JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 94, NO. D12, PAGES 14,617–14,628, OCTOBER 20, 1989

# On Dynamics and Transport Near the Polar Mesopause in Summer

MICHAEL E. MCINTYRE

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England*

Some basic aspects of dynamics and transport in polar regions within one or two scale heights of the summer mesopause are examined, and the implications for local one-dimensional photochemical modeling of those regions discussed. Included is a simple thought experiment on gravity wave breaking that throws further light on the so-called "turbulent Prandtl number" question, and related questions. The "downward control" of time averaged, zonally averaged upwelling by gravity wave breaking is noted, and a one-dimensional modeling strategy suggested in which the transport of water vapor and other constituents from below is characterized by a single parameter, defining a mean upwelling velocity inversely proportional to the mean mass density. It is suggested that, at and below noctilucent cloud altitudes, vertical mixing should be altogether neglected.

## 1. INTRODUCTION

There is a well-founded consensus that transport processes are important near the summer polar mesopause, and that gravity waves are involved. The fact that transport strongly influences a number of radiative and photochemical processes in the mesosphere, including those leading to the formation of noctilucent clouds (NLC) and polar mesospheric clouds (PMC), seems clear from many observational and theoretical studies (see, for instance, papers in this issue, and references therein including the recent review of observational constraints on transport by *Strobel* [1989], and the observational information in, for instance, *Schröder* [1974], *Olivero and Thomas* [1986], *Gadsden* [1986], and most recently *Thomas et al.* [1989].)

There are two quite distinct transport effects that might reasonably be assumed to be important. The first is the mean upwelling in summertime polar latitudes (variable, but probably within the order-of-magnitude range 1–10 $\mathrm{cm\,s^{-1}}$ at the mesopause) that occurs as part of the global scale circulation induced by breaking gravity waves. This could be a significant factor in levitating NLC/PMC particles as well as in supplying water vapor from below, the more so since colder temperatures should go with stronger upwelling over the previous few days. Upwelling velocities of several centimeters per second, giving adiabatic cooling rates of one or two tens of degrees per day in the upper mesosphere, seem to be of the right order [e.g., *Holton*, 1983; *Garcia*, this issue] to depress temperatures below radiative equilibrium by the 70 K or so that may be required for NLC/PMC formation near 83–85 km, albeit subject to uncertainties in the calculation of radiative relaxation rates in the absence of local thermodynamical equilibrium. The relevant attribute of the gravity waves is their ability to transport momentum and angular momentum from regions of the atmosphere many scale heights below. The manner in which this wave-induced momentum transport gives rise to upwelling, and hence temperature depression, is recalled in section 3.

The second transport effect that might be assumed to be important is vertical mixing of heat and constituents by the turbulence due to breaking gravity waves. For instance, such mixing seems to be needed to account for transport of atomic oxygen downward from $\gtrsim 90\,\mathrm{km}$ against the mean upwelling [e.g., *Thomas et al.*, 1984; *Garcia and Solomon*, 1985] or the downward transport of certain ions observed by incoherent-scatter radar (e.g., C. Hall, personal communication, 1988). The corresponding downward heat transport is also a potentially significant factor in the lower thermospheric heat budget and hence in the control of temperatures, albeit well above NLC/PMC altitudes [e.g., *Zimmerman and Keneshea*, 1986, *Strobel*, 1989, and references therein].

Various formulae have been proposed to model such vertical mixing in terms of a vertical eddy diffusivity $D$. For example, in observational studies in which the turbulent energy dissipation rate $\varepsilon_1$ per unit mass is inferred from measurements, it has sometimes been assumed that one may use

$$D = \beta_1 \varepsilon_1 / N^2 \tag{1}$$

where $\beta_1$ is a dimensionless number, taken to be some modest fraction of unity, perhaps 0.25 or 0.5. Here

$$N^2 = \frac{g}{\theta}\frac{\partial \theta}{\partial z} = \frac{g}{T}\left(\frac{\partial T}{\partial z} + \frac{g}{c_p}\right)$$

i.e., the square of the buoyancy or Brunt-Väisälä frequency; $g$ is the gravity acceleration, $\theta$ is the potential temperature, $z$ the altitude, $T$ the temperature, and $c_p$ the specific heat at constant pressure. The eddy fluxes of heat and long-lived constituents are usually written as

$$\rho c_p D\left(\frac{\partial T}{\partial z} + \frac{g}{c_p}\right) \tag{2a}$$

$$\rho D \frac{\partial \chi}{\partial z} \tag{2b}$$

respectively, where $\rho$ is the mass density and $\chi$ is a constituent mixing ratio. These formulae may be contrasted with the corresponding molecular formulae and assume, in the usual way, that entropy and constituents are transported advectively by small scale turbulent air



Paper number 89JD00310.
0148-0227/89/89JD-00310 $05.00

motions, to which the theory of homogeneous turbulence is approximately applicable.

Closely related to (1) is Lindzen's well-known parameterization for an upward propagating gravity wave (above its breaking altitude and below any critical levels that may exist). Lindzen's formula is equivalent to taking $\beta_2 = 1$ in

$$D = \beta_2 \varepsilon_2 / N^2 \tag{3}$$

where $\varepsilon_2$ is the local rate of loss of intrinsic wave energy per unit mass, taken by definition to be equal to the mean wave energy flux convergence per unit mass,

$$\varepsilon_2 = \frac{w_g E}{\rho_0 H_{\text{diss}}} = \frac{k(\bar{u} - c)^4}{2N H_{\text{diss}}} \tag{4}$$

where $w_g$ is the vertical group velocity, $E$ is the wave energy density, $\rho_0$ the mean mass density, and $H_{\text{diss}}$ the height scale for the dissipation of wave energy or wave action by breaking (see equation (8)). The quantity $\varepsilon_2$ may often be comparable in order of magnitude to $\varepsilon_1$ (see next section), but the two need not be equal, since when wave energy disappears some of that energy goes into increasing the basic state potential energy. Indeed we shall see that calculating the difference $\varepsilon_2 - \varepsilon_1$ is almost the same thing as calculating the vertical mixing in which we are interested, since the latter is closely related to changes in basic state potential energy; cf. (17) below. The remaining symbols in (4) have their usual meanings: $c$ and $k$ are the waves' horizontal phase speed and wave number, respectively, and $\bar{u}$ is the mean wind, so that $(c - \bar{u})$ is the intrinsic phase speed, that is to say the phase speed Doppler-shifted to the viewpoint of an observer moving with the mean wind. The last step in (4) depends upon the well-known dispersion properties

$$(\bar{u} - c)^2 = \frac{N^2}{m^2} \quad \text{and} \quad w_g = \frac{Nk}{m^2} = \frac{k(\bar{u} - c)^2}{N} \tag{5}$$

for quasi-hydrostatic internal gravity waves, where $m$ is vertical wave number; it also depends upon an assumption that

$$E = \tfrac{1}{2}\rho_0(\bar{u} - c)^2 \tag{6}$$

at breaking. The expression (6) assumes that the dimensionless wave amplitude $a$, defined as the horizontal disturbance velocity amplitude divided by $|\bar{u} - c|$, is equal to unity, and again uses the quasi-hydrostatic approximation and the fact that the wave energy $E$ is then equal to twice the mean horizontal disturbance kinetic energy. $H_{\text{diss}}$ can be defined as $w_g$ divided by the logarithmic wave energy or wave action dissipation rate. For constant $N$ and zero mean shear it can be shown within the usual ray theoretic approximations that $H_{\text{diss}}$ is equal to the density scale height,

$$H_{\text{diss}} = H_\rho \equiv \frac{\rho_0}{d\rho_0/dz} \tag{7}$$

and more generally that

$$H_{\text{diss}} = \left( \frac{1}{H_\rho} + \frac{3}{c - \bar{u}} \frac{\partial \bar{u}}{\partial z} + \frac{1}{N}\frac{dN}{dz} \right)^{-1} \tag{8}$$

This allows for the effect of shear and vertical buoyancy-frequency variations $dN(z)/dz$ at the same (ray theoretic) level of approximation as the density term. The derivation is a slight extension of that given by *Lindzen* [1981] and *Fritts* [1984]. The condition

$$a = 1 \tag{9}$$

used in deriving (6) will be recognized as the well-known gravity wave "saturation hypothesis" in its simplest form.

This paper has two related purposes. The first (section 2) is to describe and analyze a simple thought experiment on gravity waves that suggests an almost pathologically sensitive behavior of the coefficients $\beta_1$ and $\beta_2$ appearing in (1) and (3), as regards their dependence on the wave supersaturation $(a - 1)$. On the assumption that the waves break by convective overturning, it will appear that the values of $\beta_1$ and $\beta_2$ may vary by an order of magnitude or more as $(a - 1)$ varies over a modest range, as it must be supposed to do in reality. The thought experiment is highly idealized and artificial, but is enough to suggest strongly that the use of either (1) or (3) with fixed, order-unity values of $\beta_1$ and $\beta_2$ could lead to gross error. The supersaturation $(a - 1)$ can be expected to vary for instance according to the suddenness with which the waves encounter breaking conditions (see also *Weinstock* [1989]), just as can be observed for ocean waves incident on steep versus gentle beaches. The "steep beach" analogy applies to an increasing extent when values of $mH_{\text{diss}}$ and $m/k$ decrease towards unity. Constructive interference effects will cause further variability in $(a - 1)$, in realistic wave fields. The results add to earlier doubts about taking $\beta_2 = 1$ in (3) raised by the work of *Chao and Schoeberl* [1984], *Walterscheid* [1984], *Fritts and Dunkerton* [1985], *Schoeberl* [1988], and *Coy and Fritts* [1988].

The second purpose (sections 3 and 4) is to develop a possible strategy, taking account of the results just described, for representing transport in one-dimensional photochemical models of the processes that lead to NLC/PMC formation. It will be argued that the sensitivity of $\beta_1$ and $\beta_2$ to $(a - 1)$ may be less important for this purpose than might at first be thought. Indeed, a consideration of how the gravity wave field controls the mean upwelling will suggest the usefulness of a modeling strategy that altogether ignores the turbulent transport of water vapor and other chemical constituents at and below NLC/PMC altitudes, and attributes the water vapor supply entirely to mean upwelling.

## 2. An Idealized Thought Experiment on Breaking Gravity Waves

The observed structure of breaking gravity waves both in the laboratory [e.g., *Koop and McGee*, 1986] and in the real mesosphere and lower thermosphere [e.g., *Kopp et al.*, 1985, *Fritts and Rastogi*, 1985, C.R. Philbrick, personal communication, 1988] provides merely one instance of the extreme spatial inhomogeneity that seems to characterize many naturally occurring turbulent fluid flows. The point is discussed in a wider context in a recent middle atmospheric review and forward look [*McIntyre*, 1987; see sections 5 and 8]. By contrast, the classical justification of eddy-diffusive flux formulae like (2*a*) and (2*b*) involves an explicit or tacit assumption that the turbulence is nearly homogeneous.

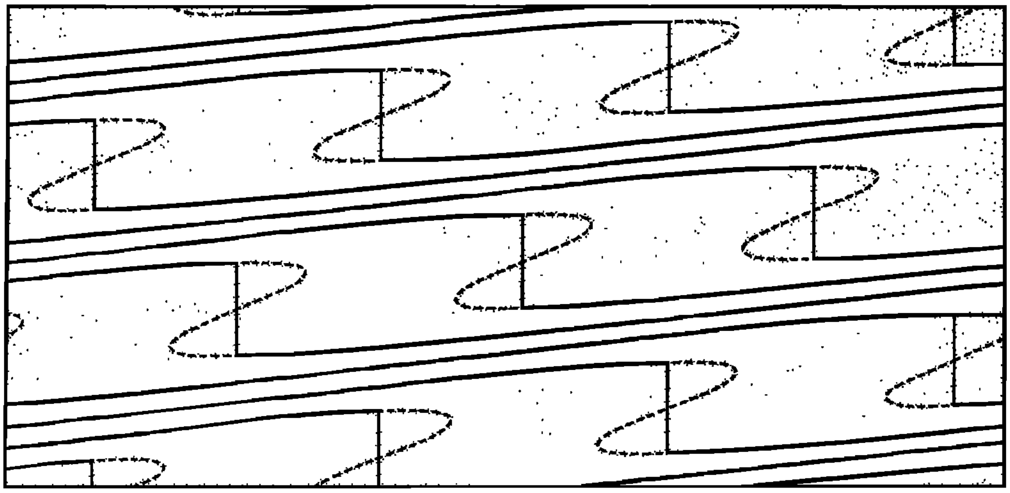

Fig. 1. Idealized, doubly periodic internal gravity wave field viewed in the $xz$ plane before and after breaking, assuming convective adjustment. The case $a = 1.66$ is shown. The solid curves are the isentropes $\theta = \theta_0(z) + \theta'(x, z, t)$ = constant after breaking, and the dashed portions show the shapes of the same isentropes in the initial wave field. The shaded regions represent the turbulent layers. In the model calculation these are idealized as perfectly mixed, with respect to horizontal momentum as well as potential temperature. Note that these mixed layers are deeper than the layers that were statically unstable initially.

More precisely, there has to be a separation of spatial scales: it is assumed inter alia that the mean gradients $\partial/\partial z$ are characterized by much larger spatial scales than the largest turbulent eddies [e.g., *Batchelor and Townsend*, 1956]. In contemplating the real atmosphere, one gets the impression that this scale separation assumption may often be one of the worst modeling assumptions one can make; therefore it is not obvious to what extent flux-gradient formulae like (2*a*) and (2*b*) should apply at all. And even if they do apply to some extent, despite the likelihood of being outside the range of circumstances for which they are justifiable, there are further questions, such as whether the same value of $D$ should be used in (2*a*) as in (2*b*). It is possible that even a long-lived tracer might have spatial fluctuations in mixing ratio that are not perfectly correlated with spatial fluctuations in potential temperature (P. H. Haynes, personal communication, 1987).

Even supposing that (2*a*) and (2*b*) do apply to a useful extent, with approximately the same $D$ values, one still has to ask whether $D$ can be evaluated from (1) or (3) and whether momentum fluxes can be similarly parameterized, as in the Lindzen parameterization, and if so with what $D$ values. There is also the distinction between $D$ values that are chosen to account for wave dissipation, in the heat and momentum equations for the wave motion, versus $D$ values that are chosen to describe mean vertical transport in a way that is relevant to photochemical modeling. For inhomogeneous turbulence there is no reason why any of these $D$ values should be the same, or even particularly close to one another, and it is not a well-defined question to ask for "the" value of the eddy diffusivity or turbulent Prandtl number without saying which choice is meant.

The purpose of this section is to analyze an idealized but physically conceivable situation that throws some light on questions like these, and illustrates the points just made. The results also strongly indicate that if formulae of the type (1) and (3) are to have any hope of conforming to reality, then we must abandon the idea that the coefficents $\beta_1$ and $\beta_2$ are constants, let alone constants of order unity or a modest fraction of unity. Instead, they must be considered to vary in a way that depends sensitively on the value of the wave supersaturation $(a - 1)$. In the case of $\beta_2$, this supports and strengthens conclusions already drawn by *Chao and Schoeberl* [1984], *Fritts and Dunkerton* [1985], and *Coy and Fritts* [1988] from arguments based on somewhat different models of breaking waves. The present model predicts even greater sensitivity to $(a - 1)$, for reasons to be explained.

For the sake of devising a thought experiment that yields definite answers in as simple and clearcut a manner as possible, while avoiding any dependence on disposable parameters, we sacrifice realism in two ways. First, we replace the real situation, typified by gravity waves arriving from below, by an initial value problem. We imagine an undisturbed, stably stratified atmosphere to which artificial external forces (but no heat sources or sinks) are applied by some hypothetical agency. This agency accelerates the fluid to produce a disturbance velocity field $u'$, and pushes the isentropic (constant-$\theta$) surfaces up or down adiabatically, in the right phase relationship to create a propagating gravity wave field. This will be called the "initial wave field." Of particular interest is a situation in which the initial wave field is supersaturated ($a > 1$) and can therefore be expected to break shortly after being set up. We assume that the breaking takes place by convective overturning.

Second, we make the Boussinesq approximation and assume that the undisturbed static stability, density, and wind are constant with height. We may then (i) go into a frame of reference in which the undisturbed atmosphere is at rest, (ii) consider the wave field to be a periodic plane progressive wave, periodic in the vertical as well as in the horizontal (Figure 1), and (iii) use the fact that such a wave is an exact, finite-amplitude solution to the Boussinesq equations of motion. On the basis of a "perfect mixing" or "convective adjustment" hypothesis, this permits a very simple analysis of the net irreversible transport across isentropic surfaces (the transport that is relevant to photochemical modeling) and its relation to the net turbulent energy dissipation $\mathcal{E}_1 = \int \varepsilon_1 dt$ and wave energy dissipation $\mathcal{E}_2 = \int \varepsilon_2 dt$, integrated over the time of the wave-breaking event. The energy budget will be free of the usual complications arising from moving-medium wave energetics, by (i), free of any ambiguities arising from net energy import or export at boundaries, by (ii), and free of doubts about the applicability of simple wave solutions at finite amplitude, by (iii). Moreover there are no disposable parameters, once the initial wave field is specified.

We further simplify the problem by assuming, as in (5) and (6), that the wave motion is quasi-hydrostatic, with intrinsic frequency $|k(c-\bar{u})| \ll N$. This assumption is not essential, but it is broadly consistent with the observational evidence, and with the idea that the breaking waves have time to undergo convective adjustment. It leads moreover to a useful conceptual simplification: we shall see that, apart from dimensional scaling factors, the only property of the initial wave field on which the results depend is $(a-1)$.

Before proceeding to a consideration of the wave energetics, we recall briefly the relationship between vertical heat transport and potential-energy change in one dimension. The left hand sketch in Figure 2a shows the vertical profiles $\theta(z) = \theta_0(z)$ and $\theta(z) = \theta_0(z) + \Delta\theta(z)$ for the initial and final states of a one-dimensional thought experiment in which an isolated layer of a stably stratified Boussinesq atmosphere is perfectly mixed by some external agency.

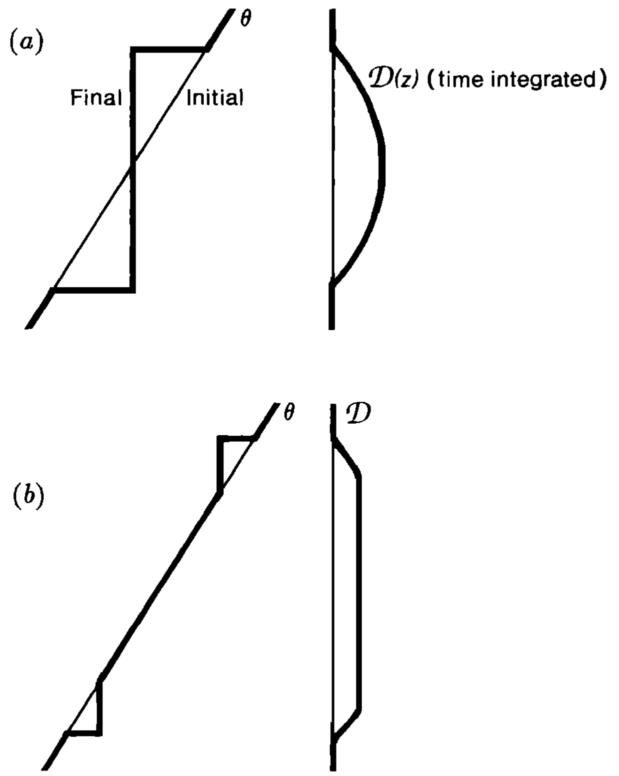


Fig. 2. Two one-dimensional thought experiments to illustrate the connection between vertical heat transport and potential-energy change in a Boussinesq model. Sketches of the vertical profiles of the initial and final potential temperature $\theta$ are shown on the left. The heavy curves on the right show the corresponding vertical profiles of the effective net "diffusivity" $\mathcal{D}$ that can be used to describe the associated heat and mass rearrangement (see text).

It will be useful to write the resulting potential-energy increase in a suggestive form involving an "effective net diffusivity" $\mathcal{D}$ expressing the time-integrated vertical heat flux in terms of the initial, undisturbed stable stratification. (Thus $\mathcal{D}$ has dimensions of length squared, rather than length squared over time.) The undisturbed logarithmic potential-temperature gradient $\theta_0^{-1} d\theta_0/dz = g^{-1}N^2$, where $N$ is the undisturbed buoyancy frequency as before. Within the Boussinesq approximation we may consider $\theta_0$ as well as the density $\rho_0$ to be constant. $\mathcal{D}$ is defined by saying that the fractional potential-temperature change $\Delta\theta/\theta_0$ shall be equal to the convergence of a time-integrated vertical flux $-g^{-1}N^2\mathcal{D}$; that is, $\Delta\theta/\theta_0 = g^{-1}d(N^2\mathcal{D})/dz$. The corresponding change $\Delta P_{\text{col}}$ in total potential energy in each column of unit horizontal area is a well-defined quantity, independent of the origin of $z$, since the one-dimensional mass rearrangement has taken place only within a finite height range, say $z_1 < z < z_2$, and satisfies $\int(\Delta\theta/\theta_0)dz = \theta_0^{-1}\int \Delta\theta dz = 0$. The potential-energy change is therefore, unambiguously,

$$\Delta P_{\text{col}} = -\int_{z_1}^{z_2} \rho_0 g z \frac{\Delta\theta}{\theta_0} dz = -\int_{z_1}^{z_2} \rho_0 z \frac{d}{dz}(N^2\mathcal{D}) dz \quad (10)$$

or on integration by parts

$$\Delta P_{\text{col}} = \int_{z_1}^{z_2} \rho_0 N^2 \mathcal{D} dz \quad (11)$$

per unit horizontal area. The corresponding $\mathcal{D}$ profile is shown on the right of Figure 2a.

Another example is shown in Figure 2b. This shows a layer throughout most of which $\mathcal{D}$, and the time-integrated flux, are constant, causing only the top of the layer to cool and the bottom to heat. Again, the resulting potential-energy change per unit horizontal area is well-defined, and is given equally well by (10) or (11).

Now consider the periodic field of finite amplitude gravity waves and let $\theta'$ and $u'$ be the disturbance potential-temperature and horizontal velocity fields. From here on, it is convenient to average vertically and to reckon energies and energy changes per unit mass. Let $A$ denote the (averaged) available potential energy of the waves per unit mass. $A$ is by definition zero before the initial wave field is set up, when the isentropic surfaces are flat, and is defined such that changes in $A$ are equal to changes in the potential energy $P$ per unit mass as long as the fluid motion is adiabatic. For the present case of constant $N$ the relevant formula is simply

$$A = \frac{g}{\theta_0} \frac{\frac{1}{2}\overline{\theta'^2}}{d\theta_0/dz} = \frac{g^2}{\theta_0^2} \frac{\frac{1}{2}\overline{\theta'^2}}{N^2} \quad (12a)$$

where the overbar denotes the average, defined in terms of the vertical averaging operator

$$\overline{(\ )} = \lim_{h\to\infty} \frac{1}{2h} \int_{-h}^{h} (\ )\, dz$$

The formula (12a) is exact, for finite amplitude waves, as long as the basic state has $N$ constant and the Boussinesq approximation holds. For an elucidation of what is involved in the adiabatic relationship between $A$ and $P$ the reader may consult *Holliday and McIntyre* [1981], *Andrews* [1981] and *McIntyre* [1988]; but ignore the statement about mixing in the first of these references; it is incorrect. The averaged kinetic energy per unit mass is

$$K = \tfrac{1}{2}\overline{u'^2} \tag{12b}$$

We denote the values of $A$ and $K$ in the initial wave field by $A_{\text{init}}$, $K_{\text{init}}$. Thus the work required to set up the initial wave field is $A_{\text{init}} + K_{\text{init}}$ per unit mass. For a sinusoidal plane wave like that depicted in Figure 1, with $c$, $k$ and $m$ all positive, we have, with a suitable choice of axes, $u' = ac\cos\{k(x-ct)-mz\}$, $\theta' = a(\theta_0/g)Nc\sin\{k(x-ct)-mz\}$, and $A_{\text{init}} = K_{\text{init}} = \frac{1}{4}a^2c^2$.

Now imagine that the waves break in such a way that the shaded regions in Figure 1 become vertically well mixed, eliminating all regions of local static instability, and that this occurs in a time very much less than a wave period. The dashed isentropes in Figure 1 (surfaces of constant $\theta_0+\theta'$) show the initial state, before breaking, and the solid isentropes show the state that results from the assumed mixing; see also the left hand diagram in Figure 3, which shows the vertical profiles of $\theta'$. This is the simplest local, mass-conserving, vertical rearrangement that eliminates all statically unstable regions; note that the depth of each mixing region, corresponding to $-z_m < z < z_m$ in Figure 3, is greater than that of the initial statically unstable region.

We also assume that the horizontal momentum $\rho_0 u'$ is mixed perfectly in the same (shaded) regions of Figure 1. For the quasi-hydrostatic waves that we have in mind, this momentum mixing assumption has a modicum of physical plausibility, since in the real, three-dimensional world the overturning would be free to take place in roll-like structures aligned along the shear of the wave motion, in the typical manner of thermal convection in shear. These would have some tendency to homogenize the longitudinal momentum. The assumed vertical profiles of $u'$ before and after mixing are shown in the right hand diagram of Figure 3.

It is not intended, of course, to suggest that these mixing assumptions can give us any more than a highly idealized model of wave breaking, even for quasi-hydrostatic waves. For instance, the turbulent layers might well, in reality, evolve toward a statically stable rather than a neutral state [e.g., *Linden et al.*, 1989], and the momentum might well be substantially rearranged but not homogenized.

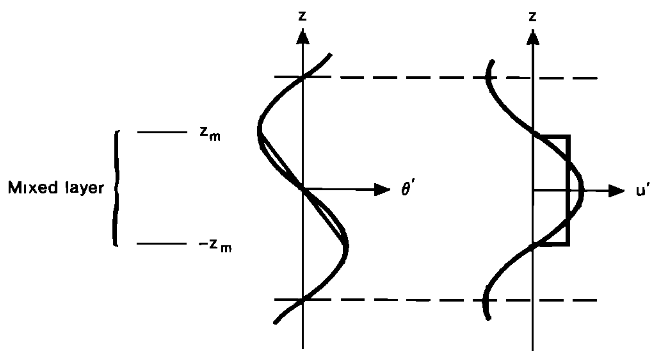


Fig. 3. Sketch of the changes in the $z$ profiles of the disturbance potential temperature $\theta'$ and horizontal velocity component $u'$ assumed in the wave-breaking thought experiment, at $x - ct = 0$. The straight line segments cut off equal areas, to conserve heat and momentum.

Quantitatively accurate modeling would demand, in reality, high-resolution numerical simulations of the fully three-dimensional turbulent problem, a task that is well beyond the scope of available computational resources even today, requiring a multitude of cases to be run in the likely event that the detailed mixing characteristics proved sensitive to noise in the initial conditions. The present model, while not pretending to be quantitative, has at least the virtue of simplicity and comprehensibility and, as already mentioned, no disposable parameters. Once the foregoing assumptions have been made, the values of several quantities of interest, including $\mathcal{E}_1$, $\mathcal{E}_2$ and the net vertical heat transport, are determined precisely and unambiguously as functions of $(a-1)$.

Denote the magnitudes of the decreases in $A$, $K$, and $P$ due to the profile changes shown in Figure 3 by $\Delta A$, $\Delta K$ and $\Delta P$, all positive by definition. $\Delta P$ is defined in the same way as in the first integral in (10), apart from the factor $\rho_0$, the sign convention, and the fact that the integral is replaced by the vertical averaging operator. $\Delta P$ is well defined for the same reason as was $\Delta P_{\text{col}}$: the vertical integral of the change $\Delta\theta'$ in $\theta'$ vanishes over each mixing layer, so that the contribution to $\Delta P$ from each layer is independent of the $z$ origin.

We may summarize the energy changes as follows. Before the experiment, the isentropes are flat, and $A = 0$, $K = 0$, and $P = P_0$, say. (The potential energy $P_0$ of the initial undisturbed state is itself ambiguous, but that is immaterial since only changes in energy are of interest.) After wave generation, $A = A_{\text{init}}$, $K = K_{\text{init}}$, and $P = P_0 + A_{\text{init}}$, and after mixing $A = A_{\text{init}} - \Delta A$, $K = K_{\text{init}} - \Delta K$, and $P = P_0 + A_{\text{init}} - \Delta P$. Note that we must expect $\Delta A$ to differ from $\Delta P$ because the mixing is a diabatic process. By definition, the net wave energy dissipation

$$\mathcal{E}_2 = \Delta A + \Delta K \tag{13}$$

still reckoning all energy changes per unit mass.

Finally, we imagine that after the waves break, the hypothetical agency conducting the experiment intervenes again and carefully removes all of the remaining wave motion adiabatically (which results in work being done upon the agency, although less work than it originally did in order to create the waves). This is the closest we can come in our thought experiment to mimicking a real situation in which the waves propagate out of the region of interest, leaving the isentropes once again completely flat. Then $A$ and $K$ decrease by amounts $A_{\text{init}} - \Delta A$ and $K_{\text{init}} - \Delta K$ respectively (taking $A$ and $K$ back to zero), while $P$ decreases by the same amount as $A$, since this final stage is adiabatic. Thus $P$ decreases by $A_{\text{init}} - \Delta A$, so that

$$\begin{aligned} P &= P_0 + A_{\text{init}} - \Delta P - (A_{\text{init}} - \Delta A) \\ &= P_0 + \Delta P_{\text{transp}} \end{aligned}$$

say, where

$$\Delta P_{\text{transp}} = \Delta A - \Delta P \tag{14a}$$

Since $\Delta P_{\text{transp}}$ is the net potential-energy increase, after the isentropes have been distorted then adiabatically flattened again, it is exactly the potential-energy increase associated with the true, irreversible, net cross-isentropic transport due to wave breaking. By analogy with (11), the corresponding "effective net diffusivity" $\mathcal{D}$ is

$$\mathcal{D} = \Delta P_{\text{transp}}/N^2 = (\Delta A - \Delta P)/N^2 \tag{14b}$$

The agency that set up and removed the waves has done net work $\Delta A + \Delta K = \mathcal{E}_2$ per unit mass over the whole experiment. This work must have been partitioned between $\Delta P_{\text{transp}}$ and the net turbulent energy dissipation $\mathcal{E}_1$ per unit mass, which is therefore exactly

$$\mathcal{E}_1 = \Delta A + \Delta K - \Delta P_{\text{transp}} = \Delta P + \Delta K \tag{15}$$

The quantities analogous to $\beta_1$ and $\beta_2$ can now be defined as

$$\mathcal{B}_1 = \Delta P_{\text{transp}}/\mathcal{E}_1 \tag{16a}$$

and

$$\mathcal{B}_2 = \Delta P_{\text{transp}}/\mathcal{E}_2 \tag{16b}$$

since we then have, from (14*b*),

$$\mathcal{D} = \mathcal{B}_1\mathcal{E}_1/N^2 \qquad \mathcal{D} = \mathcal{B}_2\mathcal{E}_2/N^2$$

the model's counterparts to (1) and (3). Notice that we also have

$$\mathcal{D} = (\mathcal{E}_2 - \mathcal{E}_1)/N^2 \tag{17}$$

In order to evaluate these quantities explicitly, let us invoke our assumption of a sinusoidal initial wave and, for convenience, take units such that $\rho_0 = 1$, $N^2 = 1$, $c = 1$ and therefore $m = 1$, that is, vertical wavelength $= 2\pi$, using the first of (5) with $\bar{u} = 0$. In these units, the initial profiles in Figure 3 give

$$A_{\text{init}} = \frac{1}{\pi}\int_0^\pi \frac{\frac{1}{2}g^2\theta'^2}{\theta_0^2}dz = \frac{1}{2\pi}\int_0^\pi a^2\sin^2 z\,dz = \tfrac{1}{4}a^2 \tag{18a}$$

$$K_{\text{init}} = \frac{1}{\pi}\int_0^\pi \tfrac{1}{2}u'^2 dz = \frac{1}{2\pi}\int_0^\pi a^2\cos^2 z\,dz = \tfrac{1}{4}a^2 \tag{18b}$$

To recover the corresponding dimensional quantities, which have dimensions of energy per unit mass, multiply by $c^2$. From the left hand part of Figure 3, we see that the half-depth of the mixed layer, $z_{\text{m}}$, a function of $a$, is the smallest positive root of

$$a\sin z = z \tag{19}$$

(Note again that this corresponds to a turbulent layer that is deeper than the initial statically unstable layer.) At $x = 0$, then, the mixed layer is $-z_{\text{m}} < z < z_{\text{m}}$, and the dimensionless $\theta'$ profile is $-a\sin z$ before mixing and $-z$ within the layer after mixing. Thus, with our sign conventions,

$$\Delta P = \frac{1}{\pi}\int_0^{z_{\text{m}}} -z(z - a\sin z)dz$$
$$= \pi^{-1}\{-\tfrac{1}{3}z_{\text{m}}^3 + a(\sin z_{\text{m}} - z_{\text{m}}\cos z_{\text{m}})\} \tag{20}$$

$$\Delta A = \frac{1}{2\pi}\int_0^{z_{\text{m}}} -(z^2 - a^2\sin^2 z)dz$$
$$= (2\pi)^{-1}\{-\tfrac{1}{3}z_{\text{m}}^3 + \tfrac{1}{2}a^2(z_{\text{m}} - \sin z_{\text{m}}\cos z_{\text{m}})\} \tag{21}$$

$$\Delta K = \frac{1}{2\pi}\int_0^{z_{\text{m}}} -(1 - a^2\cos^2 z)dz$$
$$= (2\pi)^{-1}\{-z_{\text{m}} + \tfrac{1}{2}a^2(z_{\text{m}} + \sin z_{\text{m}}\cos z_{\text{m}})\} \tag{22}$$

and, as a partial check,

$$\Delta P_{\text{transp}} = \Delta A - \Delta P = \frac{1}{2\pi}\int_0^{z_{\text{m}}} (-z + a\sin z)^2 dz$$

$$(> 0, \text{irreversible!})$$

$$= (2\pi)^{-1}\{\tfrac{1}{3}z_{\text{m}}^3 + 2a(z_{\text{m}}\cos z_{\text{m}} - \sin z_{\text{m}}) + \tfrac{1}{2}a^2(z_{\text{m}} - \sin z_{\text{m}}\cos z_{\text{m}})\} \tag{23}$$

These quantities are functions of $a$ alone, after taking (19) into account. Multiplication by $c^2$ yields the corresponding dimensional energies per unit mass. Graphs of $\Delta P$, $\Delta A$, $\Delta K$ and $\Delta P_{\text{transp}}$ are plotted in Figure 4 along with $A_{\text{init}}$, $K_{\text{init}}$, from which we gain an immediate impression of the sensitivity to the value of $(a - 1)$.

It is not hard to show, by Taylor-expanding (20)–(23), that $\Delta A$, $\Delta K$ and $\Delta P$ all behave like $(a-1)^{5/2}$, and that $\Delta P_{\text{transp}}$ behaves like $(a-1)^{7/2}$, as $(a-1) \downarrow 0$. To show this it is easiest to Taylor-expand the integrands first and then to integrate the leading order terms using $z_{\text{m}} \sim \{6(a-1)\}^{1/2}$, the latter being obtained by expanding (19), except that for $\Delta K$ it is easiest first to factorize the integrand as $-(1 - a\cos z)(1 + a\cos z)$ and then expand the right hand factor only, noting that the integral of the left hand factor vanishes. This analysis can be extended to show that these power laws would remain the same for other physically plausible rearrangements of the $\theta'$ and $u'$ fields, such as the imperfect vertical mixing envisaged three paragraphs above (13), assuming only that the mixing involves changes in $\partial\theta'/\partial z$ of order $(a-1)$ and jumps in $\theta'$ of order $(a-1)^{3/2}$ or less, consistent with $z_{\text{m}} \propto (a-1)^{1/2}$ for small $(a-1)$. Details are omitted for brevity.

Values of $\mathcal{B}_1$ and $\mathcal{B}_2$ are shown in Table 1; again, the great sensitivity to $(a-1)$ is evident. The bottom half of Table 1 probably has little physical interest; one wonders how often waves are likely to reach $(a-1) > 1$, say, before mixing becomes effective in limiting $a$, except perhaps for the most violent breaking of high frequency waves with large vertical wavelengths, in "steep beach" conditions further up in the lower thermosphere.

We may also ask what the model has to say about various "turbulent Prandtl numbers." If, for example, we were to define the effective net momentum "diffusivity" $\mathcal{D}_{\text{mw}}$ for the time-integrated wave dissipation as $\mathcal{D}_{\text{mw}} \equiv \Delta K/2K_{\text{init}}$ (in rough analogy with the fact that for $m = 1$ the contribution from ordinary molecular viscosity $\nu$ to the rate of change of wave kinetic energy $K$ is $2\nu K$), then the "Prandtl number" $\mathcal{D}_{\text{mw}}/\mathcal{D}$ takes the values shown in the fourth column of Table 1. Again, the values vary wildly. The last column compares $\mathcal{D}$ with the effective net heat "diffusivity" $\mathcal{D}_{\text{hw}}$ for wave dissipation, defined in the corresponding way as $\mathcal{D}_{\text{hw}} \equiv \Delta A/2A_{\text{init}}$. This illustrates

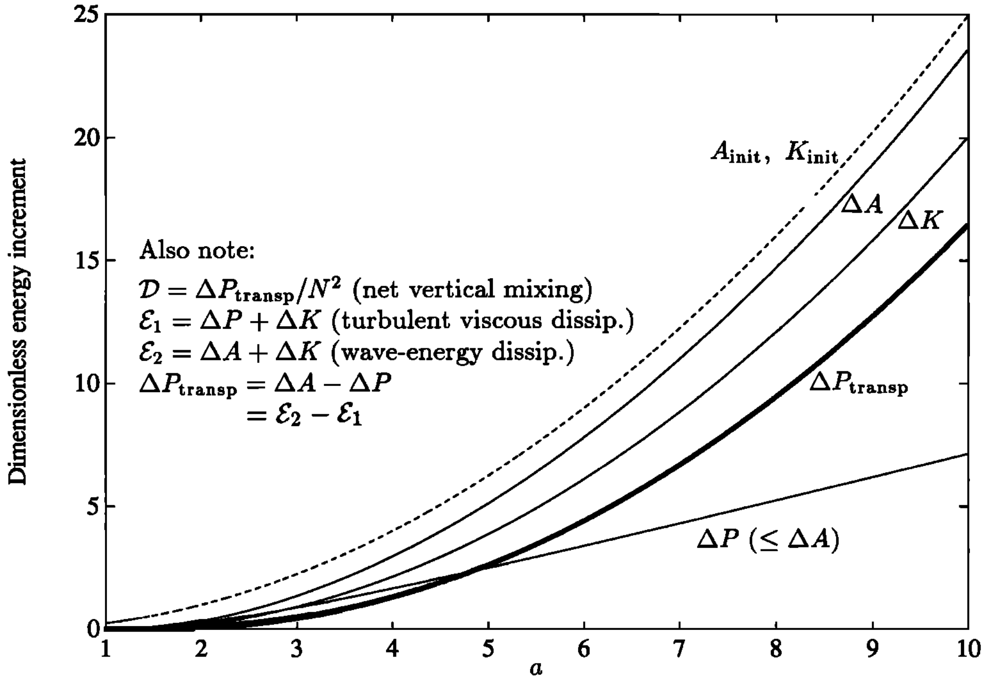


Fig. 4. Graphs of the dimensionless energy changes in the gravity wave breaking thought experiment, for different values of the dimensionless wave amplitude $a$. The heavy curve is a measure of the net vertical transport by mixing (see text). Realistic values are probably confined near the bottom left hand corner; a wider range is shown only to convey the functional character of equations (18)–(23), and as a check on consistent asymptotic behavior.

the point that "the" effective eddy diffusivity of a given quantity, in this case heat, may differ greatly according to whether it is conceived of as applying to the waves or to the basic state. For large $a$, for instance, the effective diffusivities for the waves are far smaller than those for the basic state, because the turbulent fluxes are comparable but wave-field gradients far exceed basic state gradients.

It might be thought, incidentally, that the model could be used to approach the difficult problem of backscattering from a region of wave breaking. There is no theoretical reason, of course, why a wave-breaking region should not backscatter the incident waves, although the present author is not aware of any case study where this has been unequivocally observed in the mesosphere. There are hints of it in the momentum-flux results reported by *Reid and Vincent* [1987], and there is good numerical modeling evidence in some lower-atmospheric examples involving strong mountain waves [*Clark and Peltier*, 1984]. The counterpart of this backscattering in our idealized model is the presence of Fourier components with downward group velocities in the modified wave field implied by the final profiles in Figure 3. It is curious, albeit probably not very significant, that the fundamental (wavelength $2\pi$) backscattered component vanishes, although the higher harmonics do not. The point is not pursued here because it seems too model-dependent to be of real interest.

Perhaps the most important suggestion from this analysis, even more striking than suggestions from the earlier analyses by *Chao and Schoeberl* [1984], *Walterscheid* [1984], *Fritts and Dunkerton* [1985], *Schoeberl* [1988], and *Coy and Fritts* [1988] is that waves that are not breaking too violently (corresponding to values near the top of Table 1 or the bottom left of Figure 4) will produce only weak vertical mixing. It should be weak in the sense that the dimensionless numbers $\beta_1$ and $\beta_2$ in (1) and (3) are likely to be well below unity, as are their counterparts $\mathcal{B}_1$ and $\mathcal{B}_2$ near the top of Table 1. The present model suggests this even more forcefully than its predecessors. The main difference between the present model on the one hand, and the model of Fritts and Dunkerton and Coy and Fritts on the other, for example, is that there the depth of the turbulent layer is treated as an independent parameter (related to their parameter $n$), whereas here its dependence on $(a-1)$ is built into the model, albeit in an idealized way. The present model, in other words, is based on the expectation that the turbulent layer depth $2z_m$ should decrease as $a$ decreases toward unity, and it is this that accounts for the present model's more sensitive dependence on $(a-1)$ when $(a-1)$ is small. An implicit assumption is that the static stability elsewhere in the wave field can rapidly suppress turbulence.

TABLE 1. Model Counterparts of $\beta_1$, $\beta_2$ and Two Effective Diffusivity Ratios, against Wave Supersaturation $(a-1)$.

| $(a-1)$ | $\mathcal{B}_1$ | $\mathcal{B}_2$ | $\mathcal{D}_{mw}/\mathcal{D}$ | $\mathcal{D}_{hw}/\mathcal{D}$ |
|---|---|---|---|---|
| 0.01 | 0.0019 | 0.0019 | 345.3 | 688.7 |
| 0.1 | 0.0186 | 0.0182 | 30.8 | 59.9 |
| 0.25 | 0.0448 | 0.0429 | 10.4 | 19.5 |
| 0.5 | 0.0848 | 0.0782 | 4.1 | 7.3 |
| 1.0 | 0.154 | 0.133 | 1.4 | 2.3 |
| 3.0 | 0.345 | 0.257 | 0.20 | 0.28 |
| 5.0 | 0.464 | 0.317 | 0.077 | 0.098 |
| 9.0 | 0.606 | 0.377 | 0.024 | 0.029 |
| $\infty$ | 1 | 0.5 | 0 | 0 |

The sensitivity of all these quantities to $(a-1)$ is apparent. The last two columns (see text) suggest some pitfalls inherent in the notion of "turbulent Prandtl number."

Further evidence on the nature of convective instability in breaking gravity waves has recently come from two-dimensional numerical experiments by R. L. Walterscheid and G. Schubert (Nonlinear evolution of an upward propagating gravity wave: overturning, convection and turbulence, submitted to the *Journal of the Atmospheric Sciences,* 1988). As far as can be seen from a small number of cases, the general character of the breaking waves in their model is is broadly consistent with the picture assumed here. There is, as yet, no check from a fully three-dimensional simulation, and no clear discrimination between the present model and its predecessors.

The wider implications of the present results are daunting. The sensitivity to $(a-1)$ illustrated in Table 1 suggests that in trying to understand the real atmosphere, particularly the lower thermosphere and the winter mesosphere, where waves may well break violently, an acute problem will be how to estimate effective values of $(a-1)$. These values will be sensitive to statistical properties, such as the intermittency of breaking in a realistic wave field, which will be sensitive in turn to the occurrence of constructive or destructive interference and hence to phase relations between the Fourier components in the spectrum of waves. For instance, one might get vertical mixing less weak than a simple statistical measure of $(a-1)$ might suggest, if it happened that the spectral phase relations were such as to imply relatively rare, but violent, wavebreaking events.

In the case of the summer mesosphere and NLC/PMC modeling, however, it seems likely that the weakness of vertical mixing for modest values of $(a-1)$ is the relevant consideration. We return to this question after recalling what controls the other transport mechanism that has to be considered, namely, the mean upwelling.

## 3. Upwelling in Zonally Symmetric, Quasi-Steady Models

It is now widely accepted that mean upwelling in the summer high-latitude mesosphere is a necessary part of the global scale mesospheric circulation, without which it would be difficult, if not impossible, to make sense of the observed facts, including the remarkable coldness of the summer mesopause. The summer mesopause is the coldest part of the entire atmosphere, despite a net absorption both of infrared radiation from below and of solar radiation from above. Adiabatic cooling by sustained vertical motion seems likely to be the only mechanism capable of producing the observed cold temperatures.

Dynamical considerations show that a drag force, having a certain latitudinal gradient, is needed to drive this upwelling. The force, of whose manner of functioning and domain of influence we shall be reminded shortly, is believed to come mainly from the systematic, irreversible momentum transport associated with upward propagating gravity waves directionally filtered by the mean wind structure [*Lindzen,* 1981]. Besides the theoretical robustness, and hence likely ubiquity, of this momentum transport effect (e.g., M. E. McIntyre and W. A. Norton, and references therein, Dissipative wave-mean interactions and the transport of vorticity or potential vorticity, submitted to the *Journal of Fluid Mechanics,* G. K. Batchelor Festschrift Issue, 1989), there is an accumulation of independent observational evidence for the existence of the required gravity waves in the real atmosphere [e.g., *Balsley et al.,* 1983; *Fritts,* 1984; *Reid,* 1986; *Chanin and Hauchecorne,* 1987; *Reid and Vincent,* 1987; *Röttger,* 1987; *Fritts et al.,* 1988; *Manson and Meek,* 1988; *Fritts and VanZandt,* 1989, and references therein], including some direct measurements of wave-induced momentum fluxes.

As a preliminary to discussing the real summer mesosphere, we recall briefly how the wave-driven upwelling mechanism works in a simplified context, namely, that of zonally symmetric model mesospheric circulations that vary on a seasonal time scale only, being driven by a zonally symmetric, zonally directed wave-drag force $\bar{\mathcal{F}}$ per unit mass that varies on a similarly slow time scale. An important recent model study of this kind is reported by *Garcia* [this issue]; see also the review material and historical notes in the book by *Andrews et al.* [1987].

As might be anticipated from the fact that the seasonal time scale greatly exceeds the relevant radiative relaxation time scales (e.g., the references just cited), the circulation driven by the force $\bar{\mathcal{F}}$ in these model mesospheres is quasi-steady, in the sense that the zonally averaged, or prima facie zonally symmetric, wind and temperature tendencies $\partial\bar{u}/\partial t$ and $\partial\bar{T}/\partial t$ can be neglected in the governing equations. In these circumstances the nature of the dynamical control is particularly simple to analyze. To a first approximation, the extratropical angular momentum balance takes the steady state form

$$-f\bar{v}^* = \bar{\mathcal{F}}(y,z) \tag{24}$$

and the mean meridional and vertical velocities $\bar{v}^*, \bar{w}^*$ satisfy a zonally symmetric mass-conservation equation of the form

$$\frac{\rho_0(z)}{\cos\varphi}\frac{\partial(\bar{v}^*\cos\varphi)}{\partial y} + \frac{\partial(\rho_0\bar{w}^*)}{\partial z} = 0 \tag{25}$$

where $y$ is latitude $\varphi$ times the Earth's radius and $f(y)$ is the Coriolis parameter. Here $\bar{v}^*$ and $\bar{w}^*$ are defined so as to be as directly relevant as practicable to the transport of constituents; one may use, for instance, the TEM or "transformed Eulerian-mean" equations with log pressure as the vertical coordinate $z$, suitably scaled, or even better the Eulerian-mean equations in isentropic coordinates with log entropy as the vertical coordinate [e.g., *Tung,* 1982; *Andrews et al.,* 1987]. The steady state mass-conservation equation has the same form (25) in either case, with suitable definitions. See also the discussion of "transport circulation" in *Plumb and Mahlman* [1987].

The mean density $\rho_0(z)$ falls off exponentially with altitude $z$. From (24) and (25) it follows in two or three lines of manipulation that the upwelling velocity $\bar{w}_0^*$ at a given altitude $z_0$ is related to the drag force $\bar{\mathcal{F}}$ by

$$\bar{w}_0^*(y,z_0) = -\frac{1}{\rho_0(z_0)\cos\varphi}\frac{\partial}{\partial y}\left\{f^{-1}(y)\cos\varphi\int_{z_0}^{\infty}\rho_0(z)\bar{\mathcal{F}}(y,z)dz\right\} \tag{26}$$

the integral being safely convergent at its upper limit because of the factor $\rho_0$ and the constraints on the magnitude of $\bar{\mathcal{F}}$ imposed by wave breaking. (Dimensionless amplitudes $a$ may well exceed unity, but as already discussed they are unlikely to do so by orders of magnitude; and so for gravity waves $|\bar{\mathcal{F}}| \lesssim \varepsilon_2/|\bar{u} - c| \sim w_g|\bar{u} - c|/H_{diss}$ in order of magnitude, using (4)ff., which is unlikely to increase upwards faster than $\rho_0$ decreases). A more accurate version of (26), based on a more accurate version of (24), is presented by P. H. Haynes, C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine (On the downward control principle for extratropical diabatic circulations, submitted to the *Journal of the Atmospheric Sciences*, 1989); that paper also presents an independent verification of (26) in the form of time dependent analytical and numerical calculations showing that the steady state described by (26) is indeed approached, how it is approached, and how the rate at which it is approached depends on spatial scales and radiative relaxation time scales.

The relation (26) highlights the simple but interesting fact that the quasi-steady upwelling across a given level, for example the NLC/PMC level, is controlled from above, in the sense that it depends exclusively on $\bar{\mathcal{F}}$ and its latitudinal gradient above that level. This principle of "downward control" by $\bar{\mathcal{F}}$ takes an even simpler form if the waves can be assumed to propagate straight upward. This is true to good approximation for those gravity waves that are believed to dominate $\bar{\mathcal{F}}$ in the mesosphere. The waves in question have fast vertical group velocities and periods of the order of an hour or less [e.g., *Reid*, 1986; *Reid and Vincent*, 1987]. Then

$$\bar{\mathcal{F}} \simeq -\rho_0^{-1}\frac{\partial}{\partial z}(\rho_0\overline{u'w'}) \qquad (27)$$

where $\rho_0\overline{u'w'}$ is the radiation stress or wave-induced momentum flux in the usual approximation (and notational convention) appropriate to high-frequency gravity waves. With these approximations, the integration in (26) can be carried out explicitly, giving

$$\bar{w}_0^*(y,z) = -\frac{1}{\cos\varphi}\frac{\partial}{\partial y}\{f^{-1}(y)\cos\varphi\ \overline{u'w'}\} \qquad (28)$$

To this approximation, then, the quasi-steady upwelling across a given level $z$ depends only on the latitudinal distribution of the net upward gravity wave flux across the level $z$ and not, for instance, on the height at which the waves break. This behavior is strikingly illustrated by Figure 5, taken from *Garcia* [this issue]. Three profiles of $\bar{w}^*(z)$ are shown, from three zonally symmetric model simulations, in which everything in the mesospheric circulation regime varies except the net upward flux of gravity waves incident from below, as measured by $\rho_0\overline{u'w'}$. Zonal wind profiles, mesopause altitudes, mesopause temperatures, and wave-breaking threshold altitudes all differ from experiment to experiment but, as Figure 5 illustrates, the upwelling velocity profile is the same to good approximation in all three experiments at those altitudes lying below the region where the waves are breaking, that is below about 87 km. The model is behaving as (28) predicts.

From the present perspective one might summarize the key properties of the mesospheric circulations found in Garcia's model study as follows:

1. The upwelling at a given level is controlled by the gravity wave drag above that level, and hence by the net wave-induced momentum flux across that level.

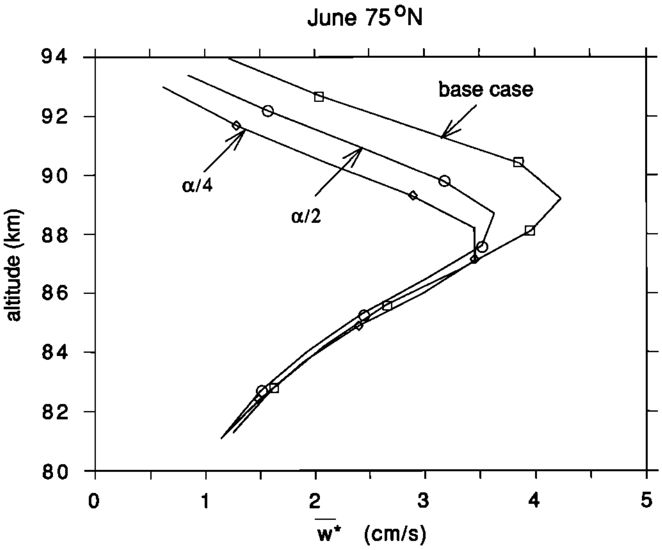


Fig. 5. Upwelling profiles from three zonally symmetric model experiments reported by *Garcia* [this issue], illustrating the "downward control" principle in the form expressed by equations (26) and (28). The mesospheric circulation regime is changed by varying the radiative relaxation time $\tau_{rad}$ while keeping other external parameters unchanged, including the parameterized upward flux of gravity waves from the troposphere. The curve marked "base case" is for $\tau_{rad}$ = 4 days, the next curve down (marked by circles) is is for $\tau_{rad}$ = 8 days, and the remaining curve is for $\tau_{rad}$ = 16 days. Mesopause altitudes and temperatures and wave breaking threshold altitudes all decrease as $\tau_{rad}$ increases.

2. The wave breaking and drag occur largely above NLC/PMC altitudes. (It is worth recalling that this is always made more likely, other things being equal, by the region of increased $N^2$ values above the mesopause. All gravity-wave theories predict an increase in the amplitude $a$, in the sense defined in section 1, hence an increase in the likelihood of wave breaking, when waves propagate upward into such a region [*Hauchecorne et al.*, 1987; *Fritts and VanZandt*, 1989]. The last term in (8) correctly indicates this tendency, albeit only qualitatively because of the ray theoretic and other approximations on which (8) is based.)

3. Below wave breaking altitudes, and therefore at and below NLC/PMC altitudes, the mean upwelling $\bar{w}^*$ on the seasonal time scale has an altitude dependence given to a first approximation by $\bar{w}^* \propto \rho_0(z)^{-1}$. (This can be seen at once from (26), since the quantity in braces becomes independent of $z_0$ when $z_0$ lies below wave breaking altitudes. It can also be seen from (28), using the fact that (27) is zero below wave breaking altitudes.)

4. Below wave breaking altitudes, and at and below NLC/PMC altitudes, horizontal advection can be neglected in comparison with vertical advection. (Statement 3 and equation (25) imply that $\bar{v}^* = 0$ below wave breaking altitudes.)

5. Below wave breaking altitudes, and at and below NLC/PMC altitudes, vertical eddy transport of heat and constituents is negligible. (This is true even in the model experiments that use the Lindzen diffusivity formula (3), according to *Garcia* [this issue], and therefore true a fortiori when the smaller values of $D$ indicated in section 2 are used.)

## 4. A TRANSPORT MODELING STRATEGY FOR THE REAL SUMMER MESOSPHERE

Can we treat the five statements just made as hypotheses that might be approximately applicable to the real summer polar mesosphere, for example at a high latitude location where rocket soundings have been taken, as in the Cold Arctic Mesopause Project [e.g., *Kopp et al.*, 1985]? If the answer were yes, then it would follow that the use of one-dimensional modeling would be strictly justifiable – provided that vertical advective transport is incorporated, with an upwelling profile of the form $\bar{w}^*(z) \propto \rho_0(z)^{-1}$, and provided that the vertical eddy diffusivity is made zero or, failing that, as small as numerical model stability permits. Such a model would be attractive as a practical research tool in that the assumed transport is physically reasonable, simple to understand and to formulate, closely consistent with two-dimensional model studies like Holton's and Garcia's, as Figure 5 illustrates, and described by a single parameter only, namely, the proportionality constant giving the strength of the upwelling. It may also be qualitatively reasonable outside the bounds of strict validity of the relations (26) and (28), as we shall indicate. The idea of introducing vertical advection into a one-dimensional photochemical model, for application to a particular geographical region like the polar cap, is not new; the usefulness of such a model has been demonstrated, for example, in the different context of understanding ozone evolution in the Antarctic stratosphere in late summer, by *Farman et al.* [1985].

In the present state of knowledge, however, it is unfortunately not possible to give a definitive answer to the question just posed. What does seem possible at present is to argue that the five statements, hereafter referred to as hypotheses 1–5, can reasonably be made the basis of a one-dimensional modeling strategy on the Occam's Razor principle. Indeed, with the possible exception of tidal effects (see *Jensen et al.* [this issue]), we shall argue that it would be difficult to justify the use of more elaborate models until far more knowledge of the three-dimensional, time-dependent mesospheric gravity wave field and circulation becomes available.

Hypothesis 5 is probably the most robust. It is supported both by direct observation (C. R. Philbrick, personal communication, 1988) and by the results of theoretical analyses like that of section 2 above. All the indications seem to be that even though we must expect some gravity wave breaking at and below NLC/PMC altitudes, the breaking is typically weak. Together with the theory of section 2 and its predecessors, this strongly suggests that we can neglect vertical mixing altogether, to an excellent approximation at and below those altitudes.

The status of the other hypotheses is less certain. Perhaps the greatest two uncertainties are first the role of transient fluctuations in the mean circulation, on time scales less than the time scale for adjustment to a downward controlled circulation, and second the role of zonal asymmetry of the mean circulation, where "mean" no longer signifies a zonal average, but instead a local average over a few gravity wave periods.

Regarding the first uncertainty, we know from ground-based observations that NLC displays may fluctuate on time scales less than a day, much faster than downward-control adjustment times. The same is true of PMC [e.g., *Olivero and Thomas*, 1986]. There is, of course, every reason to suppose that fast dynamical processes like gravity wave generation will themselves fluctuate on similarly short time scales, hence $\bar{\mathcal{F}}$ itself and the time-dependent response to it. Downward-control adjustment times for "large scale" circulations (height and length scales $H \gtrsim H_\rho$ and $L \gtrsim NH_\rho/f$) are of the order of the radiative relaxation time $\tau_{\text{rad}}$, and are longer for smaller scale circulations. Thus although the background large scale seasonal mean upwelling is almost certainly downward controlled, conforming to (26) and (28) to good approximation, short-term fluctuations in $\bar{w}^*$ need not be. On the other hand, the downward control principle has some qualitative validity outside its strict domain of applicability. It has validity to the extent that even for time scales $\ll \tau_{\text{rad}}$ at least half the mass circulation fluctuation induced by time dependence in $\bar{\mathcal{F}}$ will always be directed downward, and more like three quarters for the largest scales (more than 84% downward, for example, in the case of the second Hough mode in a 239K isothermal atmosphere with $H_\rho = 7$km and Rossby height $fL/N = 13$km). Further details are given in the paper by Haynes et al., referred to below equation (26).

Regarding the second uncertainty, similar remarks apply, but with additional effects much harder to quantify. It is hardly likely that the gravity wave drag force is either zonally directed or zonally symmetric, when one thinks of current ideas about gravity wave sources and gravity-wave filtering, to say nothing of the various possible interactions of gravity waves with tides [e.g., *Walterscheid*, 1982; *Fritts and Vincent*, 1987]. Thus it is hardly likely that the upper mesospheric circulation, which is believed to be so powerfully affected by the wave drag force, would be zonally symmetric either. However, it is conceivable that the circulation might nevertheless be more nearly symmetric than the force distribution itself, especially at the lower altitudes. Zonal winds can carry the air past geographically fixed, longitudinal nonuniformities in the force distribution and its induced circulation, so that in effect there might be some zonal averaging of the forcing in the lower part of the relevant altitude range. An additional effect of zonal asymmetry would be to contribute directly to the transient mean circulations that are superposed on the large scale, seasonal-mean circulation, the material derivative $D/Dt$ now entering the circulation dynamics in place of the partial derivative $\partial/\partial t$. Transient vertical motion would be induced by the relative advection of isentropic anomalies of potential vorticity, in just the same way as for the synoptic-scale vertical motion associated with tropospheric weather systems [e.g., section 4 of *Hoskins et al.*, 1985].

What emerges from the foregoing considerations is that, with the possible exception of tidal contributions, quantification of the actual local, transient vertical motion would require a state of the art in mesospheric observation and modeling comparable to that currently attained in operational numerical weather forecasting, in addition to which one would require synoptic, detailed observations of the time-dependent gravity wave fluxes $\rho_0(\overline{u'w'},\ \overline{v'w'})$ in three dimensions. This is a tall order, even for the next generation of space-based wind sensors. We are forced to the conclusion that for modeling purposes it is difficult to imagine doing better, at present, than provisionally adopting hypotheses 1–5 on the Occam's Razor principle. The resulting transport model is arguably far closer to physical reality in the summer polar mesosphere than one-dimensional models assuming an arbitrary vertical diffusivity function $D(z)$ but no upwelling. Moreover, a transport model having only a single disposable parameter, the proportionality constant for the upwelling velocity profile $\propto \rho_0(z)^{-1}$, would provide a more testing basis for comparison with observation than a model having all the disposable degrees of freedom associated with the arbitrary function $D(z)$.

In summary, it seems clear that the suggested modeling strategy, whose merits would of course be subject to tests against observation and, eventually, against more comprehensive observations and three-dimensional modeling, has significant attractions in the present state of knowledge. Not least among these is its avoidance of the problem posed by the eddy diffusivity concept itself, recalled at the beginning of section 2, namely the fact that that concept appears difficult to justify, in any fundamental way, under conditions typical of the real atmosphere.

*Acknowledgements.* This paper grew out of an invited presentation to the International NLC Workshop held at Boulder, Colorado in March 1988. I wish to thank the organizer, G. E. Thomas, and D. G. Andrews, M. L. Chanin, D. C. Fritts, R. R. Garcia, P. H. Haynes, A. Hauchecorne, E. Jensen, J. J. Olivero, C. R. Philbrick, G. E. Reid, G. Schubert, S. Solomon, R. A. Vincent, R. L. Walterscheid, and J. Weinstock for helpful comments and correspondence. R. R. Garcia kindly provided a copy of Figure 5 of his paper prior to publication in this issue. I wish also to thank Simon Cooper, Margaret Downing, Gail Money, and Clara Wong for gallant and generous technical support at the eleventh hour, and the Atmospheric Sciences Department of the University of California at Los Angeles and particularly C. R. Mechoso for kind hospitality during the writing of the paper. This work was supported by the U. S. Office of Naval Research.

## REFERENCES

Andrews, D. G., A note on potential energy density in a stratified compressible fluid, *J. Fluid Mech.*, *107*, 227-236, 1981.

Andrews, D. G., J. R. Holton, and C. B. Leovy, *Middle Atmosphere Dynamics*, 489 pp., Academic, San Diego, Calif., 1987.

Balsley, B. B., W. L. Ecklund, and D. C. Fritts, VHF echoes from the high-latitude mesosphere and lower thermosphere, *J. Atmos. Sci.*, *40*, 2451-2466, 1983.

Batchelor, G. K., and A. A. Townsend, Turbulent diffusion, in *Surveys in Mechanics*, edited by G. K. Batchelor and R. M. Davies, pp. 352-399, Cambridge University Press, New York, 1956.

Chanin, M. L., and A. Hauchecorne, Lidar sounding of the structure and dynamics of the middle atmosphere: A review of recent results relevant to transport processes, in *Transport Processes in the Middle Atmosphere*, edited by G. Visconti and R. R. Garcia, pp. 459-477, D. Reidel, Hingham, Mass., 1987.

Chao, W. C., and M. R. Schoeberl, On the linear approximation of gravity wave saturation in the mesosphere, *J. Atmos. Sci.*, *41*, 1893-1898, 1984.

Charlson, R. J., Noctilucent clouds: A steady-state model, *Q. J. R. Meteorol. Soc.*, *91*, 517-523, 1965.

Clark, T. L., and W. R. Peltier, Critical level reflection and the resonant growth of nonlinear mountain waves, *J. Atmos. Sci*, *41*, 3122-3234, 1984.

Coy, L., and D. C. Fritts, Gravity wave heat fluxes: A Lagrangian approach, *J. Atmos. Sci.*, *45*, 1770-1780, 1988.

Farman, J. C., R. J. Murgatroyd, A. M. Silnickas, and B. A. Thrush, Ozone photochemistry in the Antarctic stratosphere in summer, *Q. J. R. Meteorol. Soc.*, *111*, 1013-1025, 1985.

Fritts, D. C., Gravity wave saturation in the middle atmosphere: A review of theory and observations, *Rev. Geophys.*, *22*, 275-308, 1984.

Fritts, D. C., and T. J. Dunkerton, Fluxes of heat and constituents due to convectively unstable gravity waves, *J. Atmos. Sci.*, *42*, 549-556, 1985.

Fritts, D. C., and P. K. Rastogi, Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: theory and observations, *Radio Sci.*, *20*, 1247-1277, 1985.

Fritts, D. C., and R. A. Vincent, Mesospheric momentum flux studies at Adelaide, Australia: observations and a gravity wave-tidal interaction model, *J. Atmos. Sci.*, *44*, 605-619, 1987.

Fritts, D. C., T. Tsuda, T. Sato, S. Fukao, and S. Kato, Observational evidence of a saturated gravity wave spectrum in the troposphere and lower stratosphere, *J. Atmos. Sci.*, *45*, 1741-1759, 1988.

Fritts, D. C., and T. E. VanZandt, A theory of enhanced saturation of the gravity wave spectrum due to increases in atmospheric stability, *Pure Appl. Geophys.*, *130*, 399-420, 1989.

Gadsden. M., Noctilucent clouds, *Q. J. R. Astron. Soc.*, *27*, 351-366, 1986.

Garcia, R. R., Dynamics, radiation, and photochemistry in the mesosphere: implications for the formation of noctilucent clouds, *J. Geophys. Res.*, this issue, 1989.

Garcia, R. R., and S. Solomon, The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere, *J. Geophys. Res.*, *90*, 3850-3868, 1985.

Hauchecorne, A., M. L. Chanin, and R. Wilson, Mesospheric temperature inversion and gravity wave breaking, *Geophys. Res. Lett.*, *14*, 933-936, 1987.

Hines, C. O., Generation of turbulence by atmospheric gravity waves, *J. Atmos. Sci.*, *45*, 1269-1278, 1988.

Holliday, D., and M. E. McIntyre, On potential energy density in an incompressible, stratified fluid, *J. Fluid Mech.*, *107*, 221-225, 1981.

Holton, J. R., The influence of gravity wave breaking on the general circulation of the middle atmosphere, *J. Atmos. Sci.*, *40*, 2497-2507, 1983.

Holton, J. R., and M. R. Schoeberl, The role of gravity wave generated advection and diffusion in the transport of tracers in the mesosphere, *J. Geophys. Res.*, *93*, 11,075-11,082.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, On the use and significance of isentropic potential-vorticity maps, *Q. J. R. Meteorol. Soc.*, *111*, 877-946 (and *113*, 402-404), 1985.

Jensen, E. J., Thomas, G. E., and Toon, O. B., On the diurnal variation of noctilucent clouds. *J. Geophys. Res.*, this issue, 1989.

Koop, C. G., and B. McGee, Measurements of internal gravity waves in a continuously stratified shear flow, *J. Fluid Mech.*, *172*, 453-480, 1986.

Kopp, E., F. Bertin, L. G. Björn, P. H. G. Dickinson, C. R. Philbrick, and G. Witt., The 'CAMP' Campaign 1982, Proceedings of the 7th Symposium on European Rocket and Balloon Programmes and Related Research, ESA SP-229, European Space Agency, Paris, July 1985.

Linden, P. F., J. M. R. Redondo, and C. P. Caulfield, Molecular mixing in Rayleigh-Taylor instability, in *Proceedings of Congress on Compressible Turbulent Mixing*, edited by S. Orszag, in press, Springer, New York, 1989.

Lindzen, R. S., Turbulence and stress owing to gravity wave amd tidal breakdown, *J. Geophys. Res.*, *86*, 9707-9714, 1981.

Manson, A. H., and C. E. Meek, Gravity wave propagation characteristics (60-120km) as determined by the Saskatoon MF Radar (Gravnet) system: 1983-85 at 52°N, 107°W, *J. Atmos. Sci.*, *45*, 932-946, 1988.

McIntyre, M. E., Dynamics and tracer transport in the middle atmosphere: An overview of some recent developments, in *Transport Processes in the Middle Atmosphere,* edited by G. Visconti and R. R. Garcia, pp. 267-296, D. Reidel, Hingham, Mass., 1987.

McIntyre, M. E., A note on the divergence effect and the Lagrangian-mean surface elevation in water waves, *J. Fluid Mech*, *189*, 235-242, 1988.

Olivero, J. J., and G. E. Thomas, Climatology of polar mesospheric clouds, *J. Atmos. Sci., 43,* 1263-1274, 1986.

Plumb, R. A. and J. D. Mahlman, The zonally averaged transport characteristics of the GFDL general circulation/transport model, J. Atr^^s. Sci., *J. Atmos. Sci., 44,* 298-327, 1987.

Reid, G. C., Ice clouds at the summer polar mesopause, *J. Atmos. Sci., 32,* 523-535, 1975.

Reid, I. M., Gravity wave motions in the upper middle atmosphere (60-110km), *J. Atmos. Terr. Phys., 48,* 1057-1072, 1986.

Reid, I. M., and R. A. Vincent, Measurements of mesospheric gravity wave momentum fluxes and mean flow accelerations at Adelaide, Australia, *J. Atmos. Terr. Phys., 49,* 443-460, 1987.

Rottger, J., VHF radar measurements of small-scale and mesoscale dynamical processes in the middle atmosphere. in Studies of the Middle Atmosphere. *Phil. Trans. R. Soc. Lond. A 323,* 611-628, 1987.

Schoeberl, M.R., A model of stationary gravity wave breakdown with convective adjustment, *J. Atmos. Sci., 45,* 980-992, 1988.

Schröder, W., Some aspects of the study of noctilucent clouds and mesospheric circulation, *Idõjárás, 78,* 31-40, 1974.

Strobel, D. F., Constraints on gravity wave induced diffusion in the middle atmosphere, *Pure Appl. Geophys., 130,* 533-546, 1989.

Thomas, G. E., Olivero, J. J., Jensen, E. J., Schröder, W., Toon, O. B., Relation between increasing methane and the presence of ice clouds at the mesopause, *Nature, 338,* 490-492, 1989.

Thomas, R. J., C. A. Barth, and S. Solomon, Seasonal variations of ozone in the upper mesosphere and gravity waves, *Geophys. Res. Lett., 11,* 673-676, 1984.

Tung, K.K., On the two-dimensional transport of stratospheric trace gases in isentropic coordinates, *J. Atmos. Sci., 39,* 2330-2355, 1982.

Walterscheid, R. L., Inertio-gravity wave induced accelerations of mean flow having an imposed periodic component: Implications for tidal observations in the meteor region, *J. Geophys. Res., 86,* 9698-9706, 1981

Walterscheid, R. L., Gravity wave attenuation and the evolution of the mean state following wave breakdown, in *Dynamics of the Middle Atmosphere,* edited by J. R. Holton and T. Matsuno, pp. 19-43, D. Reidel, Hingham, Mass., 1984.

Weinstock, J., Superadiabatic excess and gravity wave saturation, *J. Atmos. Sci., 45,* 3519-3520, 1988.

Zimmerman, S. P., and T. J. Keneshea, Turbulent heating and transfer in the stratosphere and mesosphere, *J. Atmos. Terr. Phys., 48,* 491-507, 1986.

M. E. McIntyre, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, England.

(Received August 16, 1988;
revised December 15, 1988;
accepted February 3, 1989.)