$

where \bar{u}_- is the mean flow just ahead of the trailing edge of the wave packet. Below the lowest level of saturation, there is no permanent mean flow change due to purely transient, conservative waves (Eliassen and Palm, 1960; Charney and Drazin, 1961; Andrews and McIntyre, 1978).

An important corollary of this result is that for waves breaking only in the mesosphere, the net influence of these waves in the stratosphere is zero when the motion is conservative. Section 3 has given some indication of wavebreaking levels in a specific case.

After merger of the internal and trailing shocks in the saturated solution, the asymptotic behavior of the trailing shock seems to require that we know the

history of the shock position. Eq. (4.5) may be used to determine the amplitude of the "residual" shock left at the lowest level of saturation. This equation, together with conservation of momentum, then gives the asymptotic behavior. For the purposes of illustration an analytic method may be suitable, however, as follows: Assume that the lowest level of saturation is known, and assume that the forcing is of adequately long duration [relative, say, to the time $t_{\text{MAX}} - z_{\text{MAX}}$, where $(z_{\text{MAX}}, t_{\text{MAX}})$ denotes the initial point of saturation]. Let the area enclosed by the solution above the trailing shock (at z_i) be \mathcal{A} , where \mathcal{A} is mass-weighted as

$$\mathcal{A} = \int_0^{\bar{u}-(z_i)} \int_{z_i}^{z(\bar{u})} e^{-z} dz d\bar{u}. \quad (4.6)$$

Although it is possible to actually solve for \mathcal{A} , this leads to a cumbersome asymptotic equation for z_i ; let us replace \mathcal{A} with the approximate result

$$\mathcal{A} \approx \bar{u}_-(z_i) e^{-z_i}, \quad (4.7)$$

i.e., that the mean flow above z_i is constant and equal to its value at z_i . In the atmospheric case this is actually a good approximation since density at upper levels is exponentially small. Now the mass-weighted area \mathcal{A} decreases with time because a permanent mean flow is being formed above the lowest level of saturation. This incremental change in mass-weighted area is

$$d\mathcal{A} = -\bar{u}_f e^{-z_i} dz_i. \quad (4.8)$$

Eqs. (4.5), (4.7), and (4.8) imply the approximate permanent mean flow

$$z_i - z_{\text{sat}} = \ln \left[\frac{1 - \bar{u}_f(z_{\text{sat}})}{1 - \bar{u}_f(z_i)} \right]^2, \quad (4.9)$$

where z_{sat} is the lowest level of saturation. Fig. 6 displays a representative case for $A_B = 0.01$ and $t_f = 6$, together with a comparison to the "nonsaturated" final mean flow. Also indicated is how the trailing shock temporally traces out the permanent mean flow.

In conclusion, while we have used the term "critical layer" in conjunction with some of the above results, one should not forget that within the context of the slowly-varying theory no actual critical layer is implied if one does not exist initially. It is of interest to note that the same result is implied in Coy's (1982) "accelerated" phase speed problem insofar as the internal phase speed implied by conservation of wavenumber still defines a vertical group velocity within the WKB approximation; therefore, starting from rest the analytic model cannot generate a *true* critical layer where the internal phase speed matches the mean flow.

An important question is whether a finite Richardson number would alter this argument; critical layer formation starting from rest appears to be a distinct possibility in this case.

Of course, if a critical layer is present initially, this will inhibit further vertical propagation (Bretherton, 1966). A finite Richardson number gives a similar result with a slight propagation through the critical layer (Booker and Bretherton, 1967). Numerical simulation of this, however, remains somewhat elusive despite previous efforts (Jones and Houghton, 1971) and remains a relevant subject for further investigation (Fritts, 1982b).

Acknowledgments. The author received beneficial communication and correspondence from L. Coy, D. C. Fritts and M. E. McIntyre. This research was supported by the National Science Foundation, Atmospheric Research Section, Grant ATM79264-87. Completion of this work was supported by the National Center for Atmospheric Research, which is sponsored by the National Science Foundation.

REFERENCES

- Andrews, D. G., and M. E. McIntyre, 1978: An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.*, **89**, 609-646.
- Booker, J. R., and F. P. Bretherton, 1967: The critical layer for internal gravity waves in shear flow. *J. Fluid Mech.*, **27**, 513-539.
- Bretherton, F. P., 1966: The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Meteor. Soc.*, **92**, 466-480.
- Charney, J. G., and P. G. Drazin, 1961: Propagation of planetary scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83-109.
- Coy, L., 1982: A slowly-varying model of gravity wave, mean flow interaction in a compressible atmosphere. Submitted to *J. Atmos. Sci.*
- Davis, P. A., and W. R. Peltier, 1979: Some characteristics of the Kelvin-Helmholtz and resonant overreflection modes of shear flow instability and their interaction through vortex pairing. *J. Atmos. Sci.*, **36**, 2394-2412.
- Dunkerton, T. J., 1981a: Wave transience in a compressible atmosphere, Part I: Transient internal wave, mean-flow interaction. *J. Atmos. Sci.*, **38**, 281-297.
- , 1981b: Wave transience in a compressible atmosphere, Part II: Transient equatorial waves in the quasi-biennial oscillation. *J. Atmos. Sci.*, **38**, 298-307.
- Eliassen, A., and E. Palm, 1960: On the transfer of energy in stationary mountain waves. *Geophys. Publ.*, **17**, 44 pp.
- Fritts, D. C., 1979: The excitation of radiating waves and Kelvin-Helmholtz instabilities by the gravity wave-critical level interaction. *J. Atmos. Sci.*, **36**, 12-23.
- , 1982a: Shear excitation of atmospheric gravity waves. *J. Atmos. Sci.*, **39** (in press).
- , 1982b: The transient critical-level interaction in a Boussinesq fluid. *J. Geophys. Res.* (in press).
- Grimshaw, R., 1975: Nonlinear internal gravity waves and their interaction with the mean wind. *J. Atmos. Sci.*, **32**, 1779-1793.
- Hodges, R. R., Jr., 1969: Eddy diffusion coefficients due to in-

- stabilities in internal gravity waves. *J. Geophys. Res.*, **74**, 4087-4090.
- Holton, J. R., 1982: The role of gravity wave-induced drag and diffusion in the momentum budget of the mesosphere. *J. Atmos. Sci.*, **39**, 791-799.
- Jones, W. L., and D. D. Houghton, 1971: The coupling of momentum between internal gravity waves and mean flow: A numerical study. *J. Atmos. Sci.*, **28**, 604-608.
- Leovy, C. B., 1964: Simple models of thermally driven mesospheric circulations. *J. Atmos. Sci.*, **21**, 327-341.
- Lighthill, M. J., 1978: *Waves in Fluids*. Cambridge University Press.
- Lindzen, R. S., 1967: Thermally driven diurnal tide in the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **93**, 18-42.
- , 1968: The application of classical atmospheric tidal theory. *Proc. Roy. Soc. London*, **A303**, 299-316.
- , 1971: Tides and gravity waves in the upper atmosphere. *Mesospheric Models and Related Experiments*, J. Fiocco, Ed., D. Reidel.
- , 1981: Turbulence and stress due to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707-9714.
- , and D. Blake, 1971: Internal gravity waves in atmospheres with realistic dissipation and temperature: Part II. Thermal tides excited below the mesopause. *Geophys. Fluid Dyn.*, **2**, 31-61.
- Matsuno, T., 1982: A quasi one-dimensional model of the middle atmosphere interacting with internal gravity waves. Submitted to *J. Meteor. Soc. Japan*.
- McIntyre, M. E., and M. A. Weissman, 1978: On radiating instabilities and resonant overreflection. *J. Atmos. Sci.*, **35**, 1190-1196.
- Pitteway, M. L. V., and C. O. Hines, 1963: The viscous damping of atmospheric gravity waves. *Can. J. Phys.*, **41**, 1935-1948.
- Uryu, M., 1974: Mean zonal flows induced by a vertically propagating Rossby wave packet. *J. Meteor. Soc. Japan*, **52**, 481-490.
- Weinstock, J., 1976: Nonlinear theory of acoustic gravity waves, 1. Saturation and enhanced diffusion. *J. Geophys. Res.*, **81**, 633-652.
- Whitham, G. B., 1974: *Linear and Nonlinear Waves*. Academic Press.