

# THE SATURATION OF MIDDLE-ATMOSPHERE GRAVITY WAVES

Colin O. Hines

Arecibo Observatory, P.O. Box 995, Arecibo, Puerto Rico 00613

## 1. INTRODUCTION

A broad spectrum of gravity waves propagates through the middle atmosphere, analogous to that found in the oceans. As in the oceans, this spectrum exhibits a degree of universality over a decade or more in vertical wavenumber  $m$ : the wind power spectral density is approximately of the form  $KN_0^2 m^{-3}$  over the universal range, where  $N_0$  is the unperturbed buoyancy frequency and  $K$  is a constant variously taken to lie between 0.1 and 0.5. (In this paper, frequencies and wavenumbers are to be measured in radians per second and per meter, respectively, though it matters not for this particular purpose.) This "saturated" portion of the spectrum is found irrespective of meteorological conditions, time, place and height (Balsley and Carter 1982; Dewan et al. 1984; Tsuda et al. 1989; VanZandt 1982; Vincent 1984). The cause of saturation in the atmosphere is still a matter for debate, as is that in the oceans; the two may have — indeed, are likely to have — similar origins and differences of detail only.

The most frequently employed theory of saturation in the atmosphere attributes it to linear instability (Dewan and Good 1986, henceforth DG86; Smith et al. 1987, henceforth SFV87), though a recent paper by Weinstock (1990) makes an important challenge based on the diffusive dissipation imposed by nonlinear combination of the smaller-scale waves. The present paper outlines yet another mechanism: wave-wave interaction imposed by the advective nonlinearity of the Eulerian fluid-dynamic relations. It adapts and summarizes a three-part set of papers recently accepted, and a fourth part recently submitted, for publication (Hines, 1991a,b,c,d, henceforth H91a,b,c,d, respectively). The three parts correspond to Sections 2, 3 and 4 respectively, with the fourth part included in Section 3, and an over-all discussion is presented in Section 5.

To set the stage for oceanographers, I should remark here on two important differences between the atmospheric and oceanic cases: (1) The wave spectrum in the middle atmosphere is believed to be dominated by waves propagating their energy upward from sources at lower levels (such as winds blowing over mountains, moving cold fronts, shear in jet streams, and tropospheric convection that penetrates to or through the tropopause). Partial internal reflections may occur in the middle atmosphere, but there is no strongly reflecting single level analogous to the surface of the ocean. Instead, the upper levels act primarily as a dissipative region, the effective diffusion coefficients increasing with height  $z$ . (2) Wave amplitudes tend also to increase with  $z$ , in response to the decrease of gas density. (The amplitudes

tend to grow as  $\exp z/2H$  while the density decreases as  $\exp -z/H$ , where the scale height  $H$  is about 7 km.) In consequence, the wave-induced wind variance would be expected to grow by some five or six orders of magnitude on ascending from the tropopause (at heights of 8 - 16 km) to the turbopause (at heights of 100 - 110 km), were it not for the limiting effects of saturation and dissipation; in fact it increases by about three orders of magnitude only, but this still constitutes a wide range. The factor  $K$  in the saturated portion of the spectrum is found to remain relatively constant despite these wide ranges of anticipated and observed intensities, but the saturated range of wavenumber  $m$  is found to shift, with at least the lower bound progressing to smaller  $m$  with increase of height (e.g., SFV87). These features open to the atmosphericist a degree of freedom not available to the oceanographer, one that may inspire new modes of attack on the problem of saturation and one that provides for a testing of any new theory.

## 2. CRITIQUE OF LINEAR-INSTABILITY THEORY

The linear-instability theory of DG86 and SFV87 attributes saturation to instability engendered by the wave spectrum in consequence of the latter's growth with height. The theory is a descendant of an analysis by Hodges (1967) in which a monochromatic, upgoing gravity wave was considered. I shall represent such a wave as having phase variations given by  $[kx + ly - mz - \omega t]$ , with  $z$  the upward coordinate,  $m$  and  $\omega$  positive, and I shall term it a single "mode", there being no overlying reflector to produce a downgoing complement.

Hodges determined the Richardson number  $Ri$  as a function of phase for a single mode under approximations appropriate to much of the atmospheric spectrum. These approximations, which I also adopt, produce the dispersion relation

$$(\omega/h)^2 = (N_o/m)^2 \quad (1)$$

as in an incompressible medium (neglecting Earth's rotation),  $h = + (k^2 + l^2)^{-1/2}$  being the horizontal wavenumber, taken to be  $\ll m$ . He found that the minimum  $Ri$  in a single mode could fall below 1/4 (and so produce dynamic instability) only if it in fact fell below 0 at an appropriate phase (and so produced, at that phase, static — or convective — instability). The required condition for instability could be stated as  $\sigma_s^2 = 0.5$ , where  $\sigma_s$  is the standard deviation of the wave-induced shear, nondimensionalized (here and henceforth) by division by  $N_o$ . At greater heights, the wave amplitude was expected to remain constant (be "saturated", in later parlance), or perhaps even be reduced, in consequence of the transfer of wave energy to turbulence energy.

In DG86, arguments previously presented by Phillips (1977) for ocean waves were adapted to the atmospheric case. The wave-induced winds of the saturated portion of the atmospheric spectrum were taken to be produced by a succession of wave groups (structured in the vertical) having a range of vertical wavenumbers  $\Delta m$  proportional to  $m$ , and these were taken to become unstable — statically or dynamically — when their  $\sigma_s$  reached some critical value taken

## Saturation of Middle-Atmosphere Gravity Waves

to be of order unity. With this critical value independent of  $m$ , DG86 found that the corresponding wind variance must be proportional to  $N_0^2 m^{-2}$ , and so the power spectral density within the  $\Delta m$  group must be proportional to  $N_0^2 m^{-3}$ , the factor of proportionality being assumed to be of order unity. They then attributed this spectral form to the observable saturated wind spectrum as a whole, and thereby provided an explanation for the observations. A second argument by DG86 obtained the same spectral form on dimensional grounds, with the assumption that  $N_0$  represented the only background atmospheric parameter and  $m$  (taken to be  $\gg h$ ) the only spectral parameter relevant to instability: the dimensions of the power spectral density of the winds then required that the density itself be proportional to  $N_0^2 m^{-3}$  as before. The same spectral form was adopted by SFV87, based only on a reference back to DG86, not on any further argument.

The first argument of DG86 (and so that of Phillips) is subject to severe criticism even if one accepts the assumptions that go into it: it in fact requires yet another assumption -- a hidden assumption to date -- that the postulated wave groups must enter the observations with equal frequency of occurrence across the  $m$  spectrum. I see no basis for such an assumption, even if the  $\Delta m \propto m$  assumption might be justified on some scaling grounds (which I doubt), and so I cannot accept that the first argument is relevant. The second, dimensional argument will be valid if its assumptions are, but one of those assumptions is that wave instability is indeed the mechanism that shapes the spectrum: consistency is found, which is a necessary but not a sufficient finding to establish wave instability as the relevant mechanism. The theory of stratified turbulence, for example, as in Lumley (1964), leads to the same spectral form, as does the diffusive wave nonlinearity of Weinstock (1990). Below, I shall argue that the total wind standard deviation  $\sigma_T$  is a relevant parameter, thereby negating the purely dimensional argument, and indeed I shall argue that the  $m^{-3}$  form is only an approximate accident anyway, subject to some change from case to case and plausibly from atmosphere to ocean.

In DG86, the multiplier of  $N_0^2 m^{-3}$  was left unspecified other than that it be roughly of order unity (probably). In SFV87, on the other hand, a critical value of 0.5 was adopted for  $\sigma_s^2$  by analogy with the case of the single, monochromatic wave. This critical value was first applied to a wave group with  $\Delta m = m$  and shown to produce  $N_0^2/2m^3$  as the requisite spectral power density. Subsequently, a model spectrum of the form  $(1 + [m/m_*]^3)^{-1}$  was assumed,  $m_*$  being a characteristic vertical wavenumber determining the transition from an unsaturated (flat) portion of the wind spectrum at smaller  $m$  to the saturated, large- $m$ ,  $m^{-3}$  tail spectrum. With a further assumption as to the length of the tail -- it was taken to extend to about 200  $m_*$  -- and of the form of the frequency spectrum (assumed separable), the shear power spectral density was integrated over  $m$  and set equal to the assumed critical shear variance, 0.5. This procedure determined the absolute intensity of the spectral form and led to a wind power spectrum of  $N_0^2/6m^3$  in the saturated tail. This spectrum was claimed to be in good agreement with observations -- better agreement, in fact, than that given by the  $N_0^2/2m^3$  result of the  $\Delta m = m$  "narrow-band" group. The agreement was taken as evidence not only in favor of the instability theory, but also of the requirement to integrate the shear over the full spectrum when establishing the condition for instability --

thereby denying the individual-group model on which the form of the spectrum had first been based.

Here again, the favorable conclusion is one of compatibility rather than determinacy. It could be turned quite the other way around: the wave spectrum might be established in the model form by processes other than instability, and the SFV87 calculation would then merely establish that the modeled spectrum, if extended to the assumed maximum  $m$ , would indeed be unstable.

Since the linear-instability theory maintains consistency with observation, it cannot be disproven by the criticisms I have made here; but the strength of the arguments in favor of that theory should, I think, have been undermined. In any event, given that theory, a further point must be made: the linear-instability theory has been developed to date on the basis of linear wave theory, but the model adopted to illustrate it reveals that the waves are highly nonlinear.

If, for example, one adopts as the wind power spectral density WPSD the form

$$WPSD = (N_o^2/6m_*^3) / (1 + [m/m_*]^3), \quad (2)$$

avored by SFV87, one finds upon integration over  $m$  the wind variance

$$\sigma_T^2 = (\pi/9\sqrt{3}) (N_o^2/m_*^2) = 0.20 (\omega/h)_*^2, \quad (3)$$

where  $\sigma_T$  is the wind standard deviation as before and  $(\omega/h)_*$  is the horizontal trace speed of a wave with the characteristic vertical wavenumber  $m_*$ , use having been made of (1). This reveals that, at  $m_*$ , the horizontal trace speed is little more than twice the standard deviation of the wave-induced wind field through which the wave is propagating. Waves having greater  $m$  will have proportionately smaller horizontal trace speeds, via (1). At a point in space where the wave-induced wind  $\underline{V}$  has a component  $V_h$  (assumed horizontal, for the moment) in the direction of the wave's propagation, the total time derivative of the Eulerian fluid-dynamic equations is

$$D/Dt = \partial/\partial t + \underline{V} \cdot \nabla = (1 + V_h/[\omega/h]) \partial/\partial t, \quad (4)$$

with the  $V_h/[\omega/h]$  term representing a nonlinear interaction between the chosen wave and the whole of the wave-induced wind field (plus any background wind, in general). This interaction will clearly be of import to all waves except, perhaps, to those with  $\omega/h \geq 2 V_h$ , which will be roughly those with  $m \leq m_*$ . Integration of the (nondimensionalized) shear power spectral density corresponding to (2), namely

$$SPSD = (m^2/6m_*^3) / (1 + [m/m_*]^3), \quad (5)$$

## Saturation of Middle-Atmosphere Gravity Waves

now shows that these relatively immune waves contribute only 0.04 to  $\sigma_s^2$ . If, as has been supposed, the critical value of  $\sigma_s^2$  is of order unity (e.g., 0.5 as assumed by SFV87), it is evident that the requisite shears must come primarily from the tail portion of the wave spectrum, the portion that suffers strong nonlinear interaction. To take an extreme: at the tip of the tail, where  $m \approx 200$  m, according to the rough estimate of SFV87, the wind standard deviation would be about 72 times the horizontal trace speed. This confirms for the middle atmosphere a conclusion about the relevance of nonlinearity reached many years ago for the ocean (e.g., Holloway 1980, 1981, Munk 1981). (My own estimate of the length of the tail — to about 23 m, as is derived below — leads to a horizontal trace speed at the tip of the tail equal to about  $\sigma_s/8$ . This corresponds closely to Holloway's 1980 statement, that oceanic internal waves are too energetic by two orders of magnitude to be considered weak waves.)

These considerations reveal that, even if one wanted to pursue an instability theory of saturation, one would in principle be forced to pursue that theory nonlinearly, specifically with the effects of the advective nonlinearity  $\underline{V} \cdot \nabla$  being taken into account. As I shall argue below, the effects of this nonlinearity seem to be adequate in themselves to shape the tail into something like an  $m^{-3}$  form, at least if there is a dissipative process acting strongly at large  $m$ . This (secondary) process could be instability, if the tail extends to large enough  $m$  such that  $\sigma_s^2$  attains the critical value, whatever that may be, but it could alternatively be molecular diffusion (as I shall argue it is, above the turbopause) or the nonlinear, wave-wave diffusion examined by Weinstock (1990).

The analysis of H91a includes a derivation of the probabilities of insipient instability (i.e., the probabilities of finding  $Ri < 1/4$  for dynamic instability and  $< 0$  for static instability) for an azimuthally isotropic, Gaussian distribution of wave-induced winds. It shows that each probability increases continuously from 0 as  $\sigma_s$  increases from 0 and reaches appreciable levels (of order 0.1) for  $\sigma_s$  of order unity, thereby confirming the criteria adopted by DG86 and SFV87 for their critical shears, and providing a firm base for future application.

### 3. DEVELOPMENT OF DOPPLER-SPREAD THEORY

The intent of this section is to make an analytic estimate, necessarily crude, of the consequences of the advective nonlinearity in a spectrum of waves such as the middle atmosphere supports. For the purpose, I assume an atmosphere that is wind-free and nearly isothermal in the absence of the waves, gravitationally stratified and nonrotating. I further assume a power spectral density of x-component ( $u$ ) and y-component ( $v$ ) of wave-induced wind given by

$$Q_u^2 = (k/h)^2 Q^2(k, l, m) = \cos^2 \alpha Q^2(\alpha, h, m) \quad (6)$$

and

$$Q_v^2 = (l/h)^2 Q^2(k, l, m) = \sin^2 \alpha Q^2(\alpha, h, m) \quad (7)$$

respectively, where  $\alpha = \arccos k/h = \arcsin l/h$  is the azimuth of wave propagation and  $Q^2$  is the power spectral density of the wave-induced wind in the azimuth of the wave's own propagation. The notation within braces is intended to indicate that  $Q^2$  may be thought of as a function of horizontal-component wavenumbers  $k$  and  $l$ , or azimuth  $\alpha$  and horizontal wavenumber  $h$ , in addition to vertical wavenumber  $m$ .

The spectral components are taken to be randomly phased relative to one another, and the spectrum is taken to be broad, in which case the Central Limit Theorem establishes that the horizontal wind components will have a Gaussian distribution: the probability of finding  $u$  between  $u$  and  $u + du$  is given by

$$P_u \{u\} du = \frac{1}{\sqrt{2\pi V_u}} \exp\{-u^2/2V_u\} du, \quad (8)$$

where  $V_u$  is the variance of  $u$ , given by

$$\begin{aligned} V_u = \sigma_u^2 &= \iiint (k/h)^2 Q^2(k, l, m) dk \cdot dl \cdot dm \\ &= \iiint \cos^2 \alpha Q^2(\alpha, h, m) h d\alpha \cdot dh \cdot dm \end{aligned} \quad (9)$$

and likewise for the  $y$ -component. (The spectrum is taken to contain only going waves — that is, waves with positive  $m$ , under present convention — and the integration over  $m$  is correspondingly restricted.) For convenience, I shall assume an azimuthally isotropic spectrum, so that  $Q^2$  is independent of azimuth and the integration over  $\alpha$  may be conducted trivially. Then each of the two variances is equal to half the total (horizontal) wind variance  $\sigma_T^2$ :

$$2V_u = 2V_v = \sigma_T^2 = 2\pi \iint Q^2(h, m) h dh \cdot dm \quad (10)$$

and

$$P_u \{u\} = P\{u\} = \frac{1}{\sqrt{\pi} \sigma_T} \exp \{-u^2/\sigma_T^2\}. \quad (11)$$

Similarly for  $P_v(v)$ .

I now assume that a spectrum of the type assumed is incident on the middle atmosphere from below, where the waves can be taken to be non-interacting — i.e., linear wave theory applies. The  $Q^2$  of the spectrum there will be denoted by  $Q_i^2$ , the "i" indicating "initial". The task now is to determine how this spectrum will alter with height, as wave amplitude increases and the advective nonlinearity comes into play.

For the purpose, I consider a small packet of waves in the middle atmosphere having wavenumbers in the bin  $\Delta k \cdot \Delta l \cdot \Delta m$  at  $(k, l, m)$ . This is not to be construed as a physical wave packet, engendered by some particular source, but

## Saturation of Middle-Atmosphere Gravity Waves

rather just a mathematical construct deriving from the waves that happen to be at hand. I take it to be propagating through the irregular wind system provided by the full wave spectrum, but I take the effect of that wind system to be introduced only by the horizontal component of wind and to be treatable, in its statistical consequences, as if that component were a "background" wind, unvarying in time or horizontal position. These restrictions will limit the accuracy and to some extent the validity of the form of subsequent relations, to be sure, but it is to be hoped that they do not invalidate the general (statistical) tendencies that will be revealed.

Given these assumptions, the chosen wave packet will retain its  $\alpha$  and  $h$  unchanged; I shall take  $\alpha = 0$  for convenience at the start. In the underlying region, the packet's intrinsic frequency will be  $\omega_i$  and its vertical wavenumber will be

$$m_i = N_o k / \omega_i \quad (12)$$

from (1). At some height of interest, however, where the wave-induced wind field has x-component  $u$ , the intrinsic frequency will be Doppler shifted to

$$\omega = \omega_i - ku \quad (13)$$

and the dispersion relation (1) then establishes that

$$m^{-1} = m_i^{-1} - u/N_o. \quad (14)$$

That is to say, the spectral energy located in the bin  $\Delta k. \Delta l. \Delta m$  at  $\{k, 0, m\}$  in wavenumber space, at a point in physical space-time where the x-component of the total perturbation wind is  $u$ , will have arrived there as a consequence of Doppler shifting from the vicinity of  $\{k, 0, m_i\}$  with  $m_i$  defined by (14). Though (13) permits  $\omega$  and  $\omega_i$  to differ in sign, such an occurrence would indicate Doppler shifting through the critical condition  $\omega = 0$  at which extreme dissipation is anticipated; it will not be admitted in the present work. Correspondingly, with  $m_i$  restricted to positive values,  $m$  will be likewise restricted.

It is known from previous work (e.g., Hines and Reddy 1967) that  $(u')^2 m^{-1}$  must be height-invariant as a single mode (with perturbation wind  $u'$ , vertical wavenumber  $m$ ) propagates through a background wind system, if reflections are ignored (as in a WKBJ treatment). I take the spectral analogue of that conclusion to be that  $Q^2 m^{-1} dm$  must be invariant under Doppler shifting in the present case. (The standard exponential growth with height has not been taken into account here. It applies uniformly across the whole spectrum and so plays no explicit part in deforming the spectrum. It will be reintroduced later.) Thus, were the  $u$  of (14) the only  $u$  ever encountered, we would find that

$$Q^2 m^{-1} dm = Q_i^2 m_i^{-1} dm_i \quad (15)$$

or

$$Q^2 = Q_i^2 m m_i^{-1} (dm_i/dm) = Q_i^2 m^{-1} m_i, \quad (16)$$

where  $Q_i^2$  is the  $Q^2$  of the initial spectrum and the derivative has been taken from (13) subject to  $u$  being held constant.

In fact, one finds a whole range of  $u$ 's at a given point of space in the course of time; (16) is found only with a certain probability, given by  $P_u(u) du$  for a small range  $du$  about  $u$ . This small range corresponds to a range  $dm_i$  about the initial  $m_i$  such that

$$du = |du/dm_i| dm_i = N_o m_i^{-2} dm_i, \quad (17)$$

in which the derivative has been taken from (14) subject to  $m$  being held constant. In combination with (11) and (16), and with integration over all contributing  $dm_i$ , this implies that the  $Q^2$  appropriate to  $(k,0,m)$  is given by

$$Q^2(k, 0, m) = \int Q_i^2(k, 0, m_i) m^{-1} m_i^{-1} N_o P_u\{N_o(m_i^{-1} - m^{-1})\} dm_i. \quad (18)$$

In the present case of an azimuthally isotropic spectrum of waves, this may be rewritten in terms of  $Q^2(h, m_i)$  and  $Q_i^2(h, m_i)$ , with  $P_u$  the corresponding isotropic probability, thus:

$$Q^2(h, m) = \int Q_i^2(h, m_i) \frac{N_o}{\sqrt{\pi} \sigma_T m m_i} \exp \{-N_o^2 (m_i^{-1} - m^{-1})^2 / \sigma_T^2\} dm_i. \quad (19)$$

If the vertical wavenumbers are nondimensionalized via multiplication by  $\sigma_T/N_o$ , the transformation may be rewritten as

$$Q^2(h, M) = \int Q_i^2(h, M_i) \left[ \frac{e^{-(M_i^{-1} - M^{-1})^2}}{\sqrt{\pi} M M_i} \right] dM_i, \quad (20)$$

where

$$M = m \sigma_T / N_o \quad (21)$$

and similarly for  $M_i$ .

The factor in brackets in (20) may be thought of as a transfer function, leading from the initial to the observable spectrum in the scaled units. It is independent of  $h$  as well as azimuth, and so applies equally to the one-dimensional (in  $m$ ) spectrum obtained by integration over  $h$ , the spectrum that exhibits an approximately  $m^{-3}$  portion in the middle atmosphere. Its consequences are indicated here with the aid of Figures 1 and 2.

Figure 1 exhibits the transfer function itself, as a function of  $M$ , for a number of values of  $M_i$  in a range about  $M_i = 1$  (for which  $\omega_i/h = \sigma_T$ ). This



## Saturation of Middle-Atmosphere Gravity Waves

function may be thought of as the statistically observable spectrum  $Q_\delta^2$  derived from a delta-function input spectrum at  $M_i = M_\delta$ , this position being marked by the abscissa of the heavy dot on each curve in turn. This observable spectrum is seen to be broadened — or Doppler spread —, with its peak moving to an  $M$  value somewhat lower than  $M_\delta$  and a tail extending to higher  $M$ . From the form of (20), it is apparent that this tail asymptotes to the form  $M^{-1}$ , and yet, for  $M_\delta$  values somewhat smaller than 1, it exhibits a markedly stronger variation before reaching the asymptotic form. Indeed, for  $M_\delta = 1/2$ , it exhibits something close to an  $M^{-3}$  form over at least a decade in  $M$ , which would be consistent with what we know of the actually observed spectra. (The percentages associated with the individual curves will be explained shortly.)

Figure 2 exhibits the convolution  $Q_c^2$  of the transfer function with a step-function input — with an input spectrum that is flat (white noise) up to some cutoff value  $M_c$  (labeled and with abscissa marked by the heavy dot on each curve in turn) and is then cut off to zero. Again a tail is found, with the  $M^{-3}$  form occurring again and extending over at least a decade in  $M$ , now for  $M_c = 1/2$ .

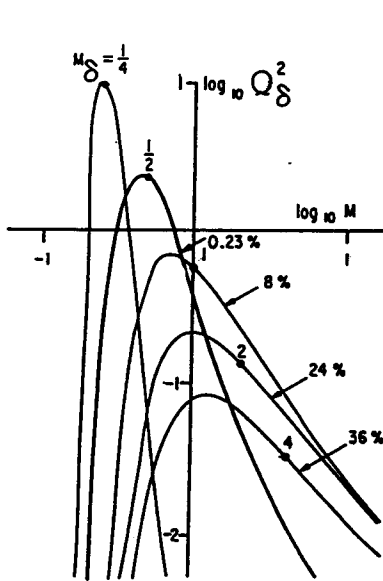


Figure 1

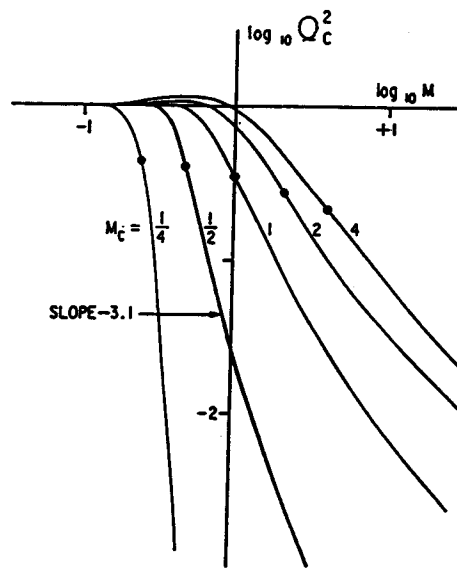


Figure 2

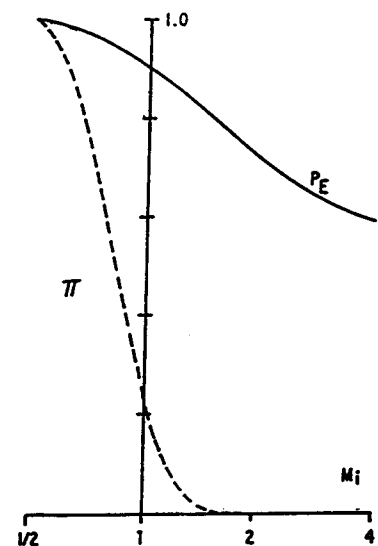


Figure 3

It would seem from these examples that the advective nonlinearity has a propensity for producing a semi-universal spectrum of the observed form, at least if some mechanism exists that would prevent incident waves having  $M_i > 1/2$  or 1 playing any substantial role in the observations. I suggest that the required mechanism is to be found in the approach of such waves to critical-layer interactions at and below the height of observation. As an approximation, I assume that waves with vertical wavenumber exceeding some maximum value  $m_M$  (scaled alternatively to  $M_M$ ) are simply obliterated.

The probability of a wave with incident wavenumber  $M_i$  having a local wavenumber less than  $m_m$ , and so escaping obliteration, is given by

$$\begin{aligned} P_E &= 0.5 + 0.5 \operatorname{erf} \{N_o/m_i\sigma_T - N_o/m_m\sigma_T\}, \\ &= 0.5 + 0.5 \operatorname{erf} \{M_i^{-1} - M_M^{-1}\} \end{aligned} \quad (22)$$

adapted from (2.22) of H91b, erf being the error function,

$$\operatorname{erf}(x) = 2\pi^{-\frac{1}{2}} \int_0^x e^{-q^2} dq.$$

This probability is depicted by the continuous curve in Figure 3 for the case  $M_M = 11.5$ , which will be justified below, although the critical-level case  $M_M = \infty$  is not much different. (The percentages marked on Figure 1 represent the complementary probability, namely that waves of the respective  $M_i$  values will indeed be obliterated.)

This  $P_E$  curve applies only locally, at the height currently of interest. To have survived to this height, the waves must have avoided obliteration at all lower levels. The probability of this successful escape is roughly the product of the probabilities of escape through each of the statistically independent underlying slabs of atmosphere. I take the depth of each such slab to be  $\lambda_c/2$ , where  $\lambda_c = 2\pi/m_c$  is the cutoff vertical wavelength of the incident wave spectrum — the transition wavelength between the small- $m$  body and the large- $m$  tail spectra — and I number the underlying slabs with the index  $n$ , counting downward from the current height of interest. I also take  $\sigma_T$  to vary with height as  $\exp\{z/4H\}$  and  $m_m\sigma_T$  to be height-independent (see below). The probability of successful escape through all underlying slabs is then the product of individual-layer factors given by (22) for each in turn:

$$\Pi = \prod_{n=1}^{n_s} (0.5 + 0.5 \operatorname{erf} [N_o \exp\{n\lambda_c/8H\}/m_i\sigma_T - N_o/m_m\sigma_T]), \quad (23)$$

where  $n_s$  is the number of slabs between the height of observation and the source height. The  $\Pi$  curve of Figure 3 represents this probability for  $M_M = 11.5$  again, with the use of  $\lambda_c = 1.07$  km and  $H = 7$  km, which are representative of the middle stratosphere (ca. 25 km height), and  $n_s = 30$ , corresponding to a source 16 km below. This curve clearly exhibits a sharp cutoff that might be approximated by a step function sited at some  $M_c$  in the range 0.5 – 1, as was wanted. (This calculation and these curves are newly presented here; they are not contained in H91b but are now submitted for publication as Part IV of the in-press series.)

A finite  $M_M$ , such as I have adopted, can be imposed by molecular diffusion (viscosity and heat conduction), and in the next section I take it to be so imposed above the turbopause. Below the turbopause, where (by definition) turbulence is to be found, linear instability might well provide the relevant mechanism. But it would be the mechanism, now, for limiting the length of the tail, not for determining the tail's form or intensity. Examination of this possibility in any detail analytically requires an analytic model spectrum.

## Saturation of Middle-Atmosphere Gravity Waves

In H91b, following VanZandt and Fritts (1989), I adopt the "modified Desaubies" spectral form

$$Q^2\{m\} = Km / (1 + [m/m_c]^4), \quad (24)$$

where  $m_c$  is a characteristic vertical wavenumber corresponding closely to  $m_i$  in Section 2, in that it marks the onset of the tail portion of the spectrum. The  $m_c$  notation is justified because this  $m_c$  is analogous also to the cutoff of incident wavenumbers in Figure 2, where too the tail may be said to begin. I take as my choice  $m_c = N_o/2\sigma_T$  (or  $M_c = 1/2$ , in Figure 2), on the empirical grounds that it seems to represent as close to observed reality as a step-function approach to the cutoff is likely to come, but with legitimacy being provided by the  $\Pi$  curve of Figure 3. This choice combines with (24), from which  $\sigma_T^2$  can be inferred by integration, to establish the value of  $K$  as  $N_o^2/\pi m_c^4$ . (The integral is only weakly dependent on the upper limit of integration  $m_M$ , provided that  $m_M \gg m_c$  as I take it to be, and so integration to infinity is appropriate.) Then (24) becomes

$$Q^2\{m\} = N_o^2 m / \pi m_c^4 (1 + [m/m_c]^4), \quad (25)$$

which approximates to  $N_o^2/\pi m^3$  in its large- $m$  tail. The intensity of this tail is fully consistent with observations (cf.  $N_o^2/2m^3$  and  $N_o^2/6m^3$  cited in Section 2), and it was derived with no assumption being needed as to the critical value of shear or even the existence of instability (as the linear-instability derivations required). Those assumptions are replaced here by the assumption of  $N_o/2\sigma_T$  as an effective cutoff of incident wavenumbers — an assumption justified by the continuous curve of Figure 3, at least for the middle stratosphere.

I would argue, further, that in the limit of small  $m$  the waves should be unaffected by dissipation or Doppler spreading and hence the power spectral density in that limit should grow with height as  $\exp z/H$ . This assumption, imposed upon (25), leads directly to

$$m_c = m_o \exp -z/4H \quad (26)$$

and so also to

$$\sigma_T = (N_o/2m_o) \exp z/4H, \quad (27)$$

with  $m_o$  a constant that would be determined by the initial energy sources, specifically those producing small  $m_i$ . These height variations are in accord with observations as I know them; specifically, the wind variance now grows with height at half the rate appropriate to a nonsaturating, nondissipating spectrum, as it was said to do observationally in Section 1.

The (nondimensionalized) shear spectrum can be obtained from (25) via multiplication by  $m^2/N_o^2$ , and then integrated to obtain the shear variance  $\sigma_s^2$ . The integral is logarithmically divergent and must be terminated at some upper

bound  $m_M$ , beyond which the spectrum is of turbulence, not of waves. Then

$$\sigma_s^2 = (4\pi)^{-1} \ln \{1 + [m_M/m_c]^4\} = \pi^{-1} \ln \{m_M/m_c\} \quad (28)$$

where  $m_M \gg m_c$  has been assumed. Now, if instability is to determine  $m_M$ , as being  $m_{Minst}$  say, this shear variance must equal the critical value  $\sigma_{scrit}^2$  for the marginal maintenance of instability, whatever that may be. Hence,

$$m_{Minst} = m_c \exp \{ \pi \sigma_{scrit}^2 \}. \quad (29)$$

If I take  $\sigma_{scrit}$  to be 1, as suggested in Section 2, then (29) yields  $m_{Minst}/m_c = \exp \pi = 23$ : the tail extends over a 23-fold range in  $m$ . This is the value that leads to the  $P_E$  curve in Figure 3. The result  $m_{Minst}/m_c = 23$  is consistent with the observations known to me, and specifically with those of Dewan et al. (1984), though perhaps for the unfortunate reason that the data do not extend reliably over a greater range. It is possible that (29) would be used best in the opposite direction, as a means of determining  $\sigma_{scrit}^2$  empirically from observed values of  $m_{Minst}/m_c$ .

In H91b, an appendix establishes that inclusion of the Coriolis force associated with Earth's rotation leads to a tail spectrum that asymptotes to  $m^{-2}$  (rather than  $m^{-1}$  as in (19)). A second appendix outlines a failed attempt to reach an  $m^{-3}$  form by the further inclusion of the wave-induced vertical wind field in the advective nonlinearity, and a third appendix deals briefly with azimuthally anisotropic conditions. I consider my failure to produce an  $m^{-3}$  asymptotic form to be serious only from one point of view, to be discussed in Section 5; specifically, I do not consider it to be serious for purposes of comparison with data. This is because I have obtained here, at least for the cases  $M_g = 1/2$  (Figure 1) and  $M_c = 1/2$  (Figure 2), the wanted  $m^{-3}$  form as a transitional form bridging the gap between low wavenumbers, which are unaffected by Doppler spreading, and high wavenumbers, which will be subject to dissipative processes and so will (probably) be altered from whatever asymptotic form they might otherwise have achieved. So far as I am aware, that is all that the observations demand.

#### 4. FORMATION OF THE TURBOPAUSE

Molecular diffusion increases with height through the atmosphere and imposes an ultimate cutoff of the wave spectrum. I suggest that turbulence terminates — the turbopause is formed — when the cutoff imposed in this manner, which occurs at some  $m_M$  given roughly by

$$m_{Mmol} = (N_o h / 2\pi\eta)^{\frac{1}{3}}, \quad (30)$$

limits the length of the tail to a value just less than that required for the occurrence of instability. (See H91c;  $\eta$  is the molecular kinematic

## Saturation of Middle-Atmosphere Gravity Waves

viscosity.) Equality of  $m_{\text{Mmol}}$  and  $m_{\text{Minst}}$  at the turbopause then requires there that

$$\sigma_{\text{Scrit}}^2 = \pi^{-1} \ln \{ 2\sigma_T N_0^{-2/3} \lambda_h^{-1/3} \eta^{-1/3} \}, \quad (31)$$

$\lambda_h = 2\pi h^{-1}$  being a (representative) horizontal wavelength. One would like to solve this equation for the height at which it is satisfied, thereby determining the height of the turbopause, but  $\sigma_{\text{Scrit}}^2$  is as yet unknown to the accuracy required for a meaningful conclusion. Instead, I use it the other way around, as a means of estimating  $\sigma_{\text{Scrit}}^2$ .

The turbopause is typically taken to lie at a height of 100 - 110 km, where  $\eta \approx 100 \text{ m}^2/\text{s}$  and  $N_0 \approx 0.02 \text{ s}^{-1}$ . If I adopt a representative  $\lambda_h \approx 50 \text{ km}$  and a representative  $\sigma_T \approx 30 \text{ m/s}$ , then (28) yields  $\sigma_{\text{Scrit}}^2 = 0.497$ , almost exactly the value chosen by SFV87. This seems to be a perfectly reasonable value of  $\sigma_s$  to accept as a condition for the termination of instability, particularly since the wave spectrum might well have been narrowed to a nearly monochromatic wave by the time turbopausal heights are reached. The same estimates combine with the modified Desaubies spectrum (25) to imply a spectral peak (which occurs at  $m_p = 3^{-1/4} m_c = 0.76 m_c$ ) at a vertical wavelength of  $\lambda_p = 25 \text{ km}$ . Such a value is frequently reported as dominating the spectrum at turbopausal heights, often with the (possibly false) identification of the observed wave as the diurnal 1,1 tidal mode. The maximum  $m$ ,  $m_H$ , corresponds to a vertical wavelength of about 4 km, which is consistent with observations.

With the model spectrum taken to be applicable down to the tropopause, it can now be employed to estimate the wave spectrum in the middle stratosphere, where it can be compared with other data. There, both  $\lambda_p$  and  $\sigma_T$  will be decreased by the fourth root of the atmospheric density ratio, as given by (26) and (27), which root is approximately 18, yielding  $\lambda_p \approx 1.4 \text{ km}$  and  $\sigma_T \approx 1.7 \text{ m/s}$ . If  $m_{\text{Minst}}^{-1}$  is similarly scaled, it yields a transition from waves to turbulence at a vertical wavelength of about 220 m, whereas if  $\sigma_{\text{Scrit}}$  is raised to 1 before the scaling is done it yields a vertical wavelength of 46 m for the transition. The observations of Dewan et al. (1984), for example, are said to exhibit a tail with log-log slope of -3.0 extending from about 1 km down to about 200 m in vertical wavelength, with some curvature (consistent with that in Figure 1 above) at smaller wavelengths, down to about 40 m. The observations are said to be unreliable outside this range, but the theory is clearly compatible with the reliable observations. Moreover, the mean wind power spectral density at a vertical wavelength of 1 km was found to be  $3.42 \text{ (m/s)}^2/(\text{c/m})$ , which converts to present units as giving a tail spectrum of  $0.306 N_0^2/\text{m}^3$ . This is to be compared with  $N_0^2/\pi\text{m}^3 = 0.318 N_0^2/\text{m}^3$  in the present theoretical model.

It should be specially noted that, above the turbopause, molecular diffusion simply replaces instability as the mechanism of dissipation, but the Doppler-spread theory continues on, otherwise unaffected in principle (until the spectrum becomes so narrow that a statistical treatment is inappropriate). There is, however, one side-effect that comes into play. At these great

heights, the statistically independent slabs of atmosphere are much thicker, so there will be fewer of them in the height range  $4H$  over which  $\sigma_T$  decreases substantially (on moving downward from slab to slab). Consequently, the product probability  $\Pi$  in (23) is unlikely to be decreased from  $P_E$  in (22) by anything like the degree illustrated by the transition from the one curve to the other in Figure 3. The cutoff of incident waves will then not be as sharp: a step-function approximation may be inappropriate and the spectral form of the tail might be more like one of the forms depicted for values of  $M_c$  equal to or greater than 1 in Figure 2.

## 5. DISCUSSION

The present analysis has adopted a necessarily approximate approach to the estimation of the effects of Doppler spreading in the middle-atmosphere spectrum of gravity waves. Despite its failings, however, it seems to have shown that these effects are important — are even necessary to include in any alternative theory of the spectral tail — and on their own are of a nature that accords with observation, without the necessity for an alternative theory.

The importance of the advective nonlinearity that gives this Doppler spreading has been recognized in oceanographic studies for more than a decade now (e.g., reviews by Holloway 1980, 1981 and Munk 1981), and some progress has been made in numerical studies that incorporate its effects (e.g., Flatte et al. 1985). The present analysis might well be carried over to oceanographic studies, and perhaps improved, to account for the approximately  $m^{-3}$  spectral form in velocity (or  $m^{-1}$  spectral form in shear) that is found there over a middle range of vertical wavenumbers.

In making the transition, one would have to include both upgoing and downgoing waves, but that change will be of minimal operational consequence. One would also have to drop the exponential growth of wave amplitude with height, a change that would have at least two important consequences, one operational and one conceptual.

Operationally, the step that led from the continuous to the broken curve of Figure 3 will not have the exponential change with height that is allowed for in (23). Instead, the full ocean depth  $D$  will act uniformly to limit the probability of escape from obliteration, and that depth will be measured in some height-independent characteristic wavelength  $\lambda_c$  ( $\sim 10$  m) to give a very large number of statistically independent slabs. The relevant probability for escape from obliteration will be given by a modified form of (23) in which all factors are identical and the required probability becomes simply the probability of escape from obliteration in one slab raised to the power  $4D/\lambda_c$ , a very large number and one that will produce an even closer approach to a step-function form than that given by the  $\Pi$  curve in Figure 3. (The relevant power will be even greater than  $4D/\lambda_c$ , if it is held that the borderline waves can propagate up and down more than once before being obliterated. The relevant power may, moreover, change somewhat with  $D$  from case to case and so perhaps produce cutoffs and spectral slopes that similarly change somewhat, observationally. These questions open the way to further investigation.)

## Saturation of Middle-Atmosphere Gravity Waves

Conceptually, the physical growth of wave amplitude with height that occurs in the atmosphere must be replaced, in oceanographic work, by the imagined growth that would have occurred, had the spectrum of waves been "turned on" with a gradually increasing intensity, a gradually increasing nonlinearity. Or, alternatively, it might be replaced simply by a straightforward calculation of the steady-state nonlinear result, if such a calculation can be said to be straightforward.

Happily enough, such a basis for calculation has in fact been laid already by Allen and Joseph (1989), but it was developed in different terms. That work adopts (in its Case III) a canonical spectrum of waves as described in a Lagrangian formulation and evaluates the appearance of the spectrum as it would be seen in Eulerian coordinates. The waves are fully linear in the Lagrangian description — their frequencies and wavenumbers lie on the dispersion-equation surface in 4-space — but are allowed to appear nonlinear in the Eulerian description, with the nonlinearity arising from the advective nonlinearity of the Eulerian equations only. That is precisely the transition that I have attempted to model here, beginning with the initial spectrum from which nonlinearities were excluded and ending with an observable spectrum in which the effects of the advective nonlinearity are included (albeit for reasons of amplitude growth with height, rather than a transition from a Lagrangian to an Eulerian description).

The Allen and Joseph work reaches (amongst other things) an  $m^{-3}$  form for the ultimate tail spectrum, a thing I have been unable to do. As noted above, inclusion of the Coriolis force (which Allen and Joseph included) would have led me to an  $m^{-2}$  form, but that is not enough. Allen and Joseph automatically included also the part of the advective nonlinearity that comes from vertical advection, a thing I have been unable to achieve successfully, and I suppose my failure is a consequence of my oversimplification of the means of handling the Doppler shifts. This leaves a missing link in the chain connecting my work to theirs, but conceptually the two approaches seem identical in the intended transition they incorporate.

The Allen and Joseph work as developed to date does not give the form of transition from the small- $m$  body to the large- $m$  tail of the wave spectrum, which is what my own analysis does succeed in doing, however approximately. It is this transition, rather than an asymptotic tail (which would in any event be deformed by dissipative processes), that is seen in the middle atmosphere, I believe. Perhaps it is this same transition in the oceans, as well.

Acknowledgments: I wish to thank Carmen Torres for her assistance in the production of this typescript. The Arecibo Observatory is operated by Cornell University with funding provided under a cooperative agreement with the National Science Foundation.

## REFERENCES

- Allen, K. R., and R. I. Joseph, 1989: A canonical statistical theory of oceanic internal waves. *J. Fluid Mech.*, 204, 185-228.
- Balsley, B. B., and D. A. Carter, 1982: The spectrum of atmospheric velocity fluctuations at 8 km and 86 km. *Geophys. Res. Lett.*, 9, 465-468.
- Dewan, E. M., and R. E. Good, 1986: Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere. *J. Geophys. Res.*, 91, 2742-2748.
- Dewan, E. M., N. Grossbard, A. F. Queseda and R. E. Good, 1984: Spectral analysis of 10m resolution scalar velocity profiles in the stratosphere. *Geophys. Res. Lett.*, 11, 80-83 and 624.
- Flatte, S. M., F. S. Henyey and J. A. Wright, 1985: Eikonal calculations of short-wavelength internal-wave spectra. *J. Geophys. Res.*, 90, 7265-7272.
- Hines, C. O., 1991a: The saturation of gravity waves in the middle atmosphere. Part I: Critique of linear-instability theory. *J. Atmos. Sci.*, in press.
- Hines, C. O., 1991b: The saturation of gravity waves in the middle atmosphere. Part II: Development of Doppler-spread theory. *J. Atmos. Sci.*, in press.
- Hines, C. O., 1991c: The saturation of gravity waves in the middle atmosphere. Part III: Formation of the turbopause and of turbulent layers beneath it. *J. Atmos. Sci.*, in press.
- Hines, C. O., 1991d: The saturation of gravity waves in the middle atmosphere. Part IV: The cutoff of the incident wave spectrum. Submitted for publication, *J. Atmos. Sci.*
- Hines, C. O., and C. A. Reddy, 1967: On the propagation of atmospheric gravity waves through regions of wind shear. *J. Geophys. Res.*, 72, 1015-1034.
- Hodges, R. R. Jr., 1967: Generation of turbulence in the upper atmosphere by internal gravity waves. *J. Geophys. Res.*, 72, 3455-3458.
- Holloway, G., 1980: Oceanic internal waves are not weak waves. *J. Phys. Ocean*, 10, 906-914.
- Holloway, G., 1981: Theoretical approaches to interactions among internal waves, turbulence and fine structure. *Nonlinear Properties of Internal Waves* (Ed: B. West). American Institute of Physics.
- Lumley, J. L., 1964: The spectrum of nearly inertial turbulence in a stably stratified fluid. *J. Atmos. Sci.*, 21, 99-102.



## Saturation of Middle-Atmosphere Gravity Waves

- Munk, W., 1981: Internal waves and small-scale processes. *Evolution of Physical Oceanography* (Eds: B. A. Warren and C. Wunsch). MIT Press.
- Phillips, O.M., 1977: *The Dynamics of the Upper Ocean*, 2nd ed., Cambridge University Press, New York.
- Smith, S. A., D. C. Fritts and T. E. VanZandt, 1987: Evidence for a saturated spectrum of atmospheric gravity waves. *J. Atmos. Sci.*, 44, 1404-1410.
- Tsuda, T., T. Inoue, D.C. Fritts, T.E. Van Zandt, S. Kato, T. Sato and S. Fukao, 1989: MST radar observations of a saturated gravity wave spectrum. *J. Atmos. Sci.*, 46, 2440-2447.
- Van Zandt, T.E., 1982: A universal spectrum of buoyancy waves in the atmosphere. *Geophys. Res. Lett.*, 9, 575-578.
- Vincent, R. A., 1984: Gravity wave motions in the thermosphere. *J. Atmos. Terr. Phys.*, 46, 119-128.
- Weinstock, J., 1990: Saturated and unsaturated spectra of gravity waves, and scale dependent diffusion. *J. Atmos. Sci.*, 47, 2211-2225.